वेदिन्-कक्ष (Vedinkaksha): A Computational Framework for a 'Sensitive' Blended Learning Platform

Thesis submitted in partial fulfilment of the requirements for the award of the degree of

Doctor of Philosophy

in

Computer Science and Engineering

by

Subrata Tikadar

Under the supervision of

Dr. Samit Bhattacharya

Department of Computer Science and Engineering

Indian Institute of Technology Guwahati

Guwahati - 781039, Assam, India

January, 2021

Dedicated to

My Parents and Grandparents

Acknowledgements

First and foremost, I would like to express my heartiest gratitude to my supervisor, *Dr. Samit Bhattacharya*. He is the first and most important person because of whom I have got familiarized with the idea of true research. His never-satisfied attitude but the silent inspiration for good works have always motivated me to do things better and better. In addition to the research ethics and guidelines, he had also provided me every facility and infrastructure which were required for conducting my research work in a smoother way. Not only in professional life, but ups and downs probably come in everyone's personal life as well. I will never forget the mental support he provided during my hard times in the tenure of my Ph.D. program.

I would like to thank all my doctoral committee members: *Prof. Rohit Sinha*, *Dr. T Venkatesh*, and *Dr. V Vijaya Saradhi* for their invaluable criticisms and suggestions which helped shaped the thesis toward the right direction with a greater clarity. I would also like to thank the HODs during the tenure of my Ph.D., Prof. Diganta Goswami, Prof. S V Rao, and Prof. J K Deka for allowing me to utilize the departmental resources including research equipment, and travel support for attending international conferences. I am also highly obliged to all the other faculty members for their help and support. I also thank the Directors, Deans, and all other administrative personnel, who have made IIT Guwahati a wonderful place for world-class research.

I am deeply thankful to all the technical staff of the Department of Computer Science and Engineering, Mr. Nanu Alan Kachari, Mr. Hemanta Kumar Nath, Mr. Bhriguraj Borah, Mr. Pranjitt Talukdar, Mr. Raktajit Pathak, and Mr. Nava Kumar Boro for extending their helping hands to solve all kinds of technical issues. I also thank Mr. Monojit Bhattacharjee, Mrs. Gouri Khuttiya Deori, and Mr. Prabin Bharali, who have smoothly handled all the official and administrative formalities. I also thank all the security personnel for their continuous help.

I would also like to thank the Ministry of Education (formerly MHRD), Govt. of India, for providing me the scholarship for pursuing the Ph.D. program. I am thankful to IFIP TC13 on HCI, and the IFIP Digital Equity Committee for the grant to attend one of the most prestigious international conferences, INTERACT 2019, in Paphos, Cyprus. I am also thankful to the IEEE Computer Society and the IEEE Technical Committee on Learning Technology for recognizing my research paper as one of the best papers in the 19th IEEE International Conference on Advanced Learning Technologies (ICALT 2019, Maceio, Brazil).

The heavenly blessings of my grandparents have always helped me to cross all the hurdles. I would like to cordially thank my parents, who are the two most important persons in my life. They have continuously inspired me to complete my thesis work, even without letting me know their deplorable homeless circumstances. I am grateful to my two younger sisters who have taken care of my parents during their hard times, and at the same time have inspired me in all the ways so that I can complete my Ph.D. without having to worry for my family.

I am fortunate to have many good friends: Anirban, Bikramjit, Debendra, Ganesh, Kangkan, Mahua, Madhurima, Pranav, Rathin, Shrestha, Subhrendu, Swagat, Ujjal, Zahir, who have helped and inspired me in many ways. I want to thank Pamela for inspiring me in all the ways without even letting me know her struggle and sacrifice for me, and for being my constant companion in every ups and downs of my life.

Lastly, but yet very importantly, I would like to acknowledge the continuous support of Dr. Ranjan Maity from the very beginning of my Ph.D. life. I am indeed grateful to all my fellow lab mates at the UCCN lab: Abdalganie, Hema, Nilotpal, Shakeel, and Ujjwal for maintaining a very good environment for work. The creative discussion, brainstorming, knowledge sharing, and spending many sleepless working nights together have helped me a lot for my gradual development as an independent researcher (possibly!). My special thanks go to *Hema*, *Nilotpal*, *Shakeel*, and *Sumantra* who have unconditionally helped me anytime, and many a time even at the cost of sacrificing their important works.

January 25, 2021 Subrata Tikadar

Declaration

I certify that

- The works contained in this thesis are original and have been done by myself and under the general supervision of my supervisor.
- The works reported herein have not been submitted to any other Institute for any degree or diploma.
- I have followed the guidelines provided by the Institute in preparing the thesis.
- Whenever I have used materials (concepts, ideas, text, expressions, data, graphs, diagrams, theoretical analysis, results, etc.) from other sources, I have given due cred it by citing them in the text of the thesis and giving their details in the references. Elaborate sentences used verbatim from published work have been clearly identified and quoted.
- I also affirm that no part of this thesis can be considered plagiarism to the best of my knowledge and understanding and take complete responsibility if any complaint arises.

January 25, 2021 Subrata Tikadar

Department of Computer Science and Engineering
Indian Institute of Technology Guwahati
Guwahati - 781039 Assam India

Dr. Samit Bhattacharya
Associate Professor
Email : samit@iitg.ac.in
Phone : +91-361-2582362

Certificate

This is to certify that this thesis entitled "वेदिन्-कक्ष (**Vedinkaksha): A Computational Framework for a 'Sensitive' Blended Learning Platform**" submitted by **Subrata Tikadar**, in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy, to the Indian Institute of Technology Guwahati, Assam, India, is a record of the bonafide research work carried out by him under my guidance and supervision at the Department of Computer Science and Engineering, Indian Institute of Technology Guwahati, Assam, India. To the best of my knowledge, no part of the work reported in this thesis has been presented for the award of any degree at any other institution.

Date: January 25, 2021
Place: Guwahati

Dr. Samit Bhattacharya
(Thesis Supervisor)

Abstract

Advanced technologies are being utilized in the field of education for an improved teaching-learning experience. Modern technologies and equipment such as smart-board, projector, and sound amplifier are being adapted in the traditional chalkboard-duster classroom system. With the advent of Information and Communication Technology (ICT), a new type of learning system, called e-learning, has also been popularized. In order to achieve the benefits of both the face-to-face classroom and e-learning systems, recently scholars have started to blend e-learning with the traditional classroom system, which is termed as a blended learning platform. Although the use of modern technologies and devices in the educational setting seems to lead to an improvement in the teaching-learning experience and outcome to a significant extent, the mere use of them does not ensure fulfillment of all the important requirements for effective teaching and learning. One such important requirement is the detection of the mental states of the students. Consideration of mental states in a classroom is very important as it directly and/or indirectly affect many critical aspects of students' present and future life such as school dropout rate, learning progress, thinking capability, and future career expectations.

However, it is challenging for a teacher to detect these states in real-time, especially when the classroom is large (having a hundred to a few hundred students). In this thesis, we propose a novel computational framework for a sensitive classroom system, which is able to automatically sense (or detect) the mental states of the students based on their emotions, involvements, and levels of classroom activities, and at the same time take some actions as per the detected states. Additionally, the system is also able to quantify the learning and understanding levels of the students on the learning platform. We have assumed a blended learning platform in the Bring-Your-Own-Device (BYOD) paradigm as the basis of the framework.

The proposed framework incorporates four novel machine learning-based computational models, which are the key research contributions of the thesis. We have built three models namely, the Touch-Affect, the Type-Affect, and the Smart-Affect,

which detect the affective states of the students from basic user-smartphone interaction data, i.e., users' touch and typing patterns and dynamics. We validated the models with the EEG data of the students. The adaptability of the models in a blended learning platform has also been tested by using the EEG data. The fourth computational model, called the In-Activity, detects the involvement of the students in study-related activities in the blended learning platform. We utilized these four models as the key components of our proposed framework for a sensitive blended learning platform.

In order to validate the proposed framework, we have built a high-fidelity working prototype for a sensitive blended learning system. We have observed that the system is able to detect the mental states of the students and take some actions based on the identified states, such as alerting the students and teacher and presenting a visualization dashboard to the teacher to perceive the overall status of the classroom based on the detected states. Based on our findings, we believe that a sensitive classroom system built following our proposed framework will help in motivating the students and teacher for greater engagement in the teaching-learning process, and in improving the learning outcome.

Contents

List of Figures

List of Tables

List of Symbols

Symbol	Description
$\surd$	Applicable
$\times$	Not Applicable
α	Significance level for statistical test
Σ	Summation
Ω_X	Rotation speed along X-axis
Ω_Y	Rotation speed along Y-axis
Ω_Z	Rotation speed along Z-axis
$\overrightarrow{ACC_X}$	Acceleration along X-axis
$\overrightarrow{ACC_Y}$	Acceleration along Y-axis
$\overrightarrow{ACC_Z}$	Acceleration along Z-axis
A_{tot}	Sum of the acceleration for a time interval
$A(R)_{tot}$	Sum of the raw acceleration for a time interval
C	The number of effectively typed characters in an interval
CE	Controlled Experiment
C_b	Backspace Key
C_e	Erased character
C_{Sp}	Special character
D_i	Duration of the i^{th} event
$Dth+$	Acceleration above a threshold value
E_i	The number of events generated during the i^{th} gesture
$H1_0$	Null Hypothesis-1
$H2_0$	Null Hypothesis-2
$H3_0$	Null Hypothesis-3
$H4_0$	Null Hypothesis-4
KC	Key Component

len	Length
max	Maximum
M_{avg}	Average memory utilization for a time interval
NS	Non-study related smartphone activities
$O(n)$	Computational complexity is Order of n
$O(n^2)$	Computational complexity is Order of n^2
p	Calculated probability of null hypothesis to be true in statistical test
ppi	Pixels per inch
P_i	Pressure generated at the i^{th} event
R_{tot}	Sum of rotation speed in a time interval
s	Second
S	Study related smartphone activities
$substr$	Sub-string
$S1$	Mental State-1 (Not Engaged)
$S2$	Mental State-2 (Ideal State)
$S3$	Mental State-3 (Frustrated)
$S4$	Mental State-4 (Feeling Good)
$S5$	Mental State-5 (Getting no Interest)
$S6$	Mental State-6 (Showing Off)
$S7$	Mental State-7 (Unable to Understand)
$S8$	Mental State-8 (Understanding but Lazy)
$S9$	Mental State-9 (Feeling Shy)
SD	Standard Deviation (sample)
T_{avg}	Average battery temperature for a time interval
$Vs1$	Visualization State 1 (Engaged and Understanding)
$Vs2$	Visualization State 2 (Engaged but Not Understanding)
$Vs3$	Visualization State 3 (Not Engaged but Understanding - Showoff)
$Vs4$	Visualization State 4 (Not Engaged and Not Understanding)
w	Time window
$x(a)$	Scalar component of the acceleration along X-axis
$x(g)$	Scalar component of the rotation speed along X-axis
$y(a)$	Scalar component of the acceleration along Y-axis

$y(g)$	Scalar component of the rotation speed along Y-axis
$z(a)$	Scalar component of the acceleration along Z-axis
$z(g)$	Scalar component of the rotation speed along Z-axis

List of Acronyms

Acronym	**Expansion**
AI	Artificial Intelligence
API	Application Programming Interface
app	Smartphone application
BYOD	Bring-Your-Own-Device
CV	Cross-Validation
DFD	Data Flow Diagram
DSSQ	Dundee Stress State Questionnaire
DT	Decision Tree
EBR	Eye Blinking Rate
ECG	Electrocardiogram
EEG	Electroencephalogram
EMG	Electromyography
ESM	Experience Sampling Method
f2f	face-to-face
GB	Giga Byte
GEW	Geneva Emotional Wheel
GSR	Galvanic Skin Response
HCI	Human-Computer Interaction
HOD	Head of the Department
HR	Heart Rate
IAPS	International Affective Picture System
ICT	Information and Communications Technologies
ID	Identity Number
IIT	Indian Institute of Technology
IMS	Instant Messaging Services

ITS	Intelligent Tutoring System
KNN	K-Nearest Neighbour
LOSOCV	Leave One-Subject-Out Cross-Validation
MCQ	Multiple Choice Questions
ML	Machine Learning
OTP	One-Time Password
OS	Operating System
PC	Personal Computer
PCA	Principle Component Analysis
PDA	Personal Digital Assistance
pdf	Portable Document Format
PG	Postgraduate
PhD	Doctor of Philosophy
ppt	Power Point Presentation
QR code	Quick Response Code
RAM	Random-Access Memory
RF	Random Forest
RNN	Recurrent Neural Network
SAM	Self-Assessment Manikin
SMS	Short Message Service
SPS	Sample per Second
SVM	Support Vector Machine
UG	Undergraduate
UK	United Kingdom
URL	Uniform Resource Locator
USD	United States Dollar
Wi-Fi	Wireless Fidelity
WSA	Work Sample Analysis

Glossary of Terms

Term	Description
Affective States	The psychophysiological constructs of the states of affect, emotion, mood, and personality traits.
Affect Induction	The process of eliciting affect and emotion (primarily in a laboratory setting).
Alert	To warn or notify someone of a specific situation.
Backchannel Data	The data related to users' interaction behavior, which are generated during online discussions and/or conversations.
Behavioral Model	Computational model that takes users' behavioral data as inputs.
Blended Learning Platform	A learning platform or system where the e-learning is blended with the traditional face-to-face learning to achieve the benefits of both.
BYOD Paradigm	A client-server based environment where the client nodes are the personal devices of the users.

Circumplex Model of Emotion	The continuous emotions are discretized on a two-dimensional circular space based on the arousal and valence.
Cognitive State	State based on the cognitive load, which primarily involves thinking and reasoning.
Computational Framework	A basic structure underlying a system or concept, which depends on computational processes.
Computational Model	A simplified description, especially a mathematical one, of a system or process, to assist computations and predictions.
Computer Vision	An artificial intelligence-based interdisciplinary scientific field that deals with computer's understanding from digital images and/or videos.
Controlled Experiment (CE)	Experiment which is conducted in a laboratory setting to control the latent issues in the data.
Cross-Validation	A technique to validate machine learning-based computational models on a limited sample of data.
Demographic Data	The socio-economic status and/or information, which are generally expressed statistically.
E-learning	Learning through an electronic medium, typically on the Internet.
Embedded Sensors	Sensors come as integrated parts of the computing devices, which may also incorporate actuators, to sense and/or interact with the real-world environment.

Empirical Study	The study based on observation and measured phenomena, which derives the knowledge from experience rather than from theory or belief.
Environmental Factors	The confound and control variables which may affect the dependent variable. Examples include light, temperature, and population density.
Involvement	Students' involvement in study-related activities in the classroom.
Likert Scale	A bipolar or even-point scaling method to measure either positive or negative response from the users for a statement without any option for the 'neither agree nor disagree' choice.
Machine Learning	An application of AI where the computers gain the ability to automatically learn even when we do not explicitly program them.
Mental States	Psychophysiological states based on emotion, involvement, and activity.
Multimodal Technique	More than one input modalities are used to compute some events or states.
Neurological Signal	The signals generated from the nervous system.
Outlier	A person or thing (data-point in the current context) that differs from all other members of a particular set.
Pearson's Chi-Square Test	A statistical test that is used to assess the goodness of fit, homogeneity, and independence comparisons; applied especially when the experiment is designed for the nominal (categorical) scale of measurement.

Peripheral Device — An ancillary device that is not part of the essential computing system but used to provide information into and get access to the information out of the computer.

Physiological Signal — The signals which are generated due to the physiological process of human beings. Examples include EBR, ECG, EEG, EMG, and HR.

Privacy-Sensitive Data — Private data which is very sensitive, processing of which may affect our privacy to a significant extent.

Sensitive Platform — A system or platform which is able to automatically sense or detect the mental states of its users.

Sensor — A device (or subsystem) that is able to detect or sense the events or environmental changes surrounding it.

Smartphone Activity — The activities which are performed using smartphones.

Survey — A research method to collect data from a specific group of users to gain specific knowledge on a topic of interest.

Survey Questionnaire — A set of questions used in a survey form.

System Architecture — A conceptual model or blueprint of a system that defines the structure, behavior, abstract view, and specification of it.

t-Test — An inferential statistical test to observe the significant difference between the means of two groups. Generally applied when the dependent variable fits a categorical (normal) distribution.

Teaching-Learning Process	Together the learning and teaching process..
User-Centric Computing	Computing that considers users' concerns in a ubiquitous computing environment.
Visualization	A technique to create visual imagery for an abstract and/or concrete idea.

“All learning has an emotional base.”

Plato (428/427 - 348/347 BC)
Athenian philosopher

1

Introduction

When we think of a classroom, the very first picture that comes to our mind is a traditional chalkboard-duster classroom, where the students and teacher participate in the teaching-learning process face-to-face (f2f). In the present days' f2f classrooms, modern technologies and equipment such as projector, smartboard, and sound amplifier are being adapted for improved teaching and learning experiences. The importance of the conventional f2f classroom is undeniable as there is no alternative to the invaluable guidance, care, and help of the teacher in real-time. Other benefits of the traditional classroom teaching are the practice of being disciplined and social, quick learning, peer assistance and knowledge sharing, and accomplishing the curriculum within the specified period of time in a structured way with the proper supervision of the course instructor. However, traditional classroom systems have the following limitations.

- **Difficulties in managing lecture contents:** It may be difficult for the instructor to present appropriate and sufficient learning materials because of limited time and classroom infrastructure. Moreover, many a time, this type of classroom systems fails to provide global learning resources. Learning contents and outcomes may be limited to and/or dependent on the knowledge of the teacher.

- **Requirement of huge effort:** Delivering the lecture in an efficient way, in terms of presentation and flow, requires a lot of effort from the teachers' side. Teachers are required to perform live – any weakness of the teachers may cause serious issues as there is no scope for a 'retake'. It is cumbersome for a teacher to find and rectify wrongly delivered contents, because lectures are delivered in real-time.

- **Limited number of students:** Only a limited number of students can take part in the classroom at a time. An increase in the number of students requires increasing resources in terms of additional teachers, more number of classrooms, and suitable classroom infrastructures.

- **Requirement of physical presence:** Students as well as the teacher are required to be physically present for taking part in f2f classroom activities. The physical presence in the classroom as per a predefined schedule at a particular venue may sometimes be problematic for both the students and teachers.

Another learning platform called 'e-learning' has been conceptually introduced in the 1960s [95], which has now been popularized because of the vast progress of Information and Communication Technology (ICT). In the e-learning platform, the students can learn anytime anywhere, which is the main benefit of this learning system. As an additional benefit, the students can access huge and global learning content in a better way. In the case of e-learning, it is easy to get access and manage enormous learning resources through the Internet, which may not be possible in conventional f2f learning. A lot of students can take part in the same time in the e-learning platform. Once the infrastructure set-up is ready, the execution cost reduces significantly. At the same time, it is also beneficial for the teacher because once the lecture material is ready, it can be reused with a negligible effort.

However, availing many of the benefits of the traditional classroom learning is difficult to obtain in the e-learning platform. As the teacher does not need to be physically present, the students may miss the active guidance, care, and help of the teacher. Moreover, as the students are not bound to take part in the learning activities at a scheduled time and venue, in this case, they may not avail of the other benefits of the f2f learning such as peer assistance and knowledge sharing. There is also a possibility of turning indiscipline and unsocial, because of the same reason.

Also, because of the lack of active supervision of the instructor students may fail to accomplish the course on time. A new teaching-learning system is therefore being introduced by the researchers to avail the benefits of both the long-established f2f classroom and e-learning system, which is termed as the *blended learning platform*.

1.1 Blended Learning Platform

We can find many definitions of blended learning, which is also known as *"hybrid learning"* or *"mixed-mode learning"* [60]. According to a classical definition by Osguthorpe and Graham [98], blended learning is the *"combination of f2f and computer-supported instruction"*. The phrase has been elaborately redefined by Hoic-Bozic et al. [59] as *"learning based on various combinations of classical face-to-face lectures, learning over the Internet, and learning supported by other technologies, aimed at creating the most efficient learning environment"*. Conceptually, blended learning may incorporate various teaching and learning methods, diverse modes of learning, and different interaction schemes. Utilizing modern technologies, different learning methods such as e-learning and traditional f2f learning may be combined in this platform. We can also think of blending various modes of learning, such as 'individual learning' where a private tutor (human or machine) takes additional care of a student in one-to-one leaning mode, and 'group learning' where students learn in groups. We may also think of incorporating and/or mixing different classroom interaction schemes such as 'synchronous interaction scheme' where both the teacher and students get involved in the interaction in real-time, and 'asynchronous interaction scheme' where the live interactions do not take place but the students can raise queries which can be addressed later by the teacher as per her/his convenience. The primary focus of the blended learning is to choose an appropriate mixture of these methods (i.e., online and traditional), modes (i.e., group and individual), and schemes (i.e., synchronous and asynchronous) to design a learning platform, which improve pedagogy in terms of learning experiences and outcomes [2] [43] [46] [47] [96] [130]. Figure 1.1 depicts the concept of a blended learning platform.

According to Al-Qahtani and Higgins [2], blended learning is more advanta-

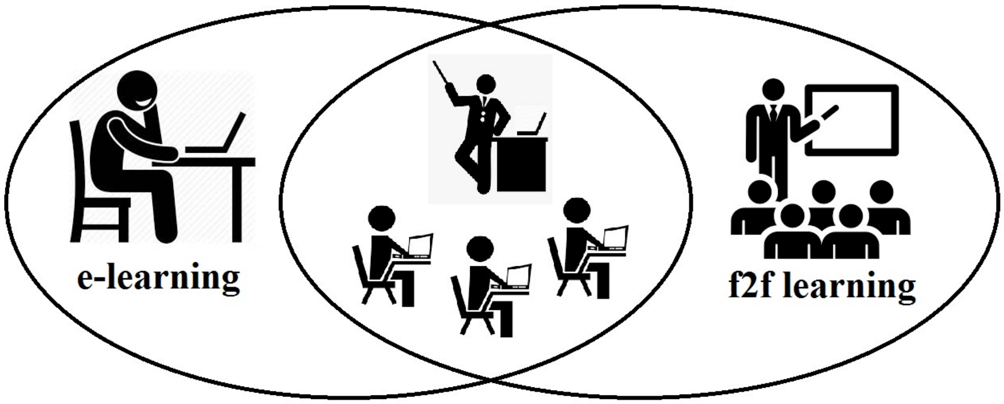


Figure 1.1: Concept of a blended learning platform

geous over e-learning and f2f learning in terms of students' achievements in terms of learning outcomes. Gitinabard et al. [43] have also found that students earn more grades in this learning platform compared to the conventional f2f and e-learning system. The research findings of Nortvig et al. [96] also indicates the efficiency of blended learning in terms of learning outcome, students' satisfaction, and their engagement in the learning process. Tikadar et al. [130] also found the efficiency of a blended learning platform in terms of classroom interactions, acceptability, and learning outcomes.

We may have many ways to design and implement a blended learning platform (based on they are defined). One such way is to utilize the Bring-Your-Own-Device (BYOD) paradigm of client-server architecture in the physical classroom. The BYOD paradigm concept, along with its benefits in the current context, is described next.

1.2 Bring-Your-Own-Device (BYOD) Paradigm

In order to exploit of the facility of ICT in the traditional learning system, a suitable classroom infrastructure is required. In most of the cases, a client-server architecture is required to be implemented to facilitate the environment of e-learning in f2f learning system [59] [130]. This architecture requires a local server, a communication medium, and a set of client nodes as depicted in Figure 1.2. The teacher and students

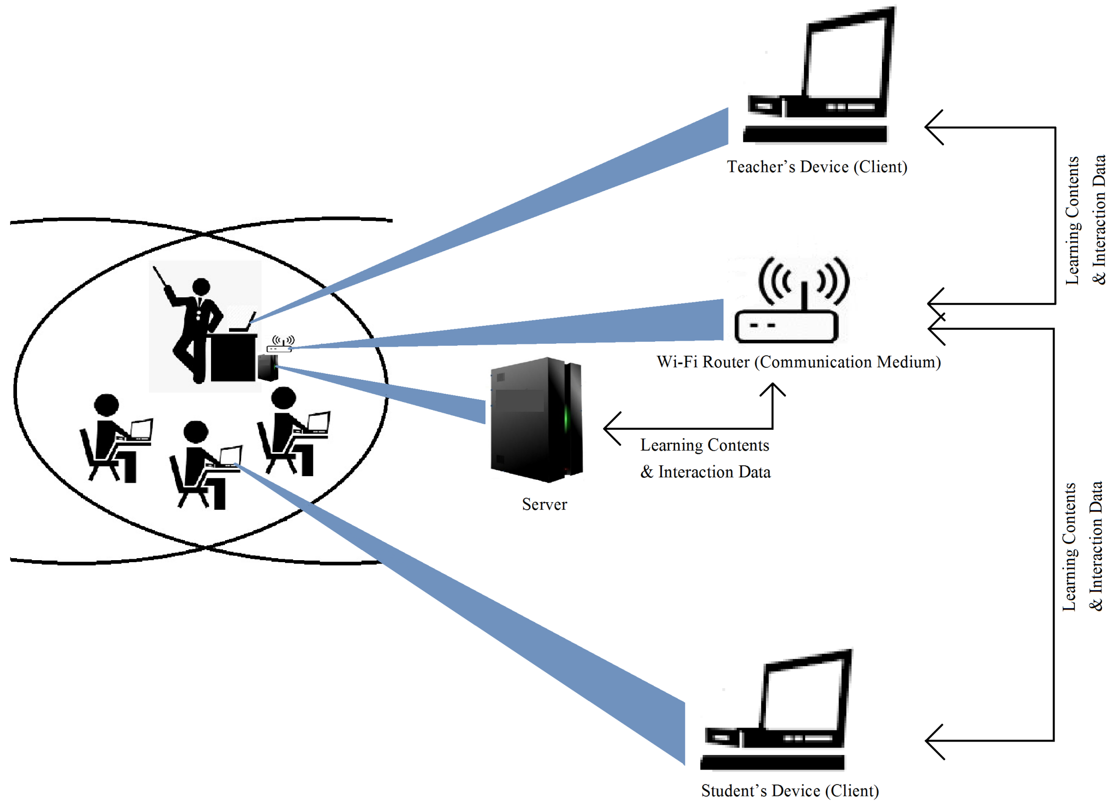


Figure 1.2: Client-server architecture of a blended learning platform

perform the classroom activities through the client nodes. The local server takes the responsibility to supply the learning materials and interfaces for the classroom interactions. It also takes the role of heavy and analytical computations on behalf of the clients, and securely stores the information and supplies them to the authentic users (client nodes) as per requirements and policy designed by administrators. It also may connect the clients to the global network, if required. All the nodes can be interconnected through a secure and reliable communication medium, e.g., via a Wi-Fi router. The client nodes may be acquired in either of the following two ways: (a) institute can arrange a fixed set of computing devices for the students and teacher, and (b) the students and teacher can bring their personal device (such as laptops or smartphones) and use them for participating in the teaching-learning process. The latter scenario, where the server and network infrastructures are provided by the institute but the client nodes are the personal computing devices of the users, is termed as the 'Bring-Your-Own-Device' (BYOD) paradigm. Although the BYOD paradigm was originally introduced for interacting with a large public display using ubiquitous personal devices [10], the same is now being adapted for many interactive systems and applications.

Therefore, it would seem prudent to embrace the BYOD paradigm for blended learning, considering its multitude of benefits [10] [34] which are listed below.

- **Portability:** Since a system under the BYOD paradigm does not require a fixed set of client devices but only a server and a portable router, its overall portability is higher.

- **Multipurpose usage of devices:** The client devices in this paradigm are not dedicated to a particular system. The devices are used for different applications and systems as per the preference of the users who own them.

- **Hygiene:** As the client devices are the personal belongings of the users, generally no one except their owners is supposed to touch the devices. Therefore, it is easy to maintain hygiene. On contrary, if a set of common devices are used by multiple users of the system, it is cumbersome to track the usage history and therefore to maintain sanitation.

- **Multiuser simultaneity:** Theoretically, there is no concept of 'additional users' in the BYOD paradigm as the number of users is always dynamic in this case. This results in multiuser simultaneity.

- **Security:** Since there is no device sharing among the users, the security of credentials and data are ensured.

- **User satisfaction and acceptability:** Users are found to be more inclined to use these systems because of all the above benefits.

- **Less operation cost:** Lastly, but yet most importantly, it drastically reduces the institute's device purchase and maintenance cost as it is not required to provide any devices to the students and teacher. The explosion of the smartphone as a ubiquitous computing device in recent times has made it easy to adapt this paradigm.

1.3 Motivation

The use of modern technologies and devices in the educational setting improve the teaching-learning experiences and learning outcomes to a significant extent. However, the mere use of the technologies and equipment does not assure fulfillment of all the important requirements for effective teaching and learning. One such important requirement is the detection of mental states of the students. Mental states, here, refer to the psychophysiological states of the students encompassing their emotions (affective states), involvements in the teaching-learning process, and the levels of performing classroom activities [6] [7] [35] [50] [62] [108] [139]. Consideration of these states in a classroom is very important as it directly and/or indirectly affects many critical aspects of students' present and future life [102] [118]. For example, lack of involvement may affect the rate of school dropouts, academic progress, and future career expectations of the students [14] [118] [102]. Similarly, Negative affect can influence the memory [83], and impair the thinking capability of a student [140]. On the other hand, positive affect helps the students to understand lessons efficiently [83]. Identifying the mental states (which comprising of affect & cognition, involvement, and activity), and at the same time, complementing the states upon requirement is therefore imperative. However, it is challenging for a teacher to detect these states in real-time, especially when the classroom is large (having more than fifty to a few hundreds of students).

Existing ways for detecting the mental states are *self-report*, *manual observation* (by the teacher or external observers), and *systematic observation* (sometimes termed as *automated measurement*) methods [102] [138]. Although the self-report based method has been found to be effective in many cases, it is afflicted with the problem of *imitation* as students may be hesitant to report their actual states, particularly when the process is not anonymous [6] [40] [138]. Moreover, declarations by the students in the self-report based method may be biased because of the *recency and primacy memory effects* – students may forget or unable to express their exact states in sequence [141]. Furthermore, conducting the survey for getting the reports and analyzing them to detect the states in real-time is very difficult [102] [138]. The observation method can also be considered as effective, but it is cumbersome and lacks temporal resolution [138]. In other words, if the teacher themself gets busy for observing the states, s/he might fail to deliver the lecture effectively and timely. In addition to the teacher, external observers might be employed for the observation. However, it may add to the monetary expenses (the number of observers increases with the increase in the number of students in the classroom, and hence their appointment costs). Moreover, the accuracy of detecting the states through external observers may be inconsistent as all the observers may not have the same level of expertise in the task [88]. Researchers have therefore proposed automated measurement methods, which consist of physiological and/or neurological signal-based models, and computer vision-based models [138] [141].

Although the models are able to detect the states consistently and with high accuracy, they are expensive (in terms of both computational as well as monetary cost). Moreover, most of the time, the models are not fit for the mobile environment – each student is required to sit at a designated position with a fixed setup. External factors (viz. lighting condition, positions and postures of the students) may also affect the accuracy of the models, particularly in the case of computer vision-based model. The physiological/neurological signal-based models involve external sensors and equipment, which causes a lack of acceptance of the models in practice. Students may not agree to wear the sensors and probes for attending the class on a regular basis. Furthermore, these models access privacy-sensitive data of the students in many cases, which may not be ethical.

1.4 Objective

The overall objective of this thesis is to present a computational framework for a ‘sensitive’ blended learning platform, which helps in detecting the mental states of the students considering their emotion, involvement, and level of classroom activities. We have chosen a blended learning system in the BYOD paradigm, called “Avabodhaka” [24] [130], as the basic platform behind the framework. In Avabodhaka, all the classroom activities including the lecture delivery, student-teacher interaction, query and response, peer-discussion, and the exam and quiz conduction are performed using mobile computers such as laptops, tablets, and smartphones. However, in our proposed framework, we have assumed only smartphones and tablets as client devices. Laptops are not considered as they do not come with the embedded sensors which are required to collect input data for the sensitive modules of our proposed framework. The blended learning platform is chosen because of its superior features as compared to the other learning platforms, in terms of teaching and learning experiences as well as learning outcomes [2] [43] [46] [47] [96]. Usages of smartphones for performing the classroom activities in the blended learning platform will help to make the environment mobile and ubiquitous. The efficiency and acceptance of this particular blended learning platform have also been tested by us through a controlled experiment [130]. The BYOD paradigm has been employed as it will help to reduce the cost of classroom infrastructure and increase the portability of the system. It would not be problematic for the students to attend lectures using their smartphones as nowadays almost every student carries at least one smartphone [61]. We have chosen the particular blended learning platform, with smartphones and tablets as client devices, for another important reason: we can utilize the sensory, interaction, and resource utilization data to detect the mental states of the students, without requiring any external sensors and equipment.

To achieve the overall objective, we have defined four sub-goals, reaching which leads to accomplishing the overall objective. The sub-goals are as follows.

(a) Systematic and real-time detection of affective states of the students for the assumed blended learning platform without using any additional sensors and equipment.

(b) Systematic and real-time detection of involvement of the students in the teaching-learning process for the assumed blended learning platform.

(c) Systematic and real-time detection of the level of classroom activities performed by the students in the assumed blended learning platform.

(d) Building and validating a computational framework to detect the mental states of the students based on their affective states, involvement in the teaching-learning process, and the level of participation in the classroom activities.

1.5 Contributions

We have made the following contributions, which essentially lead to the final objective of this thesis.

1.5.1 System Assessment

We have examined the efficiency of the assumed blended learning platform in terms of classroom interaction and learning outcomes through a controlled experiment [130]. The acceptance of the particular classroom has also been tested by us through a self-report based method. It has been found by us that the assumed blended learning platform is more efficient in terms of classroom interaction and learning outcome, compared to the traditional f2f classroom system and an f2f classroom system having a smart assistive device for classroom interaction. We have also found that the students are more inclined to accept the assumed blended learning platform compared to the other two types of classroom systems.

1.5.2 Framework for a Sensitive Classroom System

In order to make the assumed blended learning platform sensitive, we have built a computational framework for the same, which was the overall objective of the thesis. A system implemented following the proposed framework can sense the students' mental states based on their involvements, emotions, and levels of classroom activities. To reach this overall objective, we have made many other research contributions that include conducting a survey to be aware of students' behavior on performing smartphone activities, building and validating four computational models to detect their affective states and involvements in the assumed learning platform, defining the level of classroom activities in the learning system, and integrating the computational models to the proposed framework to detect the mental states as well as taking some actions based on the detected states to improve the pedagogy. These contributions are specified next in an individual manner. Figure 1.3 presents a thesis flow diagram that might help understand all our contributions more precisely.

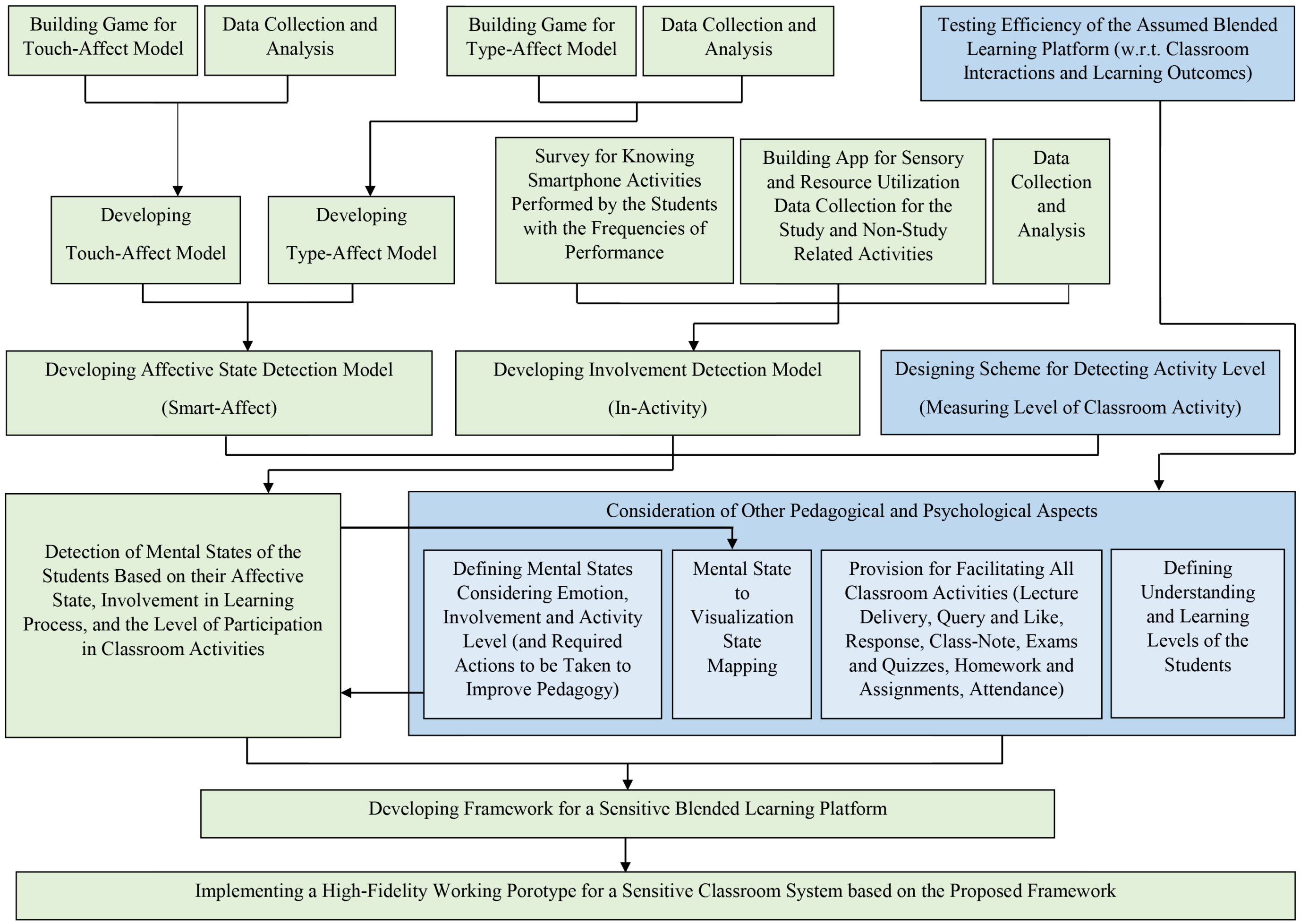


Figure 1.3: Thesis flow diagram

1.5.3 Survey on Smartphone Usage Behavior of Indian Students

We have conducted a survey throughout India to know the behavior of Indian students on smartphone usage [125]. An in-depth analysis of the survey results have helped us to become familiarized with the name and frequency of performing the study and non-study related smartphone activities by the students, both inside and outside the classroom, which upgrade the knowledge on smartphone usage by them. This state-of-the-art knowledge was required to design the proposed framework for the sensitive classroom system, particularly to develop a model to detect the involvement of the students in the assumed blended learning platform.

1.5.4 Affective State Detection Models

We have built and validated novel models to detect the affective states of smartphone users from basic user-smartphone interaction data, which are suitable for the assumed blended learning platform. Three models have been built and validated by us, which are as follows.

1. ***Touch-Affect*** – This is a novel computational model to detect the affective states of the users from their touch pattern on the touchscreen of a smartphone [131]. Taking the number of touch events created by the user per unit of time, and the average pressure generated for the events as inputs, the model detects the affective states of the users into either of the following four states: 'positive-high', 'positive-low', 'negative-high', and 'negative-low'. Here, 'positive' and 'negative'refer to the valence levels, and 'high'and 'low' indicate the levels of arousal.

2. ***Type-Affect*** – This is a novel computational model to detect the affective states of the users from their typing pattern on the virtual keyboard of a smartphone [126]. This model classifies a user's state into one of the four affective states, using the following four features while they type on the virtual keyboard of a smartphone: 'typing speed', 'touch count', 'maximum text length' and 'device shake frequency'.

3. ***Smart-Affect*** – This is a process model (based on Touch-Affect and Type-Affect) to detect the affective states of the students in the assumed blended

learning platform from the basic user-smartphone interaction behavior (i.e., touch and typing behavior) [129]. Based on the availability of input data, the process model chooses either of the Touch-Affect and Type-Affect models to compute the affective states of the users for a particular interval of time. If no input data is available for a certain interval, it shows the last identified state.

1.5.5 Involvement Detection Model

We have built and validated a computational model to detect the involvement of the students in a blended learning platform [127] [128]. By observing the device handling and resource utilization patterns, the model, named *In-Activity*, is able to specify whether a student is involved in study-related activities or performing something else which is not related to study. In other words, it detects students' involvement in study-related activities.

1.5.6 Implementation of a Sensitive Classroom

We have implemented a high-fidelity working prototype of a sensitive classroom system as per our proposed framework for the same. We incorporated the models in the system to characterize the mental state of the students into one of the nine states namely, *unengaged*, *ideal*, *frustrated*, *good*, *uninterested*, *show-off*, *unable to understand*, *lazy*, and *shy*. There are also provisions for computing understanding and learning levels, sending alerts to the students and teacher, and offering a visualized classroom to the teacher based on the detected states.

1.6 Thesis Outline

The thesis has been organized into seven chapters including the current chapter as **Chapter 1**, entitled "Introduction", which briefly introduces the problems tackled in the thesis, along with a description of the basic terminologies to understand the overall work. The overview of the remaining six chapters is as follows.

Chapter 2, entitled "Background", presents the existing works which are directly and/or indirectly related to the proposed framework for a sensitive classroom system. This chapter reports the existing sensitive classroom systems, and the works related to the components of the same, with their critical analyses to establish the novelty of the proposed framework.

Chapter 3, entitled "Vedinkaksha: A Framework for a Sensitive Classroom", describes the concept of the sensitive classroom in detail. This chapter elaborates on the entire process and working principle of the sensitive classroom, along with the details of its each component.

Chapter 4, entitled "Behavioral Study on Smartphone Usage by Indian Students", presents a survey on smartphone usage behavior by Indian students. This chapter describes the importance and methodology of the survey along with its results related to the thesis work.

Chapter 5, entitled "Empirical Study Details", presents the details of all the empirical studies conducted to build and validate the computation models, which are the major components of the proposed framework for the sensitive system.

Chapter 6, entitled "Computational Models", reports all the computational models involved in the Vedinkaksha framework. This chapter contains a detailed description of each of the computational models, which have been built and validated by us, and essential for the proposed framework.

Chapter 7, entitled "Conclusion and Future Scope", concludes the thesis by presenting a working prototype of a sensitive classroom system that has been built following the Vedinkaksha framework, along with the discussion on avenues for future research based on the current work.

"It is true that contemporary technology permits decentralization, it also permits centralization. It depends on how you use the technology."

Noam Chomsky (1928)
American linguist

2

Background

2.1 Introduction

There are few works in the literature, which are directly and or indirectly related to the proposed framework for a sensitive blended learning platform. The directly related works refer to those research works which have been conducted for building the systems and frameworks for a sensitive learning platform, which are able to detect the mental states of the students. Whereas, the indirectly related works indicate the studies which have been conducted not for building any such framework or system but related to the key components of the framework based on which the mental states are defined, i.e., the various models to detect the affective state, involvement, and activity level of the students. In this chapter, we will discuss the existing related works to critically analyze them for establishing the novelty and importance of our work. This discussion will also help us to understand why the existing works are not suitable to be readily adapted to the proposed framework.

2.2 Mental State Detection Methods

The existing works for detecting the mental states of the students (or their components) can broadly be categorized into three groups based on the way of detection: self-report, manual observation (by the teacher or external observers), and systematic observation (sometimes termed as automated measurement) methods [102]

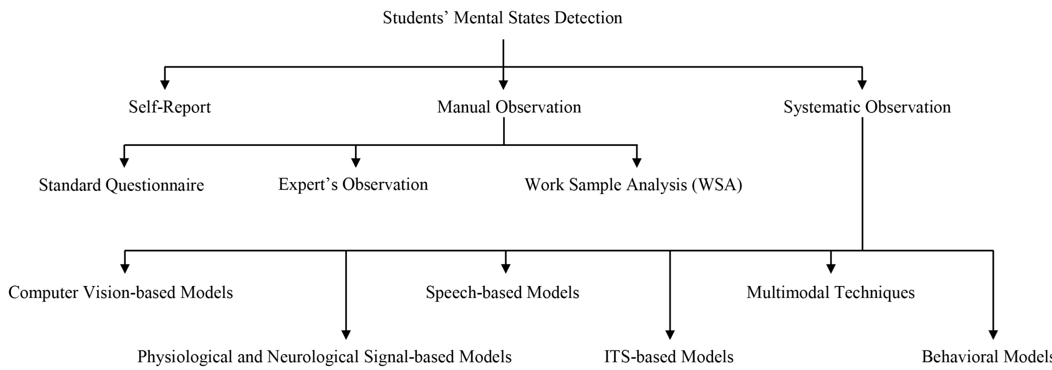


Figure 2.1: Hierarchical diagram for mental state detection methods

[138]. The groups and subgroups of the mental state detection methods are summarized in Figure 2.1 in the form of a hierarchical diagram and discussed in the following subsections.

2.2.1 Self-report

In this method, students are asked to report their states through some standard questionnaires. Most of the time, the students are asked to report the states indirectly. The students report on the factors, which are analyzed to define the states. The factors, here, refer to the affective, behavioral, and cognitive aspects, together with which define the mental states of the students. The factors include distraction, attention, amount of mental effort spent, task persistence, response time and level of task completion, and emotional reaction to and interest in learning activities [22] [102] [138]. Examples of self-report based studies to detect the mental state include [73], [86], and [118]. Larson and Richards [73] conducted a self-report based study on 392 American high school students to record their boredom (as a state and a trait) as well as to know the reasons for the same. Through this study, they have also analyzed the role of students' location, age, ability, experience, and socioeconomic status in the boredom. The study results confirmed that the boredom increases in the case of academic subjects like science and mathematics, compared to non-academic subjects such as music, art, and gym (the names of the courses might have been mentioned here just for examples with respect to a student of science/engineering background - for an art student, the examples might be the opposite). The boredom may also increase while performing passive activities such as listening to the lectures compared to active participation in interactive classroom activities and lab courses. It also increases during schoolwork times compared to non-schoolwork times. The researchers have also observed that elder students get more bored during homework and less bored in performing classwork, which is the opposite in the case of younger students. However, they did not find any significant role of gender and socioeconomic status in the cause of the boredom state. Matthews et al. [86] also conducted self-report based studies, comprising of 'Dundee Stress State Questionnaire' (DSSQ) [87], to collect evidence for fundamental state dimensions of subjective state in performance settings. Through this study, they have identified distress, engagement, and worry as the three fundamental state dimensions. Shernoff et al. [118] followed the Experience Sampling Method (ESM) to conduct a self-report based study on 526 American high school students to verify the influencing factors for students' engagement. By analyzing the reports of the students, which were collected through a Likert type of response scale, they have found that the engagement is significantly influenced by the phenomenological aspects, instructional methods, and school subjects. Their study results indicate that the phenomenological aspects such as concentration, interest, and enjoyment have a high positive correlation with the engagement. The influence of instructional method on the engagement has also been witnessed – it has been observed that students become more engaged during individual and group works compared to while listening to lectures and watching videos. They have also found that students get more engaged in non-academic subjects than in academic subjects, which corroborates the findings of Larson and Richards [73] on the same aspect.

The self-report based method has been found to be effective as it helps us to find the states of the students, and at the same time, the cause of the particular states. Through a standard questionnaire, with the detection of emotion, involvement, and level of activity of the students, it is possible to know the reasons for them. For example, students may get inattentive when the lecture is too hard to understand or too easy that cause boredom [73] [118].

However, the self-report based method has some limitations. Some students may hesitate to report the actual state, whereas other students may not [138] [141]. To elaborate, some students may think that there is nothing wrong with reporting the states honestly. On the other hand, many of them may feel embarrassed to report the states, which are not expected from them (e.g., 'disengaged'). As a consequence, there is a chance of identifying fake states. Not only that; many times students are unable to assess their own behaviors, cognitions, and affective states accurately [102]. Even if they are able to do so, they may make mistakes to appropriately recall the states of particular moments. According to Hermann Ebbinghaus (a German psychologist), the human beings have a tendency to start recalling the most recent incident first and best ('recency effect'), and from the rest of a list of incidents, they can recall the first few items most frequently than the middle items ('primacy effect'). The reports may, therefore, be biased by the 'recency and primacy memory effects' [138]. Furthermore, conducting the survey and analyzing the same to know the states of the students in real-time is cumbersome, and sometimes it is impossible. We cannot ask the students to take part in the survey during lectures. Similarly, it is not possible for a teacher to analyze students' reports (collected through the survey questionnaire) to find out their mental state in real-time.

2.2.2 Manual Observation

An alternative method to know the states of the students is manual observation, which is sometimes used for confirming or triangulating self-report questionnaires [102] as well. Confirming, here, refers to verifying the results obtained by another method; whereas triangulating indicates its incorporation with the other methods so that a final conclusion can be made from the common results obtained from all the methods. In the manual observation method, as the name suggests, the teacher manually observes the students. The teacher is asked to respond to a standard questionnaire (e.g., [102] [119] [120] [123]). The teacher, based on the provided questionnaire, either reports her/his subjective opinion on the states of the students or fills up a set of checklists and rating scales for an objective measure that is supposed to indicate the states. Sometimes, the observation takes place in an indirect way, which is termed as 'Work Sample Analysis' (WSA). In WSA, pieces of evidences are collected by observing the problem solving and analysis capability of the students, which act as factors for their mental states [11].

The observation method overcomes all those limitations of the self-report based method, which are originated because of the consideration of direct feedback of the students. The method is therefore free from the identification of fake and biased states. However, the observation method requires a huge effort and involvement from the teacher's side, which may create a problem in the flow of lecture delivery. It is very difficult for a teacher to deliver the lecture, and at the same time, to observe the states of the students and report them through some questionnaires. The difficulty level increases in the case of a large classroom, where there may be a hundred to a few hundred students. In this scenario, external observers such as additional instructors and/or teaching assistants may be appointed, who might help the instructor in observing the states. However, it is expensive in terms of monetary cost – the number of observers will increase with the increase in the number of students in the classroom. For many of the academic institutes, it is next to impossible to afford the cost. Moreover, in this case, inconsistency in the detection can occur, as not all the observers may have the same level of expertise in the task. Literature reports some works where the involvement is detected by an external observer by analyzing the video of the classroom activities [65]. In this case, also, the issues of both the affordability as well as the expertise of the observer come into the picture. Moreover, this method requires expensive classroom infrastructure and supportive tools (e.g., multiple high-quality cameras in the classrooms, communication medium and protocols for streaming the videos and identified states, and a separate room and PC with video analysis software for the expert), which add to the expenses. An alternative way to overcome these limitations is systematic observation, which is also known as the automated measurement method.

2.2.3 Systematic Observation

The limitations of the manual observation and self-report based methods can be addressed by the systematic observation method. In this method, a computing system plays the role of the observer on behalf of the instructor. The system involves one or more computational model(s), which observe the factors and analyze them to detect the states.

As the systematic observation method involves neither the teacher nor the students for state detection, the problems of fake and biased state detection can be overcome. At the same time, both the students and teachers can be involved in the teaching-learning process without interruption, unlike the self-report and manual observation methods. The systematic observation method also drastically reduces the manual effort required for state detection by the teacher, which is an additional effort with the lecture delivery and other teaching activities. As some valid computational models are involved in the systematic observation method, it is expected that state detection is more consistent in this method compared to the self-report and manual observation methods. It also can help to reduce the increasing monetary cost for state detection with the increased classroom size – up to a certain limit, the models involved in this method are generally independent of classroom size.

There are many techniques, and therefore, many models to detect the states systematically. The advantages and disadvantages of them, along with the adaptability of the techniques in the current context are described next.

A. Computer vision-based models

Visible changes in physical appearances, such as changes in facial expression and body language (gesture and posture) of the students, are analyzed to detect their states. The detection is done systematically which incorporates image processing and computer vision based models.

For example, Whitehill et al. [138] proposed a classroom system to detect students' emotional and behavioral engagement, and to correlate it with task performance. The backbone of their system is a computer vision-based model, which detects the level of engagement of the students by observing their facial expressions and head movements. The model detects five levels of engagement, namely, 'not engaged at all', 'normally engaged', 'engaged in task', 'very engaged', and 'unrecognized'. The literature contains many such computer vision-based models where students' emotional states are identified from their facial expressions and gestures-postures [9] [17] [18] [36] [80] [137].

These models are found to be outstanding and related to the current context, as facial expression and body language are the manifestations of emotional states which is one of the most important factors of the mental state of the students.

However, there are some issues which may deter to apply the model in practice. First of all, the technique is expensive in terms of both the computation and monetary costs. The involvement of computer vision and image processing makes the technique computationally expensive. The devices, which are required to run the expensive computing modules, should be of high configuration, and hence costly. Secondly, a fixed setup is required to be arranged for each student for this type of model to be worked. For example, the classroom system proposed by Whitehill et al. [138] requires a webcam (in a fixed position and hence students have to sit at a fixed position), along with an iPhone for every student. However, in a real-time f2f classroom or in our assumed blended learning platform, we cannot assume that each student will always face toward a camera unless a very special classroom setting as proposed in [113] and [138]. Thirdly, external factors such as ambient light, the quality and number of cameras, and the siting direction of the students may affect the accuracy of the model.

B. Physiological and neurological signal based models

The use of physiological and neurological signals has also been observed in the literature for user-state detection. Electroencephalogram (EEG), electromyography (EMG), electrocardiogram (ECG), galvanic skin response (GSR), eye blinking rate (EBR), and heart rate (HR) are used as models' inputs to detect the states.

For example, AlZoubi et al. [3] used three physiological signals, to be specific, the ECG, EEG, and GSR to recognize affective states into one of the eight states, namely 'boredom', 'confusion', 'curiosity', 'delight', 'engagement', 'surprise', and 'neutral'. Kolodyazhniy et al. [68] assessed physiological and neurological signals and found distinct patterns in autonomic, respiratory, and facial muscle activity for 'fearful', 'neutral', and 'sad' states. Hazlett [55] examined facial EMG and found that EMG of the zygomaticus muscle becomes significantly larger when human beings are in positive valence as compared to when in negative valence. On the other hand, the EMG of corrugator muscle becomes greater during negative valence compared to positive valence. This is because the zygomaticus muscle controls smiling, whereas the corrugator muscle determines the frowning. Healey and Picard [56] used ECG, EMG, skin conductance, and respiration to determine three levels of stress, viz., ‘high’, ‘medium’, and ‘low’, with high accuracy.

As the states are detected by observing the physiological and neurological signals, most of the time the detection of the state through this technique are found to be very accurate, and hence to be useful in many applications (e.g., games [55], healthcare [25] [112], and transport [56]). Nonetheless, it is hard to adapt them in the classroom context in practice because of many reasons. First of all, this technique requires elaborate setups involving expensive sensors and probes. Secondly, students may get irritated if they are asked to wear the probes and sensors while attending a lecture on a regular basis. Thirdly, in this technique, the students are likely to be aware of the fact that their states are being monitored. In that case, they may not behave naturally. As a result, fake states may be detected [3] [105].

C. Speech based models

Speech technologies have also been utilized for building models to detect the affective and activity states [69] [92] [107]. The models detect the states mainly from three speech parameters: ‘voice quality’, ‘utterance timing’, and ‘utterance pitch contour’ [92].

These models are supposed to be useful in many important application areas, such as healthcare, forensic, robotics, and so on [69]. However, it is very difficult to adapt them to our proposed framework because it may not be assumed that students will always be talking inside a classroom. Even if that is the case, it is very hard to recognize and analyze the voice of a particular student in the presence of others’ voices without a very special arrangement of equipment and setting. Moreover, the accuracy of these methods depends on the corpora used for training the models used in these methods [107]. This is because the same word may be pronounced and expressed in different ways based on the ethnicity and gender of the users [107].

D. ITS based models

The literature contains another set of works, which is termed as ‘Intelligent Tutoring Systems’ (ITSs). The problem solving and analysis capability and style of the students are systematically observed to provide immediate and customized feedback and instruction. Both the observation and feedback provision is done systematically without requiring the intervention of a human teacher. This helps the students to learn things in a better way with greater care, as it seems to be one-to-one private tutoring [28] [97]. “AutoTutor” [49], “ANDES” [41], “ATLAS” [133], “WHY2” [48], “SHERLOCK” [79], and “PAT” [28] are some of the examples of popular ITSs. ITSs simulate the dialogue pattern of a typical human tutor, and therefore, students get a personal computerized tutor and learn things in a better way by doing extensive practice. Although the ITSs are great AI contributions in the field of education technology, they are limited to the concept of a private tutor. On the other hand, our proposed system is for regular classrooms, which takes additional care of the students as per their mental states on behalf of the teacher and provides an efficient learning environment in terms of outcome and learning experience. At the same time, it provides a better teaching experience for the teacher and reduces the manual effort required to take care of the students as per their learning state of mind.

Nonetheless, there are few works in the literature, where the ITSs have been used as a platform for collecting the required data to detect the mental states of the students. The ‘student model module’ [97] of the ITSs, i.e., the problem solving and analysis capability and style, along with the knowledge domain of the students are considered to capture their mental states. Conceptually this technique can be termed as automated WSA.

For example, Botelho et al. [19] explored deep learning method, to be specific the recurrent neural networks (RNNs), to detect the mental state of a student into either of ‘bored’, ‘frustrated’, ‘confused’, ‘concentrated’, and ‘unknown’, based on an ITS called “ASSISTments” [57]. d Baker et al. [29] developed a model, which can detect the mental state of the students as either of the ‘bored’, ‘confused’, ‘engaged’, ‘frustrated’, and ‘undefined’ from their semantic actions in the interface of “Cognitive Tutor Algebra I” [109]. A similar model was developed by Paquette et al. [99] based on another ITS, namely the “Inq-ITS” [45].

These models were found to be suitable for Intelligent Tutoring Systems (ITS). However, in these models, the features used for detecting the affective states are based on students’ problem solving and data analysis capability and style. As a result, they work only for science, engineering, and mathematics fields [5] [48]. Moreover, the models have been built for specific courses and they require content access. Furthermore, the models for detecting the affective states are dependent on specific ITSs. These characteristics of the ITS-based models indicate that they might not be adaptable in our assumed blended classroom setting which is independent of course and content.

E. Multimodal technique

There exist a few works in the literature, where the multimodal technique has been used to detect the mental state of the students. In this technique, as the name suggests, more than one input modalities are considered to detect the states. To illustrate, physiological and neurological signals can be considered with facial expression to detect the states, with the expectation for availing a higher accuracy compared to unimodal models.

For example, Dragon et al. [32] and Woolf et al. [141] developed an affect-aware tutoring system based on the "Wayang" ITS [8]. The system takes four different types of inputs to detect the mental states of the students: behavioral input (pressure applied by the students while handling the mouse), seating posture, physiological signal (skin conductance), and facial expression. Conati [27] considered heart rate, skin conductance, and eyebrow position of the students to detect their mental states. Conati's state detection model was built based on an educational game, called "Prime Climb", which was developed by the EGEMS group of the University of British Columbia. The classroom system proposed by Kim et al. [67] also requires requires several high-end cameras, microphones, haptic gloves, and a 'massively-parallel computing component' for detecting the affective state of a student.

Although the multimodal techniques have been found to be superior (in terms of accuracy) compared to the unimodal models, they are very expensive both in terms of computational and monetary costs. For instance, the affect-aware tutor [32] [141] requires a subsystem for analyzing facial expression, a posture analysis seat, a pressure sensing mouse, and a wearable skin conductance sensor, for every individual student. The system proposed by Conati [27] requires an EMG, GSR, and HR sensors for each student to detect her/his state. Availing and using these sensors, equipment, and infrastructure for a regular class is quite cumbersome. These, therefore, may not be adaptable in practice [71]. Moreover, these systems are not usable in the context of mobile and ubiquitous environment like our target blended learning classroom environments – they work on special platforms like ITS or educational game.

F. Behavioral Models

These models use behavioral data as inputs to detect the states. The behavioral inputs include device handling pattern, keystroke pattern, and touch pattern on the touchscreen while interacting with the device. There are some major advantages of considering the behavioral data as inputs. It requires neither expensive sensors and hardware nor significant infrastructure and setup. Moreover, computation is much less compared to the other techniques. Furthermore, in most of the cases, the affective data are collected without the knowledge of the users, which in turn minimizes the fake/imitation of affect.

Epp et al. [33], Khanna and Sasikumar [66], [82] Lv et al. [84], Matsuda et al. [85], Vizer et al. [136], and Zimmermann et al. [144] identified affective states from the users' keystroke patterns and dynamics. In all these works, the inputs were taken from a physical keyboard. Additional sensors (e.g., additional pressure sensors) were required to capture the keystroke pattern and dynamics. Moreover, the typing pattern on the physical keyboard and typing pattern on the virtual keyboard may be different (unlike typing on a physical keyboard, we rarely use all our fingers while typing on smartphones). The models are, therefore, not suitable to adapt to our assumed blended learning platform, as we assumed only smartphones to be used in the learning platform.

Keystroke data from virtual keyboards were used by Lee et al. [77] and Lee et al. [76] for identifying the affective states. However, their approach was not unobtrusive. They used the *"International Affective Picture System (IAPS)"* for emotion induction before data collection. The affective state at the time of induction and at the time of data collection may vary. Similarly, for the verification of their approach, they collected feedback from the participants using the SAM (Self-Assessment Manikin) method after finishing the tasks. The affective states at the time of the task and at the time of taking the feedback may differ. Moreover, the SAM method itself may change the affective state of the user.

We found some works where touch gestures have also been considered for detecting the affective states. Smartphone interaction data were used by Bauer and Lukowicz [15], Ciman et al. [25], and Sano and Picard [112] to identify users' stress levels for healthcare applications. Our target application areas are not the healthcare systems but an affective classroom system. Restriction on identifying just stress and no-stress states is not sufficient for the target application [141]. Moreover, the inputs in some of these works ([112], [15]) were not just touch gestures; they used personal data like the phone call, SMS, and user location. We did not intend to use any privacy-sensitive data, as it may not be ethical. Gao et al. [39] identified user emotion at the time of playing a game. However, most of the features in their model were defined in the context of game-playing, e.g., the number of enemies killed, the distance between two players, and so on. The model is therefore not suitable for the current context. Moreover, some of the features used in this model are computationally expensive. For instance, 'Average stroke length' and 'Average stroke Directness Index' have a computational complexity of $O(n^2)$. Features used in our approach, on the other hand, have only $O(n)$ complexity in terms of computations. Furthermore, in their approach, emotional data were collected using an expensive device (iPod$^{\text{TM}}$), which is not expected to be available in every circumstance. Shah et al. [116] attempted to overcome these limitations. They used a comparatively smaller set of features and used inexpensive devices (mobile phone priced ∼150 USD) for data collection and processing. However, some of those features like 'deviation in the number of strikes' and 'deviation in the number of taps' are difficult to know beforehand. Most importantly, none of these studies have been conducted targeting the classroom environment, and hence, may not suitable for the same.

Urh and Pejović [132] deployed an app (smartphone application) to detect whether a user is involved in 'an easy task' or 'a hard task'. The app predicts the involvement by observing the values of the gyroscope, accelerometer, and microphone sensors; along with the time, current location information, activity details (reported by the Google Activity Recognition API), information about nearby WiFi and Bluetooth devices, charging status of the battery, volume-settings information, ambient light, calendar events, and the screen-status. The work cannot be adapted in our proposed framework as a component to detect the involvement in study-related activities. This is because both the study and non-study related tasks may be easy as well as hard. For instance, a student may be engaged in playing a hard game as well as an easy game. In both cases, the student has not been involved in study related activities. Similarly, a student may be engaged in finding the meaning of a word, which seems to be an easy task. At the same time, the student may be engaged to find the answer to a conceptual question which may be very hard for her/him. Here, in both cases, the student is involved in classroom activities. Moreover, their app accesses private data, which we do not want to use for ethical reasons.

Harada et al. [54] used the smartphone's accelerometer data to observe the body movement of its user. They used it to recognize the physical activity of the students and the teacher. Consequently, they found that higher synchronization of body-movement between students and the teacher indicates better and constructive discussion. In case the body movement was not properly synchronized, the discussions were found to be not constructive. As they used accelerometer data to indirectly observe the behavior of the students (and the teacher) to recognize their activity, the work is related. However, it is different in the sense that their proposed model does not predict the involvement of the students in the classroom. Instead, it quantifies the outcome of the discussion between the teacher and students once they are involved/engaged in the discussion.

2.3 Analysis of Literature

After categorizing the existing works in the literature into groups and subgroups based on the methods applied for detecting the states, we have critically analyzed them keeping the following eleven questions in mind.

1. Do they require additional sensors and equipment for the detection of the states?

2. Do they involve high computational and/or monetary expenses?

3. Do they require privacy-sensitive data?
4. Do they involve huge manual effort?
5. Do they depend on a particular system or course?
6. Can they be directly affected by environmental factors?
7. Are they able to provide consistent results?
8. Are their results reliable?
9. Are they suitable for real-time state detection?
10. Are they suitable for detecting the states in educational settings?
11. Are they suitable to adapt to the assumed blended learning platform?

This is because the overall objective of the thesis is to present a computational framework for a sensitive classroom system that can sense the mental states of the students in a systematic, real-time, ubiquitous, and unobtrusive way, which is suitable to easily implement in a classroom in practice in a cost-effective manner, without any additional equipment and sensors, independent of any particular system and course, without using privacy-sensitive data, and with minimal effect of external factors. The literature has, therefore, been analyzed with the help of the eleven questions constructed based on this. The rationale behind the construction of these eleven questions is as follows.

Consideration of inspecting the requirement of additional sensors and equipment is important as their incorporation may result in the unsuitability of the systems and/or models in the current context. They may add to the monetary as well as the computational cost. Consequently, the system may be infeasible to be adapted in many institutes. Moreover, the students may not agree to wear additional sensors for attending lectures on a regular basis. As a result, the system may not be suitable to be deployed in a real-time classroom setting and hence, for real-time state detection. The involvement of high computational costs also involves greater infrastructure settings which ultimately add to the monetary cost. For many institutes, it may not be possible to afford such costs. The use of the personal and private data of the users may not be ethical as they may be misused and cause serious issues. If the detection method requires huge manual effort form the students and teacher, it may affect the teaching-learning process instead of improving the same. Also, the detection method should not be dependent on a particular system (e.g., particular ITS), course (e.g., science, engineering), or the materials of the particular course, so that it can be applied independently for detecting the states of the students for all courses. In other words, the detection method should not be somewhat where it can detect the mental states of science students but unable to detect the states of language students, for example. Consideration of reviewing the effect of environmental factors for the detection method (e.g., lighting condition, and seating postures of the students in the case of computer vision-based method) is also important as it may affect the accuracy and consistency in the detection. Verifying the reliability and consistency of state detection is essential as it may have an influence on the applicability of the system and/or model. For taking some action based on the identified states, to improve the current classroom situation, detection of the states in real-time is also important. There may exist a model that is suitable to detect a particular state for a specific application (e.g., stress detection in healthcare application). However, the same may not be adequate and/or enough for the current context. Therefore, assessing the suitability of existing works in educational settings is essential. At the same time, it should also be assessed whether a model and/or system is suitable for the particular blended learning platform. This is because there exist different types of classroom settings, and one particular state detection technique may not be adapted for all of them because of the infrastructural differences.

We, therefore, assessed the existing works in the literature through the above mentioned eleven questions so that we could find out a suitable state detection method which can be adapted in our assumed blended learning platform to make it sensitive. We aimed to select a method that has '$\times$' (the answer is 'no') for the first six questions, and '$\surd$' (the answer is 'yes') for the rest five questions among the eleven. The summary of the analyses is presented in Table 2.1, and are described as follows.

Alongside the existence of few advantages (e.g., identification of the cause of the states along with their detection), the self-report method have many disadvantages such as unreliability, bias, and requirement of huge effort and involvement of the teacher and students which cause interruptions in the teaching and learning activities. As a result, the method is found to be infeasible to apply for real-time state detection. The manual observation method does not involve the students. Hence, students can take part in learning activities without interruption. The problem of fake and biased state detection can also be reduced to a significant extent in the manual observation method. However, it still requires a huge effort and involvement of the teacher, and hence cause interruptions in lecture delivery. Utilizing the method for real-time state detection may not be possible. There is a probability of inconsistency in state detection as well in this method (when external observers are hired for state detection). Therefore, the manual observation method is also not suitable in the current context.

The third alternative, i.e., the systematic observation method, seems to be suitable for real-time state detection. However, the most of the existing models, which follow the systematic observation method (computer vision, physiological and neurological signal, speech, ITS based models) may not be applied in our assumed blended learning platform due to several constraints, such as the involvement of expensive computing and monitory costs, lack of infrastructure, the effect of environmental factors, imitation in the state, user dissatisfaction because of the requirement of additional and wearable sensors and probes, infeasibility, and dependency in other systems and specific courses.

Although only the behavioral models of the systematic observation method are seemed to be suitable to apply in the current context (as they may have '$\times$' for the first six questions, and '$\surd$' for the rest five questions), the existing works in this category suffer from at least one of the following issues: unsuitable in mobile and ubiquitous environment, imitation in detected states, unsuitable for educational setting and application, the involvement of high-computational complexity, and requirement of privacy-sensitive data access.

Table 2.1: Summary of the analyses of existing related works for state detection and their categories based on the detection method

Method	**Sub-method**	**Requirement of**				**Dependency on particular system course**	**Effect of environmental factors**	**Consistency**	**Reliability**	**Suitable for**		
		additional equipment	**high computation/ monetary cost**	**privacy sensitive information**	**huge manual effort**					**real-time state detection**	**educational setting**	**assumed blended learning platform**
Self report	–	×	×	×	√	×	×	√	×	×	√	×
Manual Observation	Observation by teacher	×	×	×	√	×	×	√	√	×	√	×
	Observation by external observer(s)	×	√	×	√	×	×	×	√	×	√	×
Systematic Observation	Computer vision based models	√	√	×	×	×	√	×	√	√	√	×
	Physiological and neurological signal based models	√	√	×	×	×	×	√	√	√	×	×
	Speech based models	√	√	×	×	×	√	×	√	√	×	×
	ITS based models	×	√	×	×	√	×	√	√	√	√	×
	Multimodal approach	√	√	√	×	×	×	√	√	√	√	×
	Behavioral Models(s)	×	×	×	×	×	×	√	√	√	√	√

We, therefore, motivated to build some novel computational models which address all the identified issues, and utilizing which we can build a computational framework for a sensitive blended learning platform which can systematically detect the mental states of the students based on their affective states, involvement in study-related activities, and level of participation in the classroom activities. Hence, we aimed to focus on building behavioral computational models for real-time detection of affective states, involvements, and activity levels of the students, which are suitable to adapt in the assumed blended learning platform to make it sensitive.

2.4 Chapter Summary

In this chapter, we have presented the exiting works for detecting the mental states and their components. We have found that the existing works can be classified into three major groups based on the methods applied for detecting the states: self-report, manual-observation, and systematic observation methods. After the critical analyses of the literature survey, we have found that the self-report and manual observation methods are not suitable to adopt in the current context. Although only the behavioral models (belong to the systematic observation method) seem to be suitable, the existing behavioral models in the literature suffer from many issues that restrict them to be adapted in our assumed blended learning system. Thus, we aimed to build and validate several novel behavioral models addressing all the identified issues so that we can utilize them for building a framework for a sensitive classroom system that is able to detect the mental states of the students based on their emotion, involvement, and activity level. The proposed framework and the novel models are presented in the subsequent chapters.

"If we teach today's students as we taught yesterday's, we rob them of tomorrow."

John Dewey (1859 - 1952)
American educational reformer

3

Vedinkaksha: A Framework for a 'Sensitive' Classroom

3.1 Introduction

The word "Vedinkaksha" consists of two Sanskrit words: 'Vedin' and 'Kaksha'. The first part, i.e., 'Vedin' means sensitive, and the second part, i.e., 'Kaksha' means room/classroom. The two words collectively refer to 'sensitive classroom'. The term 'sensitive' is used because a learning platform built under the umbrella of the proposed framework is able to sense (or detect) the mental states of the students automatically. The process is unobtrusive as it requires neither direct nor indirect feedback from the students. In order to detect the mental state of the students, Vedinkaksha considers their emotions, involvement status, and the levels of classroom activities (Figure 3.1). At the same time, the learning platform can automatically take some precautionary measures based on the detected states, for improved learning and teaching experiences. It sends alert signals to the uninvolved and inactive students so that they may become involved and active. The instructor is also alerted, if the majority of the students are not involved in the teaching-learning process, are frustrated, feel uninterested, or unable to understand the learning contents delivered by the instructor. Additionally, it provides the facilities for quantifying learning and understanding levels of the students and visualizing

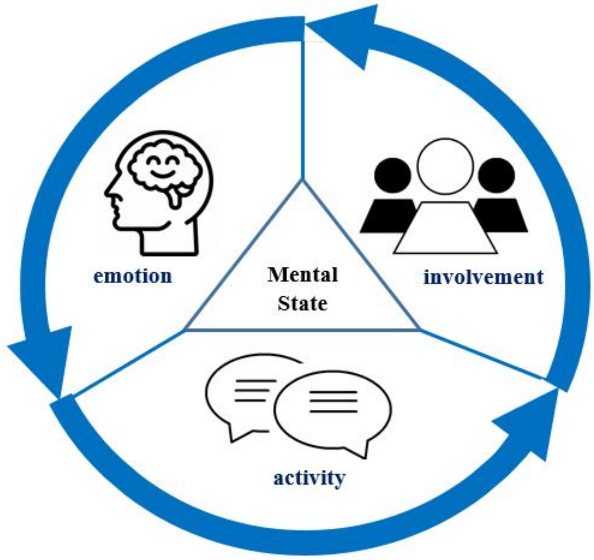


Figure 3.1: Components of mental state in the context of learning progress

the classroom based on the identified states.

In this chapter, we will present the proposed framework in details. We will elaborately discuss the system architecture of a learning platform under the framework, along with the working principle of detecting the mental states in the framework. This is followed by an explanation for each of the key components of the framework. We will also illustrate how the various computational models can be incorporated to detect the mental states of the students based on students' emotions, involvements and levels of classroom activities.

3.2 Avabodhaka: A Blended Learning System

We have considered "Avabodhaka" [24] [130] as the basis of our proposed framework for a sensitive classroom system.

In Avabodhaka, although both the teacher and the students remain present in a live face-to-face classroom, they perform all the classroom activities through mobile computers such as laptops, smartphones, and tablets. Avabodhaka runs on a client-server architecture (depicted in Figure 3.2) in the BYOD paradigm. A local server provides the necessary infrastructure for the classroom activities and the contents for the same. All the devices used by the students and the teacher

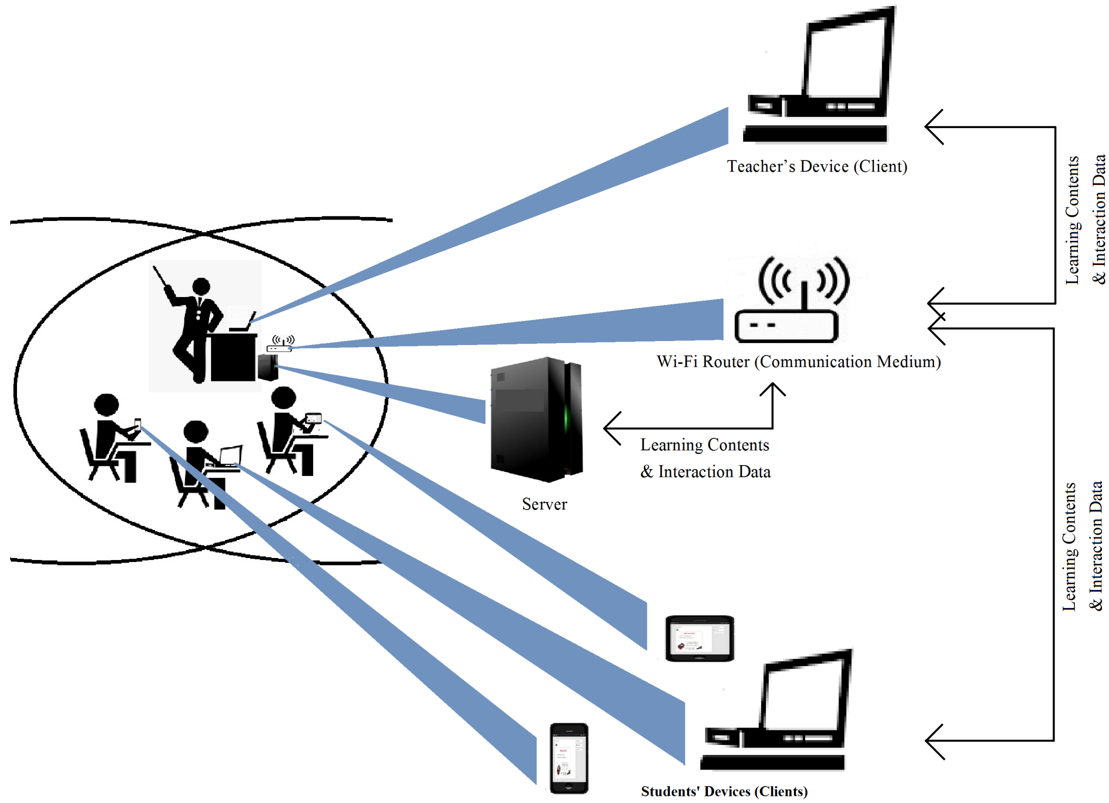
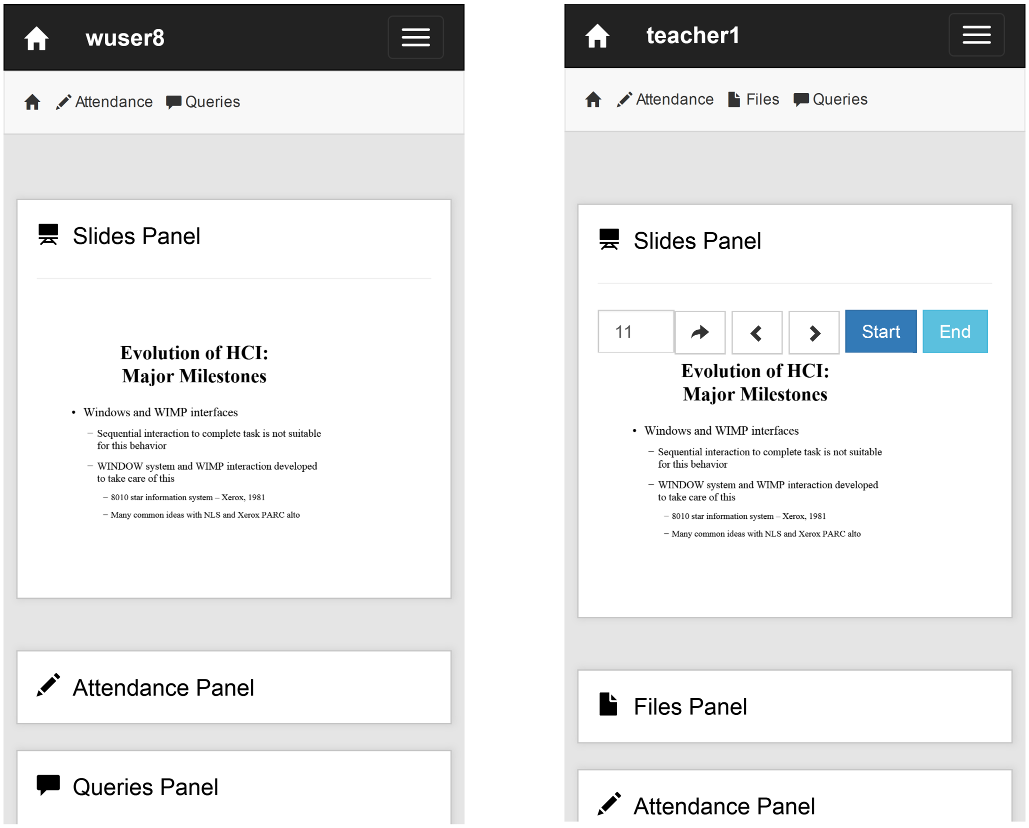


Figure 3.2: Semantic diagram of Avabodhaka system

act as clients. The teacher and the students can take part in classroom activities using Avabodhaka after logging in to the system with their credentials. Avabodhaka facilitates the following classroom activities (discussed in the next five sub-sections) for the teacher and the students, with an expectation for greater interactions between the students and the teacher as well as among the fellow students, which helps in improving teaching-learning experience with better learning outcomes compared to the traditional f2f classroom system [130].

3.2.1 Lecture Delivery

The teacher can upload the lecture materials (pdf/ppt) to the Avabodhaka server and open the same from the lecture delivery panel. Once opened, it is streamed to all the devices, which are connected to the server in a lecture session. If the teacher changes the slide, it is propagated to all students' devices. It may be noted that only the teacher can change the slide. Figure 3.3 demonstrates the lecture panels of



(a) Lecture Panel for Students **(b) Lecture Panel for Teacher**

Figure 3.3: Screenshots of Avabodhaka's Lecture Panel

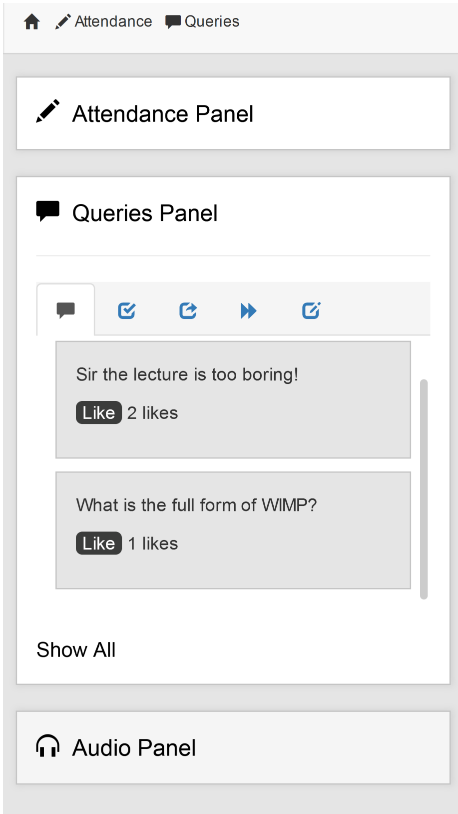
Avabodhaka for both the teacher and students.

3.2.2 Audio Streaming

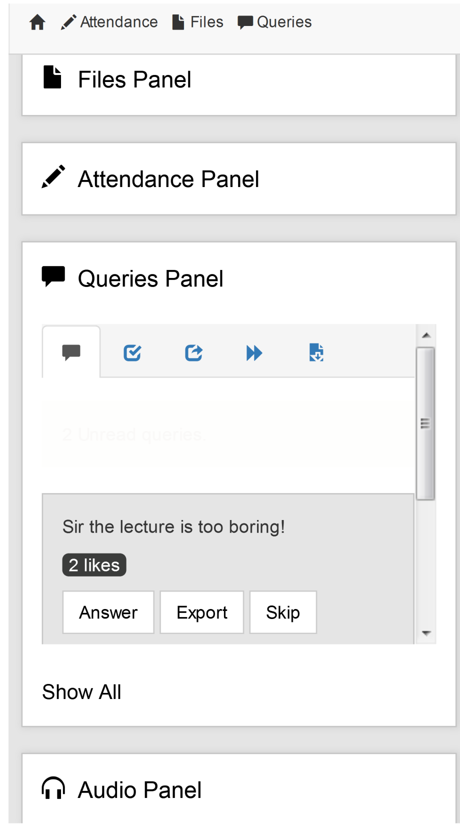
The speech of the teacher is also live-streamed to all the devices that belong to the students. If the voice of the teacher is not naturally audible, which may occur in the case of a large classroom, students can plug an earphone/headphone into their device to hear the voice smoothly and loudly.

3.2.3 Query and 'Like'

The 'query panel' of the system facilitates the students with an interface to anonymously raise queries to the teacher. If the students have any queries during the lecture, they can post the queries in this panel. With the teacher, all the other students can view the queries. However, only the teacher can reply to the queries.



(a) Query Panel for Students



(b) Query Panel for Teacher

Figure 3.4: Screenshots of Avabodhaka's query panel

In case another student has the same query, s/he can 'like' it instead of retyping it. Based on the number of 'likes', the queries are sorted and displayed in descending order so that the teacher can answer the queries and doubts as per their priorities. The most liked query gets the highest priority and is placed at the top of the query-list, which is supposed to be answered first. Figure 3.4 depicts the 'query-like' mechanism of the system. The query posting mechanism has been kept anonymous to minimize the hesitations for asking questions from shy students.

3.2.4 Examination

There is also a provision for conducting quizzes and examinations in Avabodhaka. A teacher can upload questions along with the answers. The teacher can post three types of questions: the 'Multiple Choice Question' (MCQ), 'exact-match', and 'short-answer' types of questions. Students can choose one option out of the

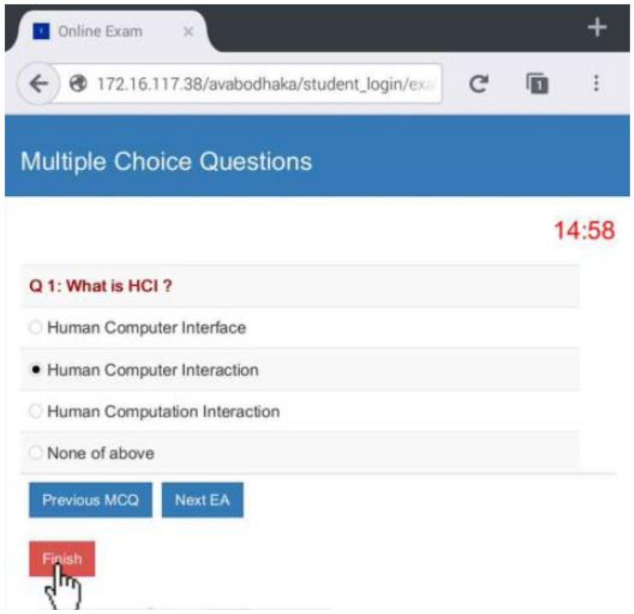


Figure 3.5: Screenshots of Avabodhaka's Examination Panel

given multiple options for the MCQ type of questions, whereas, in case of an exact-match type of question, they are supposed to enter a particular keyword that should be exactly matched with the unambiguous answer. For the short-answer type of questions, students are expected to type a sentence. The system evaluates this type of question, based on a word matching algorithm. When the teacher wishes, s/he starts the examination session. The system keeps all the answers hidden unless the exam is over. Students, who are logged in, can take the exam. They can respond to the questions by either checking the radio buttons (for the MCQ type of questions) or entering the text from the keyboard (for the exact-match and short-answer type of questions). Once the scheduled time gets over, the system automatically evaluates the responses and shows the results to the students. The results are also stored in the server's database for future reference. Figure 3.5 represents screenshots of 'examination panel' of the system.

3.2.5 Attendance

The system also provides a facility for recording attendance. The teacher, as per their preference, activates the 'attendance panel', and only then the students get the scope for marking their attendance.

We have chosen this particular blended learning platform as the basis of our proposed framework for a sensitive classroom because of the following two reasons. Firstly, it provides a better interface for students-teacher interactions, and consequently better learning outcomes. Through a controlled experiment [130], we have found that both the students and the teacher are found to be interested to adopt the learning platform for performing regular classroom activities. Also, they interact more while using this system compared to the traditional f2f classroom system. It has also been observed that students are able to secure higher marks by using the Avabodhaka system. Secondly, we can utilize the infrastructure of the system to use the sensory, backchannel, and resource utilization data for the detection of mental states of the students, without requiring any external sensors and equipment.

3.3 Issues with Avabodhaka

Although Avabodhaka is proven to be a better platform for teaching and learning activities, there is no concept of 'sensitivity' – it is unable to detect the mental states of the students. For making it 'sensitive', i.e., for proposing Vedinkaksha, we required a modification in the system architecture of Avabodhaka, and had to add some new features to the system. The changes in the system architecture and features, along with their roles in implementing the sensitivity are described below.

3.3.1 Changes in System Architecture

To propose Vedinkaksha, we have modified the system architecture of Avabodhaka in the following way. We have assumed that all client devices of Vedinkaksha should be either smartphones or tablets, whereas, in the proposal of Avabodhaka, client devices could be any of the mobile computers such as smartphones, tablets, laptops, and palmtops. We have assumed only smartphones and tablets as the client devices because of wide availability of these devices nowadays [61]. Consideration of smartphones and tablets as the client device is, therefore, most suitable for the BYOD paradigm. There is another very important reason for this consideration: we can utilize the presence of some embedded sensors in these devices for collecting data to ubiquitously detect the mental states of the students. The four most important

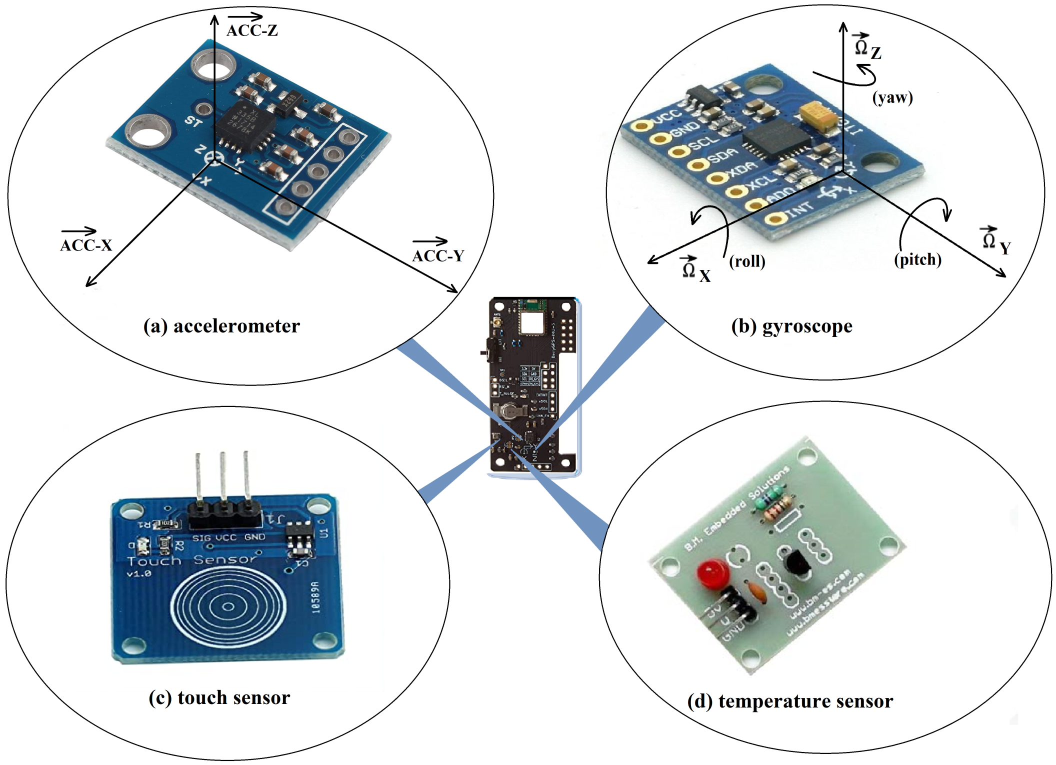


Figure 3.6: Four embedded sensors which are available with every smartphone nowadays

sensors found in almost every smartphone are the accelerometer, gyroscope, touch, and temperature sensors (Figure 3.6). The usefulness of these sensors in the current context is presented as follows.

A. Accelerometer

The accelerometer is one of the sensors which is known as 'cheap but accurate'. Generally, we find a three-axis accelerometer embedded in smartphones and tablets. When the device moves toward an arbitrary direction, we can record the scalar components along the X, Y, and Z axes (Figure 3.6(a)) of the actual acceleration. The acceleration value toward the arbitrary direction can be derived from these three scalar components.

B. Gyroscope

When a device moves, it not only accelerates in a particular direction but may rotate or tilt as well. This angular movement is detected through another useful sensor called gyroscope. When an angular movement along an arbitrary axis occurs, the gyroscope detects its components along X, Y, and Z axes which are known as 'roll', 'pitch', and 'yaw', respectively (Figure 3.6(b)). The rotation value along the arbitrary axis can be derived from these three scalar components.

The rotation and acceleration values together can help us detect the shake and movement pattern of the device, and hence handling pattern of its user. The device handling patterns of the users have been used as input features for the behavioral models which have been developed by us and are essential to be incorporated in the proposed 'sensitive' framework.

C. Touch Sensor

Another important user behavior in the current context is the interaction pattern on a touchscreen device. The touch sensor (Figure 3.6(c)) captures the touch events and their characteristics, which indicates the interaction behavior of individual users. The behavioral differences of the users have also been found to be useful in building computational models for the framework.

D. Temperature Sensor

Resource consumptions may vary for different types of applications, which can be availed from the system logs. In addition to the dynamic resource consumption patterns, there may be changes in the physical properties of the devices and/or its parts depending upon the type of applications. One such important perceivable physical change is the varying temperature of the battery and hence of the device, which is measured by the temperature sensor (Figure 3.6(d)). In most cases, the temperature sensor comes as an integrated component of the battery. The varying temperatures of the battery for different applications are also found to be useful for building one computational model for the current context.

3.3.2 Changes in System Features

With the modification of the system architecture, we added four new system features in the Avabodhaka for the proposal of Vedinkaksha. The added features are supposed to provide scopes for collecting essential data, using which the system can detect the mental states of the students. As an additional benefit, these features may make the learning platform more attractive and useful as well. The details of the four new features are as follows.

A. Class-note

During the lecture, students can add notes through the 'Classnote' panel. The notes are organized as per the sequence of the slides and locally stored in the students' devices. The server records the class note taking logs, i.e., the notetaking time with the identification information of the students who add the notes during the lecture session. The class note feature contributes to the framework in two ways. Firstly, it helps to collect the typing data which are utilized as inputs for one of the affective state detection models. It may be noted that the typing data are available through the query panel as well. However, it is expected that the probability of getting the typing data during lecture sessions will be increased due to the addition of this feature. This is because, many a time, students may not have a query but may wish to note down some important information for future references. Secondly, the note-taking logs, along with the logs for other classroom activities, help to quantify the level of classroom activities.

B. Random Popup

In a random interval, a pop-up message is displayed to the student's device, which disappears automatically after five seconds. Students are supposed to respond to the popup message within this specific period of time. The 'random popups' help to measure students' attentiveness in the classroom. It also can help to measure the activity level as it is one of the classroom activities in the proposed framework.

C. Live-quiz

The teacher can initiate instant quizzes through this panel during the lecture. The quiz should consist of a few questions about the topic which is currently being taught. The questions of the 'live-quiz' may be added by the teacher at the beginning of the lecture or anytime during the lecture. The responses to the live-quizzes help to measure the understanding level of the students in real-time. The server logs the scores earned by the students for these quizzes. These scores can be used as one of the grade processing elements, and hence can help in measuring the learning levels as well.

D. Alert and Visualization

The detected mental states of the students might be underutilized if those are not informed to the teacher in real-time. We, therefore, devised of presenting a visualized classroom to the teacher based on the detected states in real-time, which might help in real-time assessment of the class. We also devised of sending alert signals to the students and the teacher based on the states of the students for the improvement of the pedagogy. The students are supposed to be alerted when they are not engaged in the teaching-learning process, or fail to understand the lessons but pretend otherwise. An alert, in the form of a motivational message, can also be sent to the students who feel shy to participate in the classroom activities. We have also devised of alerting the teacher when a majority of the students are found to be uninvolved in the teaching-learning process, frustrated, uninterested, or unable to understand the learning contents delivered by the instructor, so that s/he may tailor the way of lecture delivery and/or lecture content.

With the addition of these new features and changes in the system architecture in Avabodhaka, we built a framework for a ‘sensitive’ classroom called Vedinkaksha, which can detect the mental states of the students and may take some actions for an improvement in the pedagogy. The detailed description of the framework, along with the working principle of the same is presented in the subsequent sections.

3.4 Proposed Framework for a Sensitive Classroom

3.4.1 Infrastructural Requirement

We have assumed the same infrastructure requirement as specified in Avobodhaka, for Vedinkaksha to work. The only difference, as mentioned earlier, can be found in the assumption of client nodes. The full list of infrastructure requirements for Vedinkaksha is given below.

1. Smartphones or tablets for all the students and the teacher (BYOD), where a

smartphone application (app) will run to securely facilitate the interfaces for all the classroom activities, and provision for conveying the state information as well as the alert signals. The smartphones/tablets should have at least the following four sensors embedded in them: accelerometer, gyroscope, touch sensor, and temperature sensor.

2. A peripheral device (we assume a smartwatch), which can receive the alert signal for the teacher, and at the same time displays the details of the signal.

3. A local server, which can process the raw sensory, resource utilization, and backchannel data to compute the mental states of the students, and at the same time stores this sensitive information securely and serves the same upon a requirement to the authentic users. The server also stores and supplies the learning materials for all the specified classroom activities, i.e., lecture delivery, examination conduction and evaluation, attendance recording, and 'random popup' generation. It also stores the activity logs of both the teacher and students and processes them on demand. The local server also helps the clients to connect to the Internet for accessing global learning contents, if the academic policymakers allow.

4. A Wi-Fi router, which connects and communicates with all the client nodes with the server.

3.4.2 System Architecture

Figure 3.7 depicts the system architecture of a learning platform under the Vedinkaksha framework. The flow of inputs and outputs along with their roles in the 'sensitive' framework are presented as follows.

- **Login Credentials:** Both the teacher and students should log in to the system by entering their credentials. These user credentials are required to uniquely identify the users so that the sensitive data can be collected and processed for the individual users. It would also help to securely send sensitive data, information, and signals to the authentic users.

- **Peripheral Device’s ID:** We have devised of sending alert to the teacher through a peripheral device for minimizing the interruption in lecture delivery (e.g., a smartwatch). The peripheral device could be connected to Vedinkaksha via unique identification information.

- **Teaching and Examination Contents:** Once the students and the teacher are connected to Vedinkaksha, it should get activated. The teacher would be able to upload the materials of lecture delivery and examination conduction, and access them whenever required. After opening the lecture content by the teacher, it should be shared with all the students. If the teacher changes the contents it should be propagated to all the client devices. However, the students would not be allowed to change lecture contents. The teacher would also be able to upload questions with their answers beforehand and initiate examinations at any desired time.

- **Query and Backchannel Data 1:** There should be a provision for raising queries by the students for clarifying their doubts and taking part in discussions. We have devised of a poll-based ranking system of the queries like in Avabodhaka. Through this query posting and ‘liking’ mechanism, Vedinkaksha might be able to collect the touch and typing pattern of the students, which specify the behavioral signatures of the individual students and might be helpful for identifying their affective states. With the query and ‘like’, the students might be engaged in responding to ‘random popups’ and ‘live-quizzes’ as well. These responses’ history could be termed as ‘backchannel data’. Other backchannel data received from the students’ side might be devices’ temperature, power, and memory utilization details. We term all the backchannel data availed from the students’ side as ‘Backchannel Data 1’. These backchannel data could also be used for detecting students’ emotions, involvements, and activity levels.

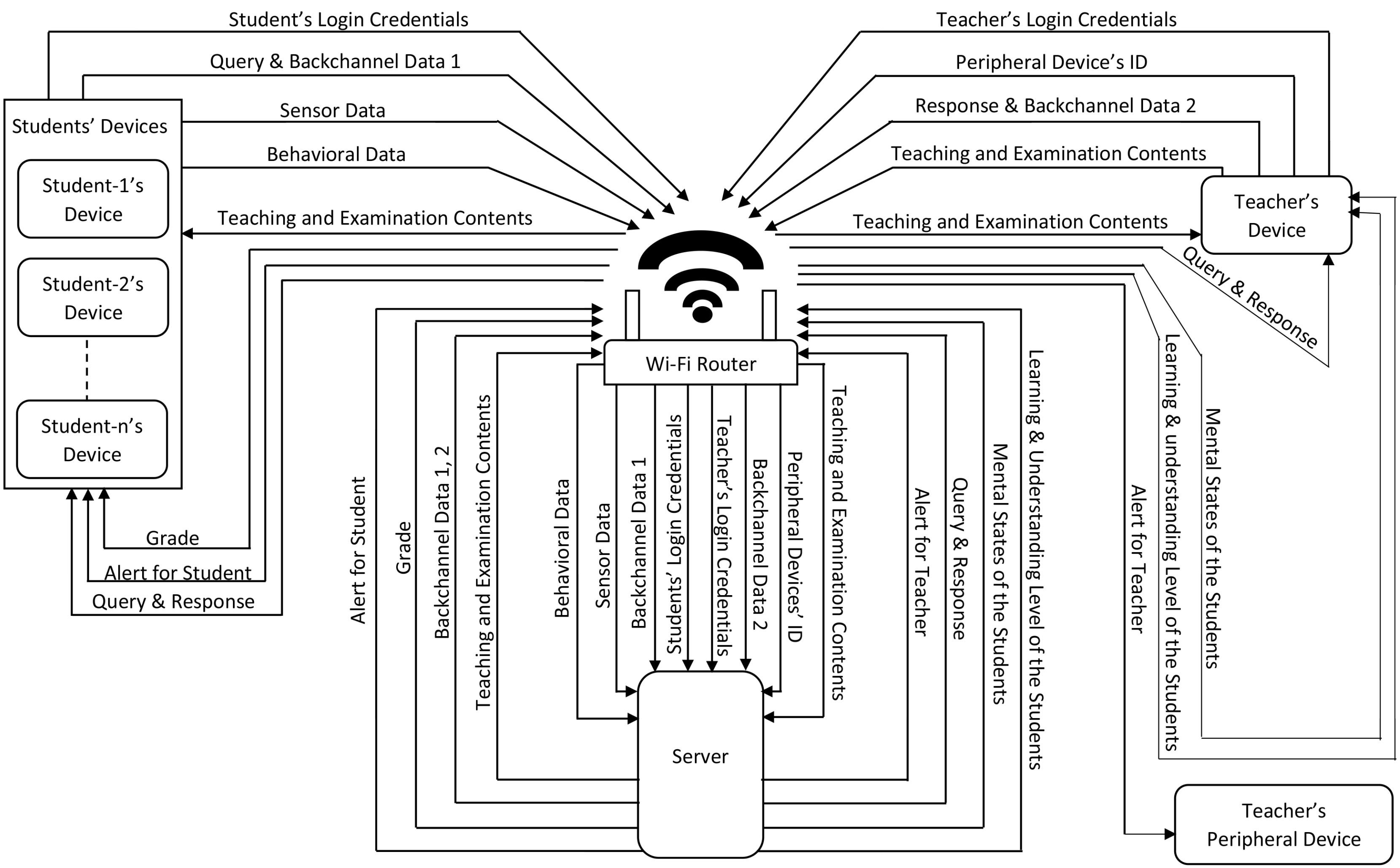


Figure 3.7: System architecture of a sensitive classroom which follows the Vedinkaksha framework

- **Response and Backchannel Data 2:** The teacher should be able to answer the queries, or ignore a query (if s/he thinks that the question is irrelevant in the current context). It can be noted that some backchannel data may be availed from the teacher's side as well, which we term as 'Backchannel Data 2'. The examples of 'Backchannel Data 2' are the number of queries responded to, the number of alerts sent (particularly for not understanding the teaching content) to the teacher, and the number of times the alerts are acted on. These data might also be useful for taking some smart decisions such as the requirement of resending the alert to the teacher.

- **Sensor Data:** Readings from the embedded sensors of students' smartphones could also be collected and utilized by the sensitive modules of Vedinkaksha, which detect the affective states, involvement, and activity levels of the students. The data of accelerometer, gyroscope, temperature, and touch sensors might be useful in this regard.

- **Behavioral Data:** Along with the sensory data, the activity data related to the behavior of individual users might also be helpful for state detection. Users' typing pattern (typing speed, the number of characters typed at a time without making an error, smartphone shake frequency while typing), touch pattern (the number of touches, the pressure generated while touching), and device handling pattern of the users (e.g., whether the device is being shaken frequently) are the examples of such behavioral data which might be useful in this context. It can be noted here that many a time the behavioral information is availed by processing the sensory data.

- **Mental State of the Students:** Based on the sensory and behavioral data along with the resource utilization pattern and activity log, various computational modules of Vedinkaksha would be able to detect the affective states, involvement, and activity level of the students. Based on these, it would be able to detect the mental states of the students, which is detailed in the next section.

- **Grade:** Grade would be automatically calculated, by the examination module of Vedinkaksha, based on the responses of the students in the examinations and quizzes. It would be displayed to the students immediately after the examinations and quizzes get over and stored in the server for future references. Grade history would be useful for determining the learning level of the students.

- **Learning and Understanding Levels of the Students:** Here, the term ‘understanding’ refers to the short-term grasping of the lecture content, whereas the term ‘learning’ indicates the long-term outcomes from the learning process. The level of understanding of the students might be computed from the response-history of the live-quiz. It could also be detected by observing their activity, emotion, and involvement. The learning level of a student might be defined based on the historical data on the understanding level along with the grade earned by her/him.

- **Alert:** A student should be alerted if an undesired state is detected. For example, if a student is found to be not engaged, s/he might be alerted by a device vibration along with a string of texts having some warning and/or motivational message. The teacher might also be alerted if the majority of the students are found to be in some undesired states or they are unable to understand the learning contents.

The above system architecture provides sufficient infrastructure for detecting the mental states of the students, and take some smart actions based on these states for improved teaching and learning experiences. The first and most important aspect in this regard is to detect the mental states of the students, which is the primary focus of this thesis. The state detection can be followed by taking some actions to complement and/or comprehend them. The next section presents the working principle for mental state detection in detail, along with the smart actions which can be taken by the framework based on the identified states.

3.4.3 Working Principle

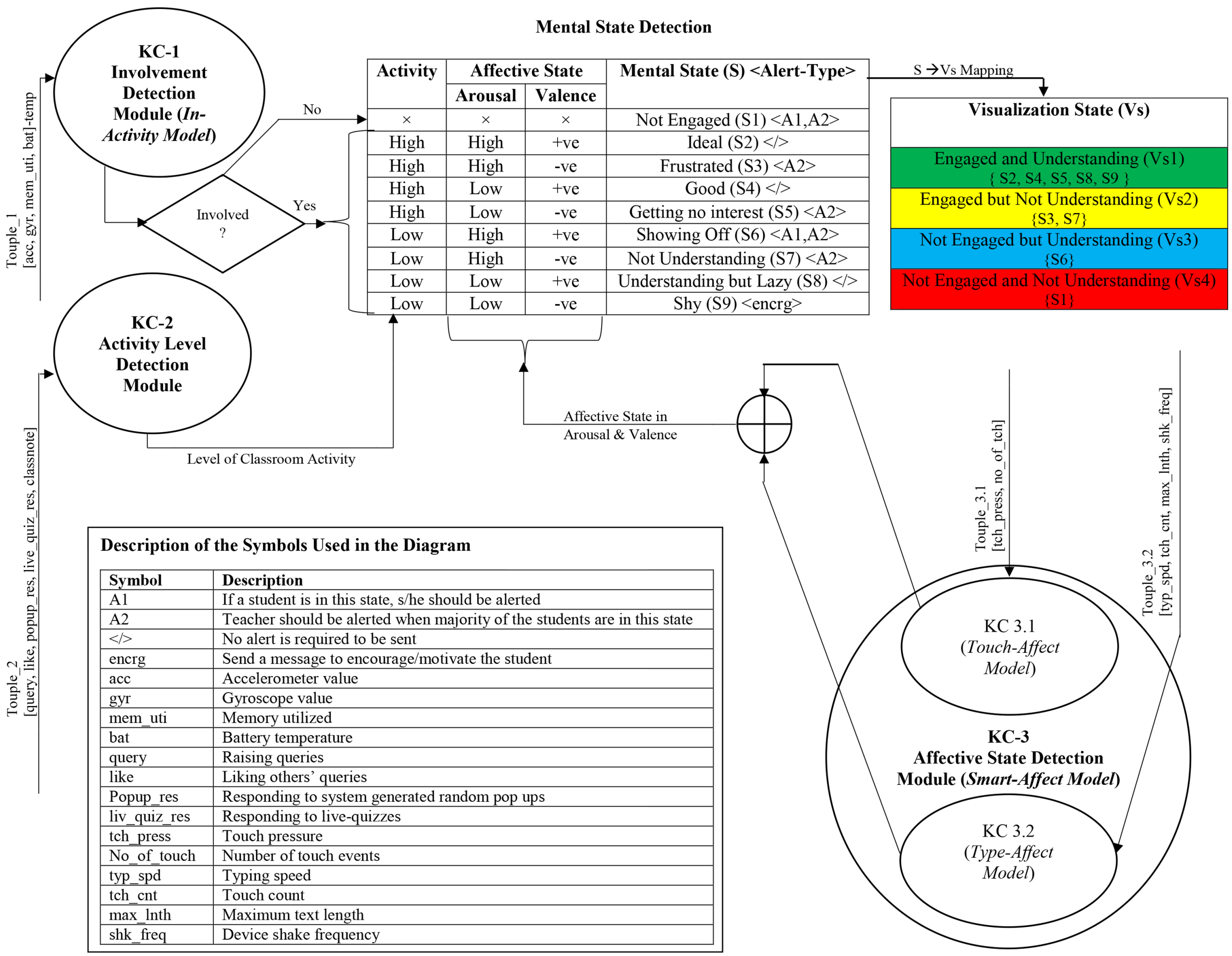
The mental states are defined based on three key factors: involvement, activity level, and affective states of the students [6] [7] [35] [50] [62] [108] [139]. These three key factors are computed by three key components (KCs) of the framework: KC-1, KC-2, and KC-3 that have been built and validated by us. The working principle for detecting the mental states with the help of these three KCs is depicted in Figure 3.8, and is described as follows.

KC-1 is responsible for detecting the involvement of the students. It decides whether a student is involved in learning activities in Vedinkaksha or doing something else (e.g., playing a game, watching a movie, and performing social networking). KC-1 observes device handling and memory utilization patterns, along with the battery temperature to detect the involvement. The device handling pattern is identified from the values of accelerometer and gyroscope sensors of the smartphone.

KC-2 classifies the level of classroom activities as ‘high’ or ‘low’. If a student performs at least one activity from the list of the classroom activities within a certain interval of time, their activity level is considered to be high. Otherwise, their activity level is considered low. The list of classroom activities is as follows: making a query, liking a query, responding to random popups, participating in the ‘live-quiz’, and taking class notes.

KC-3 is a process model to detect the affective states of the students from their behavioral data, i.e., from the touch and typing pattern on a smartphone/tablet. This is composed of two subcomponents: KC 3.1, and KC 3.2. KC 3.1 detects the affective states of the students from their touch behavior, whereas KC 3.2 detects the affective states from students’ typing behavior. Based on the type of the available input (touch data/typing data) KC-3 collects the input data string and sends it to a particular subcomponent which returns the affective state as one of the following four states based on the Circumplex model of emotion [106]: ‘high-positive’, ‘high-negative’, ‘low-positive’, and ‘low-negative’. Here, ‘high’ and ‘low’ indicate the levels of arousal, whereas the ‘positive’ and ‘negative’ indicate the levels of valence. This particular emotional model has been found to be better suited to the classroom context than the rest [141], which is the reason behind its consideration in our work.

If the students are found to be involved in other activities, Vedinkaksha defines the state of the students as ‘Not Engaged’ in the learning process. This is the S1 state (as shown in Figure 3.8) during which Vedinkaksha sends an alert signal to



Activity	Affective State		Mental State (S) <Alert-Type>
	Arousal	Valence	
×	×	×	Not Engaged (S1) <A1,A2>
High	High	+ve	Ideal (S2) </>
High	High	-ve	Frustrated (S3) <A2>
High	Low	+ve	Good (S4) </>
High	Low	-ve	Getting no interest (S5) <A2>
Low	High	+ve	Showing Off (S6) <A1,A2>
Low	High	-ve	Not Understanding (S7) <A2>
Low	Low	+ve	Understanding but Lazy (S8) </>
Low	Low	-ve	Shy (S9) <encrg>

Visualization State (Vs)
Engaged and Understanding (Vs1) { S2, S4, S5, S8, S9 }
Engaged but Not Understanding (Vs2) {S3, S7}
Not Engaged but Understanding (Vs3) {S6}
Not Engaged and Not Understanding (Vs4) {S1}

Description of the Symbols Used in the Diagram

Symbol	Description
A1	If a student is in this state, s/he should be alerted
A2	Teacher should be alerted when majority of the students are in this state
</>	No alert is required to be sent
encrg	Send a message to encourage/motivate the student
acc	Accelerometer value
gyr	Gyroscope value
mem_uti	Memory utilized
bat	Battery temperature
query	Raising queries
like	Liking others' queries
Popup_res	Responding to system generated random pop ups
liv_quiz_res	Responding to live-quizzes
tch_press	Touch pressure
No_of_touch	Number of touch events
typ_spd	Typing speed
tch_cnt	Touch count
max_lnth	Maximum text length
shk_freq	Device shake frequency

Figure 3.8: Working principle of mental state detection and smart actions to be taken in Vedinkaksha

warn the student so that s/he may get involved in the learning process. In this state, the levels of classroom activities and her/his emotional state are not considered.

If the student is found to be involved in the teaching-learning process, Vedinkaksha considers her/his affective state and the level of performing classroom activities to detect her/his mental state as one of the eight states namely, the ‘Ideal’ (S2), ‘Frustrated’ (S3), ‘Good’ (S4), ‘Bored’ or ‘Uninterested’ (S5), ‘Showing Off’ (S6), ‘Not Understanding’ (S7), ‘Understanding but Lazy’ (S8), and ‘Shy’ (S9) states. To illustrate, when a student is involved in the teaching-learning process, s/he is highly active in performing classroom activities, and her/his emotional state is ‘high-positive’, s/he might be enjoying the learning with enthusiasm. This might be considered as the ideal state of a student in the context of classroom environment. We, therefore, termed the particular mental state as the ‘Ideal’ state (S2). In this case, there is no requirement for sending an alert signal. On the other hand, if a student is found to be in a ‘high-negative’ emotional state instead of being active and involved, s/he is probably trying her/his best but not satisfied. In this case, s/he might be in a ‘Frustrated’ (S3) state. If the majority of the students are in this state, the teacher should be sent an alert signal. In case, a student is found to be ‘low-positive’ emotional state while s/he is ‘involved’ and her/his activity level is ‘high’, the student is probably happily enjoying the learning. In other words, although the student presumed to be not very excited, still in a ‘Good’ state (S4). In this case, neither the student nor the teacher should be alerted. When a student is found to be ‘involved’ but her/activity level is ‘high’ and s/he is in ‘low-negative’ emotional state, Vedinkaksha defines her/his mental state as ‘Bored’ (S5). In this case, the student should not be alerted but the teacher might be informed about the state. When the majority of the students are in this state, the teacher should be sent an alert signal, so that s/he can change the topic or start an advanced topic. When a student is found to be ‘high-positive’ while s/he is ‘involved’ but her/his activity level is ‘low’, it is presumed that the student is unable to understand but ‘Showing Off’ (S6) as a good student. In this case, the student should be alerted. The teacher should also be alerted if most of the students are in this state. In case the affective state of a student is ‘high-negative’ while s/he is ‘involved’ but her/his activity level is ‘low’, it is presumed that the student is ‘Not Understanding’ (S7) the teaching contents.

For this state, the student should not be alerted, but the teacher should be informed about it when a majority of the students belongs to this state. If a student is found to be 'low-positive' while s/he is 'involved' but her/his activity level is 'low', the student might be understanding well and enjoying the class but not participating in the classroom activities because of laziness. In the case of detecting a 'Lazy' student (S8), no alert signal is required to be sent. Instead, a motivational message can be sent to the student. When a student's affective state is 'low-negative' and her/his activity level is 'low' but s/he is 'involved', the student is probably 'Shy' (S9). The shy students should be encouraged by sending a motivational message so that they can actively participate in classroom interactions and discussions. In addition to alerting the students and teacher based on the identified states, there should be a provision for informing the identified states to the teacher, so that s/he can assess the class in real-time. Vedinkaksha is able to inform the teacher about the states in a simpler form so that it might take minimal effort to perceive the states, and hence assess the classroom in an efficient way. We propose to group together the similar types of mental states (in terms of students' engagement and understanding levels) into four visualization superstates for this, as depicted in Figure 3.8.

It is challenging to compute these mental states. There exist several research issues for building the KCs. The two most important issues here are:

1. How to unobtrusively detect the affective states of the students in the assumed blended learning platform without using any extra equipment and sensors, and without the knowledge of the students.

2. How to unobtrusively detect students' involvement in teaching-learning process for the assumed learning platform without using any extra equipment and sensors, and without the knowledge of the students.

To address the above-mentioned research issues, we have conducted many studies and built four computational models. The first one is a novel computational model that can detect whether a student is involved in teaching-learning activities, or doing something else which is unrelated to study. This model works as the KC-1 for Vedinkaksha. The second one is a process model that can detect the affective states of smartphone users from basic user-smartphone interaction data, i.e., from touch and typing behavior of the users. This model acts as the KC-3 for detecting the mental states of the students. The cornerstones of the KC-3 are two computation models that play the roles of KC 3.1, and KC 3.2. The novel minimalist computational model of KC 3.1 is used to detect the affective state of a smartphone user from her/his touch pattern on the touchscreen. On a similar note, the novel computational model of KC 3.2 detects the affective state of a smartphone user from her/his typing pattern on the smartphone's virtual keyboard. All these four models, which have been built by us through behavioral (**Chapter** 4) and empirical studies (**Chapter** 5), are detailed in **Chapter** 6.

3.5 Chapter Summary

The framework for a sensitive blended learning platform, named as Vedinkaksha, which is able to sense the mental states of the students, has been proposed. While proposing the framework, we have considered a blended learning platform, called Avabodhaka, as the basis of the framework. We have discussed how few changes in the system architecture and features of the existing blended learning platform have helped us build the framework for a sensitive learning platform. The working principle for detecting the mental states, and taking some smart actions aiming to comprehend and/or complement the states for an improved teaching-learning experience has also been discussed. The proposed framework involves some computational models, through which the sensitivity of a learning platform has been achieved. The building and validation processes of these models have been presented in the subsequent chapters.

"Tell me and I forget. Teach me and I remember. Involve me and I learn."

Benjamin Franklin (1706 – 1790)
American polymath

4

Behavioral Study on Smartphone Usage by Indian Students

4.1 Introduction

During the development of computational models for Vedinkaksha, we required the knowledge of the smartphone usage behavior of the students. The knowledge was essential particularly for building the model for involvement detection. This model would detect whether a student is involved in performing study-related or non-study related activities with her/his smartphone. Therefore, it was required to know the list of smartphone activities performed by the students, along with the frequency of performing the activities, both inside the classroom and in general. Although we required to know about smartphone activities inside the classroom, we were also interested to know the behavior in general. This is because, we assumed (and later confirmed by the result of the study) that students may tend to frequently perform the activities inside the classroom, which are performed frequently in general as well.

When we explored the literature for the knowledge, we found that most of the existing studies regarding these are either old or based on a specific group of students (e.g., students of a certain academic class of a specific institute of a particular location [78] [100] [114]), or have not been conducted on Indian students.

The results of old studies might not serve our purpose as the behavior may change over time. Several reasons may exist for such changes. One of the reasons is technological progress. For instance, if we consider the usage of Short Message Service (SMS) over time, it can be found that the average number of SMS usage in 2005 in the UK and Denmark was from 21 to 36 per day [12]. A few years later, in 2011, students used to send fifteen SMS/day on average – although the number is little bit higher in the case of girls, i.e., 39 SMS/day [117]. In 2014 less than half of the British students (only 41%) used to send more than five[1] SMS/day. Nowadays, we rarely use SMS for general communication purposes. This is required only for some official purposes, to receive OTPs, and sometimes in case of emergency. Although few stakeholders use the SMS for advertisement and promotions, those may be considered as commercial and marketing purposes - not for personal use. The probable reasons may be the availability of advanced multimedia supported texting services, e.g., Instant Messaging Services (IMS). Other reasons for changes in preferences and habits of the students on performing smartphone activities may include their age group, class of study, financial status, gender, and ethnicity. For example, it has been found by Park and Lee [101] that there are differences in behavior based on the gender of the students. Schroeder et al. [114], and Han and Yi [53] found that the dissimilarities in the behavior may occur due to different age groups and classes of study. As per their observations, college students text more frequently than school students. Nonetheless, the frequency gradually starts decreasing when they get older. Chen et al. [23] found that although both the male and female students are equally addicted to the smartphone, the types of applications used by these two groups are different. In 2018, Nayak [94] conducted a study to know the relationships among smartphone usage, addiction, and academic performance based on students' gender. He found that although female students are more addicted to using smartphones, the effect of addiction is more on male students.

The behavior may also be changed because of the ethnicity of the students. As per the report of Bowen and Pistilli [20], 40% of American students used iPhone in

[1] https://www.statista.com/statistics/466675/frequency-of-using-main-mobile-phone-for-sms-texting-uk/.

2012. After five years in India, the iPhone share is only 11% [1]. This is one of the examples where we can observe that the behavior may differ based on ethnicity. One of the reasons for such differences may be the different socio-economic backgrounds of the students.

We should not build a model (or we should not make a decision) for Indian students based on the behavior of foreign students, where the model's inputs are behavioral data, and the behavior may change because of the ethnicity. Also, we should not use out-of-date data for building such a model, as the behavior may change with the passage of time. At the same time, we did not find any recent study which has been conducted on Indian students. This is instead of having a huge number of the student population in India. We, therefore, were motivated to conduct an up-to-date study of smartphone usage, especially on Indian students. The subsequent sections present the details of the study and its results along with the utilization of the same for developing the Vedinkaksha framework, particularly for the model to detect the involvement.

4.2 Method

We conducted a quantitative survey. A pilot study was conducted to prepare a survey questionnaire. The survey was conducted both online (via an electronic medium) and offline (pen-paper method). We conducted an online survey for college students and an offline survey for the school students. The survey data were analyzed thoroughly to have knowledge of the smartphone usage behavior of the students. The conclusions on the findings were made based on the statistical analyses of the results. The details of the survey methodology are described as follows.

4.2.1 Study Design

We conducted a descriptive type of quantitative survey [74]. The questions used in the survey were semi ordered categorical, and close-ended [21] [74]. We have used the term 'semi-ordered categorical' because the options for responding to the questions were in between categorical and ordinal. We conducted the survey following the guideline from [74].

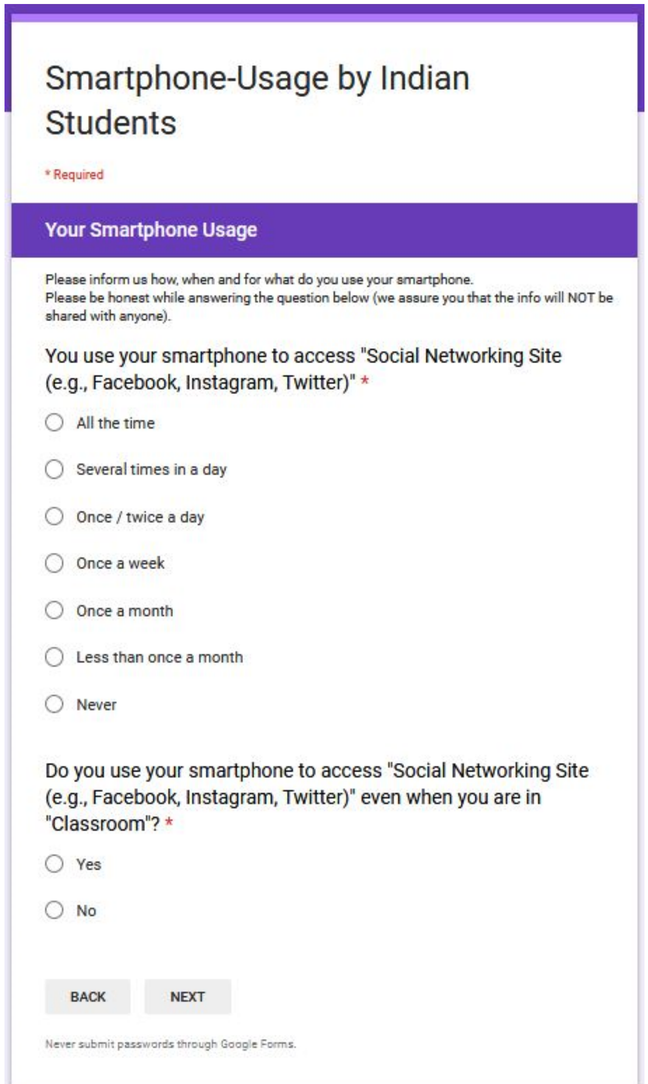
A. Identification of Smartphone Activities

In order to know the names of the smartphone activities which are generally performed by the students, we conducted a pilot study. The study helped us to know all the possible activities performed by the students in general and inside the classroom. We conducted the pilot study with 121 UG and PG students (69 male and 52 female, were in the age group of 20-30 years) from an institute of national repute in India. In the pilot study, each participating student was provided a blank sheet and asked to report the name of all the activities performed by them using smartphones. The data collection procedure was anonymous and the participants were informed that no personal information will be retrieved or disclosed. At the time of inviting the students through a formal e-mail, the participants were informed that they might come to our research lab for participation anytime – as per their convenience, and the study would take a maximum of ten minutes. We also informed the students that participation was voluntary. However, we gave chocolate to each of the participants at the end of the experiment as a token of appreciation.

We collected the data, i.e., the anonymous reports about the name of the smartphone activities from the students for one week. During this one week, we collected data from 121 participants in total. We compiled all these data and found a list of 29 smartphone activities, viz., accessing location-based information, accessing weather information, banking & other financial transaction, booking cabs, booking hotel/ticket (movie, travel), browsing shopping sites/apps, calculator, capturing pictures, checking and setting task scheduler, checking date and time, instant messaging, listening to music, playing games, random activities, reading/writing blog/microblog, reading news, recording video, recording voice, sharing media, SMS, social networking, surfing educational sites, taking class note, video call, viewing stored pictures, voice call, watching live-streaming, watching movie/video, and writing e-mails.

B. Preparation of Survey Questionnaire

The survey questionnaire was prepared based on the result of the pilot study, following the guidelines of Lazar et al. [74]. For each of the twenty-nine activities,


Smartphone-Usage by Indian Students

* Required

Your Smartphone Usage

Please inform us how, when and for what do you use your smartphone.
Please be honest while answering the question below (we assure you that the info will NOT be shared with anyone).

You use your smartphone to access "Social Networking Site (e.g., Facebook, Instagram, Twitter)" *

- All the time
- Several times in a day
- Once / twice a day
- Once a week
- Once a month
- Less than once a month
- Never

Do you use your smartphone to access "Social Networking Site (e.g., Facebook, Instagram, Twitter)" even when you are in "Classroom"? *

- Yes
- No

BACK NEXT

Never submit passwords through Google Forms.

Figure 4.1: An example of the questions used in the survey questionnaire

identified in the pilot study, we made a question having two parts. In the first part, each student was asked about the frequency of performing a particular activity. There were seven options for the participants, among which they had to select one. In the second part of each of the questions, they were asked to report whether or not the particular activity is performed by them inside the classroom. An example of the questions used in the survey questionnaire is shown in Figure 4.1. An additional question was kept to know if the participant performs any activities other than the specified twenty-nine activities, those were identified from the pilot study.

At the beginning of the survey form (link of which has been provided in the footnote[2]), we asked questions to collect demographic data of the participants. The data include the 'academic-class', 'age', 'annual family income', 'gender', 'name of the institute', and 'native place' of the participants.

Answering each of the questions in the survey form was mandatory for participating in the survey. The only exception was there for the question related to the annual family income. In the survey form, we did not ask the name of the participating student. The Reason for the anonymity was to target more responses and most importantly to get honest responses from the participating students. Someone might hesitate to provide exact information if the form was not anonymous. We wrote the questions in simple English (explanation and examples were also there, wherever required) to make them easily understandable by every student. This was because the students of India belong to diverse backgrounds in the context of the medium of instructions (there are 22 official languages in India). We devoted special care at the time of preparing the survey questionnaire for getting honest and correct responses from the participants. Special care was also taken to minimize other latent issues of conducting an online survey. We followed the guideline from Cairns and Cox [21] and Lazar et al. [74] for these. It might be thought that instead of self-report we might log the actual usage details. However, it is cumbersome to log the activity-details throughout a big country like India. Moreover, the logging of the data may be unethical as the data may have many privacy-sensitive contents.

C. Target Participants

Through this survey, with the usage frequency, we also wanted to observe the behavioral differences in performing smartphone activities based on students' gender, academic class, and ethnicity. For this, we required data from every region of the country covering all academic levels. We targeted participants from all types of academic institutes: secondary and higher secondary schools, general degree colleges, engineering colleges, and medical and pharmaceutical institutes across India. The students of three academic levels, namely, school students, undergraduate students, and postgraduate students were focused in this study. School students, here, refers

[2]https://docs.google.com/forms/d/e/1FAIpQLSdL9jS3dy3QOFjgZeHAhFYMA7JRcsntMUp6T5nxLQAjiMZzgA/viewform?usp=sf_link

to the students of secondary and higher secondary levels. The students of primary schools have not been considered in this study.

D. Two Versions of Survey Form

It is generally easy to formally reach the college students as most of these institutes have some email facilities for the students. However, the same is hard in the case of school students. We, therefore, prepared an offline version of the survey form with the same set of questions that were there in the online version. Printed copies of the offline versions of the form were used to collect the data from school students. We used ‘Google Form’, and ‘Microsoft Office Word’ for preparing online and offline versions, respectively.

4.2.2 Procedure

For the online survey, we first collected the e-mail addresses of the Deans, Directors, HODs, Principals, and other administrators of various reputed institutes throughout India. Then we wrote a formal e-mail to these administrators explaining the purpose of the study and asked for their permission to conduct the survey among their students. In the e-mail, we mentioned that upon agreeing they could forward the email to their students requesting voluntary participation in the survey. The link (URL) of the online survey form was also provided in the e-mail, mentioning that only those students should participate who use smartphones. In the email, as well as in the survey form, it was mentioned that the data collection procedure was completely anonymous and the data would be used solely for research purposes – not for any commercial purpose. It was also mentioned that by no means the identities of the participants would be retrieved or disclosed. We approached 133 engineering colleges, 89 general degree colleges and universities, and 28 medical and pharmaceutical institutes throughout India via e-mail.

In the process of data collection from school students, we physically met the principals of six secondary and higher secondary schools with prior appointments. In the meetings, we discussed the purpose of our study and asked permission to conduct the survey. We also assured the principals, and later the participating students, of the anonymity and confidentiality of the data. Each of the six principals permitted

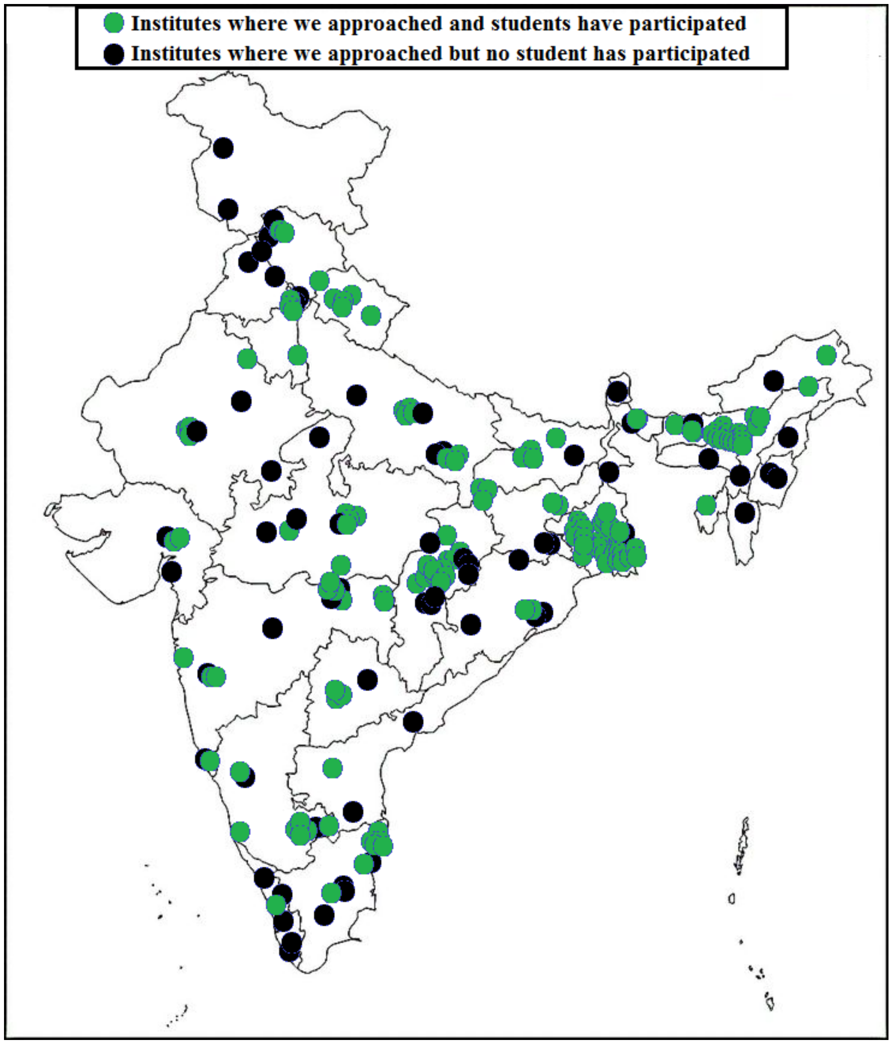


Figure 4.2: Distribution of academic institutes we approached and got participants throughout India

us to conduct the survey and fixed a particular date for conducting the survey. We took helps from a few teacher volunteers to distribute the printed version of the survey form, and to collect the same after students' responses.

4.2.3 Participant Details

We approached 256 academic institutes in total, including the schools. However, we got responses from 188 institutes (96 engineering colleges, 74 general degree colleges, 12 medical and pharmaceutical institutes, and 6 secondary and higher secondary schools). The distribution of all these academic institutes throughout India is depicted in Figure 4.2.

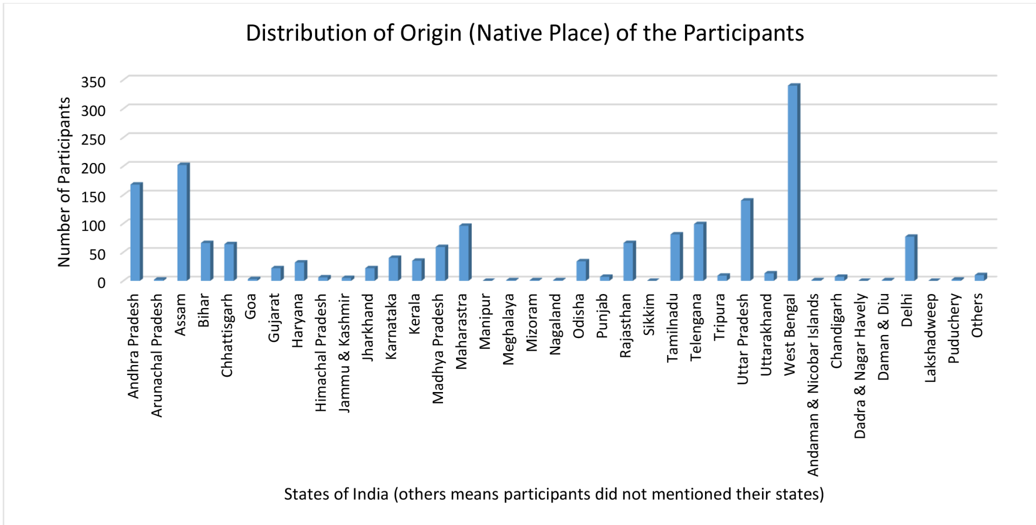


Figure 4.3: State-wise distribution of the students who took part in the survey

We got data of 1711 students (excluding outliers that has been specified in the next subsection) from 188 academic institutes throughout India. Among them, data from 1469 students (belonging to 182 institutes) were collected through the online survey. The rest of the data (of 242 students) were collected through the offline survey.

Most of the 1711 participants were in the age group of 13 to 31 years. We got participants from each state and union territory (every part of the country). They belonged to all academic levels (except the primary level) and a variety of socio-economic backgrounds. The overall distribution of the participants' gender was as follows. Male=76% and female=24% (for college students, male=79% and female=21%; whereas for the school students, male=52% and female=48%). All the demographic information have been graphically depicted in Figure 4.3 – 4.8. The information indicate that we had a good sample of Indian students.

4.3 Analysis of Survey Data

4.3.1 Outliers

We collected data from 1784 students (1522 online, and 262 offline participants) in total. However, some of them were discarded as outliers. We manually scrutinized all the responses of the participants for selecting the outliers. We discarded the data

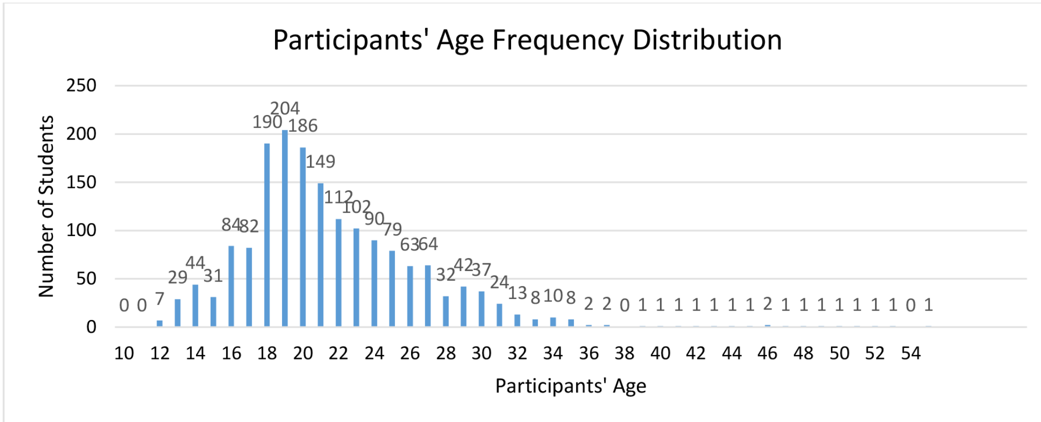


Figure 4.4: Age frequency distribution of the participating students

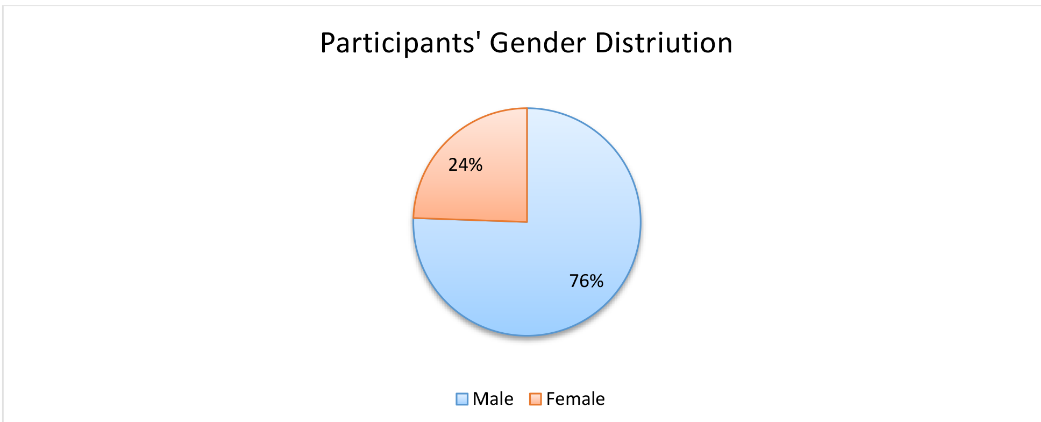


Figure 4.5: Gender distribution of the students who took part in the survey

of those participants who provided exactly the same answers for all the questions (for example, some of the participants selected 'never' for the first part and 'no' for the second part of each of the questions), who did not answer all the questions (in case of school students), or who indirectly showed their irritations by using offensive words as well as slang in the 'comments' field in the survey form. We assumed that these students might not have participated sincerely and/or honestly. Thus 73 of them (53 from the online and 20 from the offline survey) were rejected as outliers, and rest 1711 participants' data were used for further analyses.

4.3.2 Tools

We used Microsoft Office Excel 2013 for data visualization and analyses. The Microsoft Office Excel 2013 was used for performing statistical tests as well.

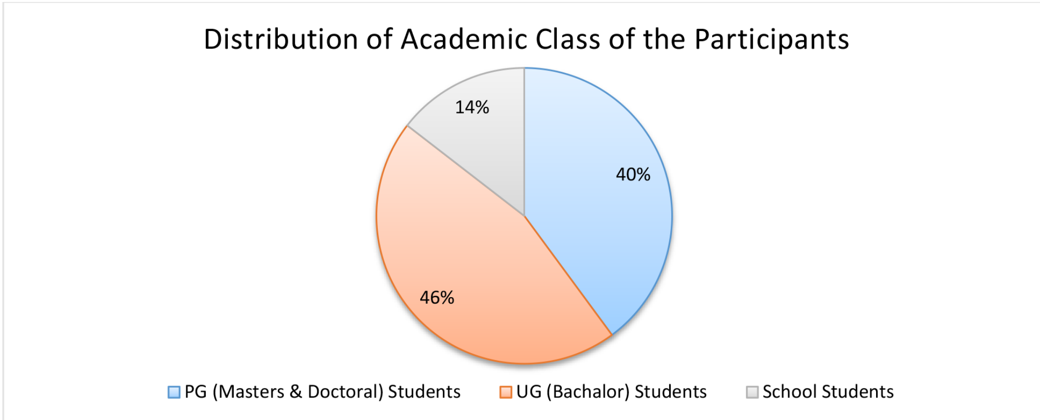


Figure 4.6: Distribution of academic classes of the participating students

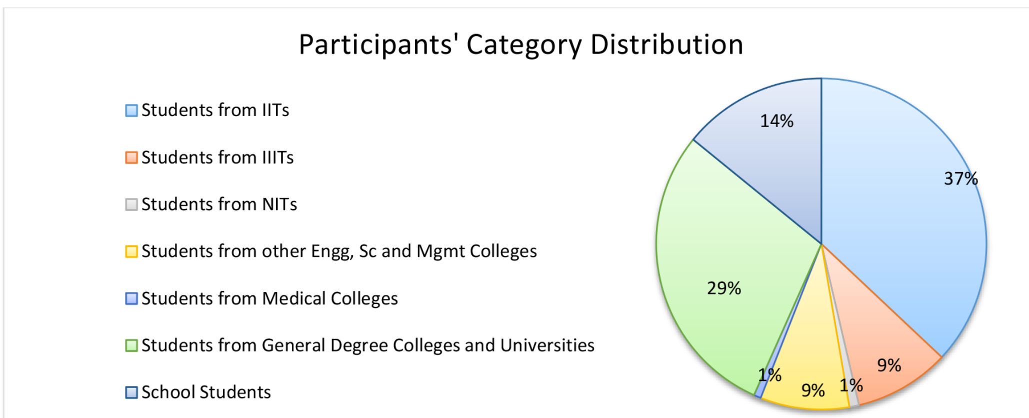


Figure 4.7: Distribution of the categories of the participating students

4.3.3 Statistical Test

As questions used in our quantitative survey was semi ordered categorical and close-ended type, and the conclusions were made by comparing two unequal groups (e.g., male and female students, school and college students, students of North East and South regions), we performed 'Pearson's Chi-Square Test' for statistical significance tests mentioned in this work. We claimed the behavioral difference for a particular activity between two groups as significant, only when the p-value was found to be less than the α-value. We considered α as 0.05. This means that the result found was significant with 95% confidence, only when the p-value was found to be less than 0.05.

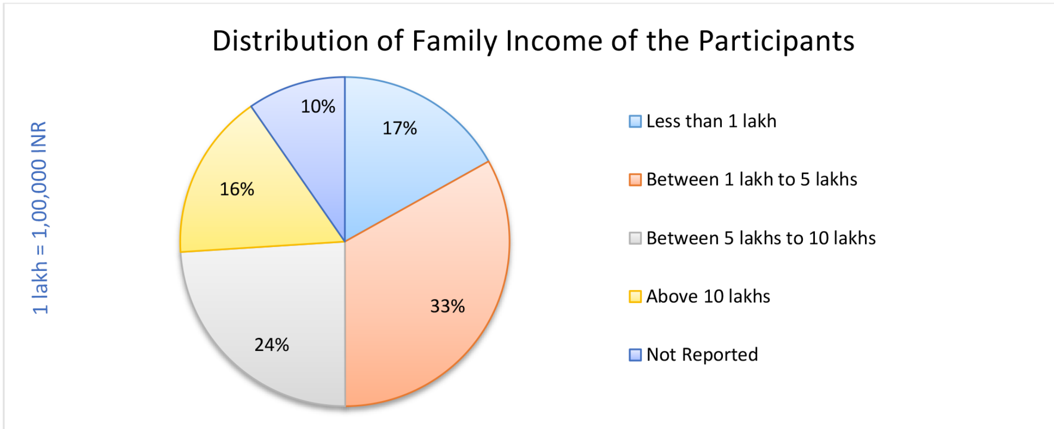


Figure 4.8: Distribution of the family incomes of the participating students

4.4 Results

4.4.1 Frequency of Performing Smartphone Activities

A. Frequency in General

The overall survey statistics are presented in Figure 4.9 in the form of a bar chart. To illustrate, the blue colored portion of the bottom-most bar (labeled 29) of the figure means that 29% of the students perform instant messaging all the time; next to that, the orange-colored portion of the same bar (labeled 53) indicates that 53% of the students perform the instant messaging several times in a day; and so on. The activities are put in ascending order, from top to button, based on their frequency of performances.

B. Frequency inside the Classroom

The frequency of smartphone activities performed by the students inside the classroom is presented in Figure 4.10. In this figure, the darker shaded portion of the bar indicates that the specific percentage of the students perform the particular activity inside the classroom; whereas the lighter shaded portion of the same bar indicates that the specific percentage of the students do not perform the specific activity in the classroom with their smartphones. For example, in the bottom-most bar, the dark portion (labeled 66) denotes that 66% of the students use their smartphones to check the date and time inside the classroom. Rest 34% of the participating students

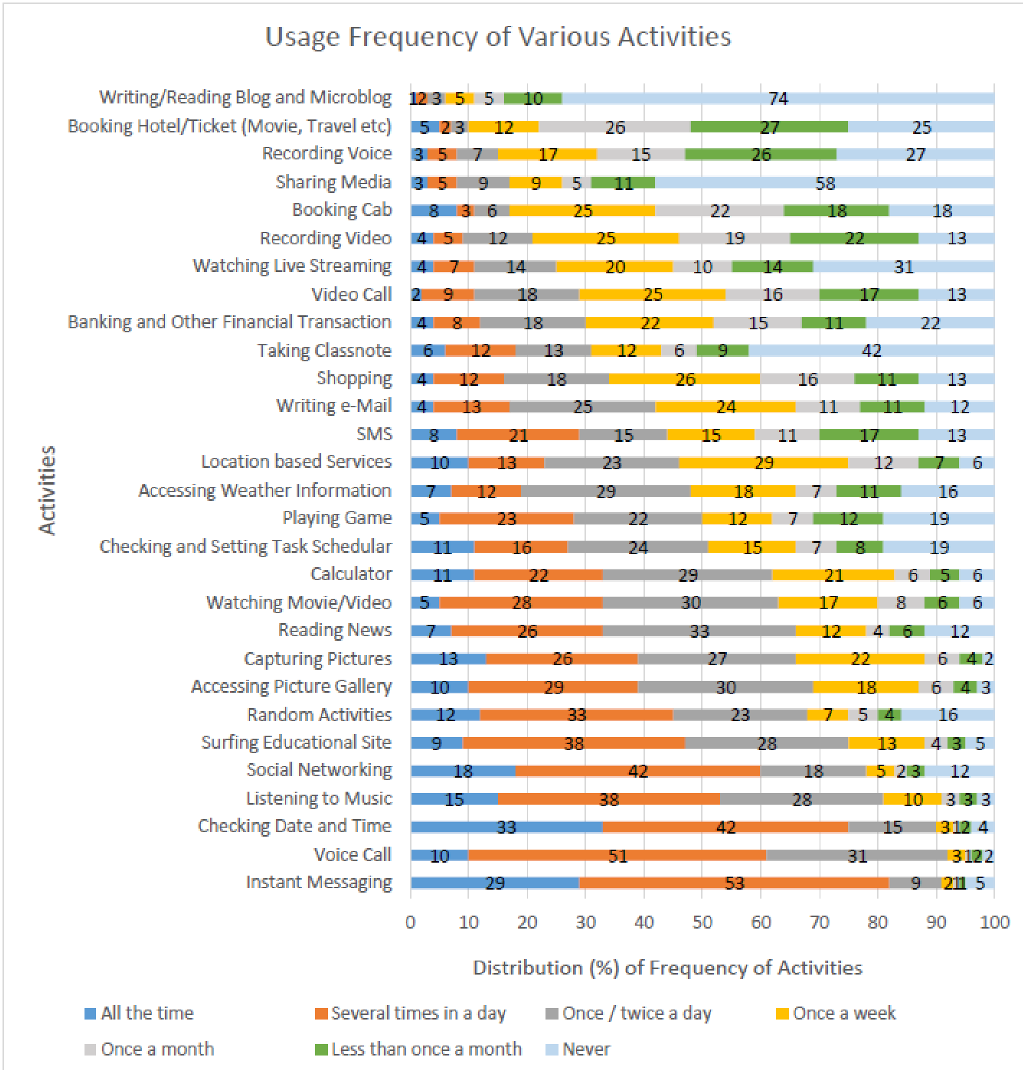


Figure 4.9: Overall usage statistics of smartphone activities in general

do not use their smartphone to check the date and time inside the classroom.

It can be observed that inside the classroom, students have a tendency to perform all the activities which are performed most frequently by them in general (except the 'voice call' and 'taking class-note').

C. Frequencies of Additional Activities

It may be recollected that we additionally asked each student (through the 30^{th} question) to know whether s/he uses her/his smartphone for any other activities other than the specified 29 activities in the survey form. About 75% of the participants

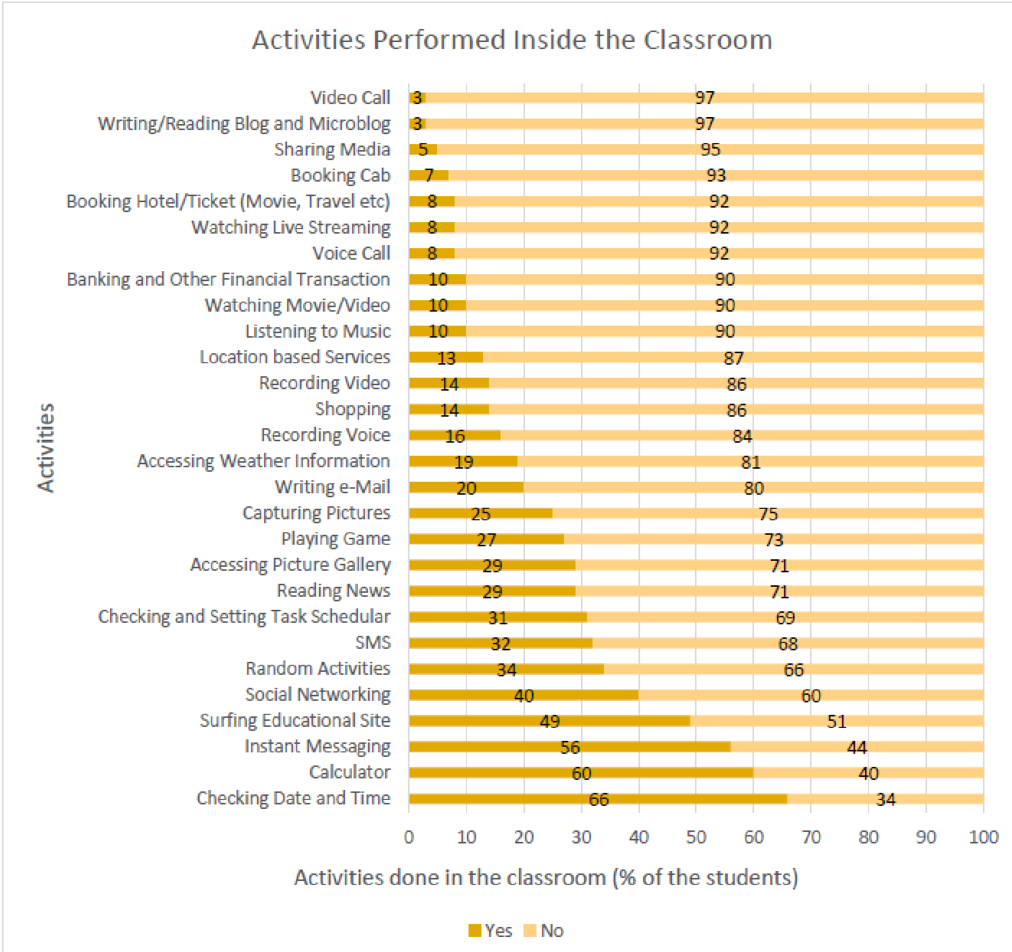


Figure 4.10: Overall usage statistics of smartphone activities inside classrooms

have reported that nothing is done by them other than the specified twenty-nine activities. However, some students (among the rest 25%) have mentioned that they perform various other activities. The details of the other activities are as follows. 47 students (3% of the total participants) have reported that they read e-books (academic e-books (1.5%), non- academic e-books including novels and magazines (1.5%)). 30 students (2% of the total participants) have reported that they watch porn on their smartphones. 26 students (1.5% of the total participants) have informed that they use the device to set alarms. A fraction of the total participating students have reported that they use the devices for health monitoring, photo and video editing, participating in discussion forums including 'Quora', and use the smartphone as a torch. Very few students (0.5% of the total participants) have reported that they also develop apps on their smartphones. Many smaller fractions of the total participating students (2-5 students) have reported various other activities including checking the stock market, paying bills and maintaining a digital wallet, playing digital musical instruments, using dating apps, using a dictionary app, using as a stopwatch, searching jobs, and writing the diary. However, as per the report of the students, they perform these activities rarely. Moreover, none of these students has reported that these activities are performed inside the classroom, except the two students who reported that they seldom use a dictionary inside the classroom.

4.4.2 Behavioral Differences based on Academic Level, Gender, and Ethnicity

In addition to identifying the frequencies, we also analyzed the behavioral difference in performing smartphone activities based on students' academic class, gender, and ethnicity. The details of the analyses are presented as follows.

A. Gender-based Differences

After analyzing the data, we have found some gender-based differences (see Table 4.1) on smartphone usage as follows. Male students have been observed to check date & time, do random activities, perform banking & other financial transactions, perform instant messaging, perform social networking, play games, read the news, surf educational sites, and write e-mails more than female students. On the other hand, the female students have been observed to access picture galleries, send SMSs, and use the mobile camera to take pictures more frequently than the male students. Significant differences in the behavior of males and females have been found for all of these activities.

B. Academic Level based Differences

We have analyzed this aspect in two dimensions: comparing the behavioral difference of school students with college & university students, and comparing the undergraduate students with the postgraduate students. It can be noted that, in India, generally 'school students' means the students of the standard of I to XII; whereas 'college students' refers to the undergraduate (UG) students, and 'university students' indicates the postgraduate (PG) students which include masters as

Table 4.1: Differences in behavior based on gender

Gender Compared		**Differences in frequency of performing the activities *(Male means male students, and female means female students)***	**Significance difference has been found while tested statistically *(Pearson's Chi-Square Test with α=0.05)***
Male & Female	**In general**	**Male perform more frequently than Female:** Playing game, instant messaging, social networking, surfing educational sites, writing email, reading news, banking & other financial transactions, checking date & time, random activities.	Playing game (p=0.0008), instant messaging (p=0.00005), social networking (p=0.0001), surfing educational sites (p=0.000005), writing email (p=002), reading news (p=3.84×10^{-07}), banking & other financial transactions (p=2.71×10^{-16}), checking date & time (p=0.04), random activities (p=0.01).
		Female perform more frequently than Male: SMS, capturing pictures, accessing picture gallery.	SMS (p=0.0004), capturing pictures (p=0.001), accessing picture gallery (p=0.0008).
	Inside classroom	**Male perform more frequently than Female:** SMS, playing game, instant messaging, surfing educational sites, writing email, reading news, banking & other financial transactions, checking date & time, random activities.	Playing game (p=0.0001), instant messaging (p=0.00005), social networking (p=0.001), reading news (p=0004), banking & other financial transactions (p=005), checking date & time (p=0.00003).
		Female perform more frequently than Male: Listening to music.	Listening to music (p=005).

well as doctoral students. The two levels of analyses based on the academic class are important because many school students do not use smartphones during school hours (however, we wanted to know their preferences, if allowed). It is very important to know their preferences as these indicate their habits and future addiction to smartphone activities, which consequently affects their social and personal life [75]. The observed dissimilarities of smartphone usage by the students based on their academic levels are reported as follows.

(a) Difference between school students and college (& university) students – It can be observed from Table 4.2 that the school students want to play games, record video, record voice, send and receive SMS, share media, and watch live streaming more frequently than the college & university students actually perform those (for school students, we asked their willingness to do the various activities inside the classroom as many times they are not allowed to use smartphones inside the classroom). On the other hand, the college & university students access picture gallery, check and set task scheduler, check date & time, perform banking & other financial transactions, perform instant messaging, perform random activities, do social networking, make video calls, make voice call, read news, surf educational site, and write email more frequently than the school students. However, the significant differences have been found in the case of accessing picture gallery, banking

& other financial transaction, checking date & time, checking & setting task scheduler, instant messaging, playing games, random activities, reading news, recording video, recording voice, sharing media, SMS, social networking, surfing educational sites, video call, voice call, watching live streaming, writing emails; while tested statistically.

Inside the classroom, the following activities are performed more frequently by the UG and PG students compared to the school students wished to be performed: accessing picture gallery, banking & other financial transaction, calculator, checking & setting task scheduler, checking date & time, instant messaging, playing games, random activities, reading news, recording voice, sending SMS, social networking, surfing educational sites, and writing emails. The result of the statistical test depicts that significant differences have been found ($p \leq 0.04$) for all these activities except banking and other financial transactions. On the other hand, the following activities are wished to be performed by the school students more frequently than the college and university students actually perform: listening to music, recording video, sharing media, video call, voice call, and watching live streaming. Among these activities, significant differences have been found only in the case of listening to music, making video call, and voice calls. Possible reasons for the behavioral difference may include underage of school students, restriction of bringing smartphones in the school (although we collected the willingness of the students for performing the activities inside the classroom, the actual activities performed by them may not be the same when they will be allowed).

(b) Difference between UG and PG students - We have also observed some differences in the behavior of UG students from that of PG students. The UG students are observed to listen to music, perform random activities, perform social networking, play games, record video, and use the device as a calculator more frequently than the PG students. On the other hand, the PG students are observed to do shopping, perform banking & other financial transactions, read the news, and write e-mails more than the UG students. Differences were found to be significant for all these four activities.

Inside the classroom, the UG students capture picture, check date & time, perform random activities, listen to music, perform banking and other financial

Table 4.2: Differences in performing activities based on the academic level

Academic Level Compared		**Differences in frequency of performing the activities *(School means school students, and college means college students)***	**Significance difference has been found while tested statistically *(Pearson's Chi-Square Test with α=0.05)***
School & College (and University)	**In general**	**School want to perform more frequently than College performs:** SMS, playing games, sharing media, recording video, recording voice, watching live streaming.	SMS (p=0.00002), playing game (p=5.54×10^{-10}, watching live streaming (p=1.54×10^{-10}), sharing media (p=1.17×10^{-12}), recording video (p=2.40×10^{-16}), recording voice (p=9.52×10^{-27}).
		College perform more frequently than School want to perform: Voice call, video call, instant messaging, social networking, surfing educational site, writing email, reading news, accessing picture gallery, banking & other financial transactions, checking date & time, checking and setting task scheduler, random activities.	Voice call (p=5.14×10^{-40}), video call (p=9.39×10^{-12}), instant messaging (p=3.99×10^{-86}), social networking (p=2.72×10^{-34}), surfing educational site (p=3.01×10^{-39}), writing email (p=8.67×10^{-47}), reading news (p=1.14×10^{-33}), accessing picture gallery (p=9.00×10^{-16}), banking & other financial transaction (p=2.99×10-49), checking date & time (p=2.51×10^{-20}), checking & setting task scheduler (p=8.25×10^{-08}), random activities (p=8.40×10^{-36}).
	Inside Classroom	**School want to perform more frequently than College perform:** Voice call, video call, listening to music, watching live streaming, sharing media, recording video.	Voice call (p=0.02), video call (p=0.00003), listening to music (p=0.00007)..
		College perform more frequently than School want to perform: SMS, playing game, instant messaging, social networking, surfing educational site, writing email, reading news, accessing picture gallery, banking & other financial transaction, recording voice, checking date & time, checking & setting task scheduler, calculator, random activities.	SMS (p=0.00004), playing game (p=2.66×10^{-10}), instant messaging (p=1.88×10^{-53}), social networking (p=1.03×10^{-24}), surfing educational site (p=1.31×10^{-11}), writing email (p=1.22×10-910^{-09}), reading news (p=2.09×10^{-12}), accessing picture gallery (p=1.29×10^{-11}), checking date & time (p=1.10×10^{-50}), checking & setting task scheduler (p=3.66×10^{-14}), calculator (p=1.95×10^{-30}), random activities (p=6.06×10^{-17}).
UG & PG	**In general**	**UG perform more frequently than PG:** Playing games, listening to music, social networking, recording video, calculator, random activities.	Playing games (p=6.68×10^{-21}), listening to music (p=7.02×10^{-16}), social networking (p=0.0001), recording video (p=0.00002), calculator (p=0.01), random activities (p=0.00001).
		PG perform more frequently than UG: Writing email, reading news, shopping, banking & other financial transactions.	Writing email (p=0.0002), reading news (p=0.0005), banking and other financial transactions (p=0.01), shopping (p=0.01).
	Inside classroom	**UG perform more frequently than PG:** Playing game, listening to music, instant messaging, social networking, surfing educational site, reading news, capturing picture, banking and other financial transaction, recording video, recording voice, checking date & time, calculator, random activities, taking class note.	Playing game (p=3.62×10^{-32}), listening to music (p=9.19×10^{-07}), instant messaging (p=2.17×10^{-10}), social networking (p=8.54×10^{-17}), surfing educational site (p=9.61×10^{-08}), reading news (p=0.000002), capturing picture (p=5.87×10^{-11}), accessing picture gallery (p=0.00001), recording video (p=5.17×10^{-07}), recording voice (p=0.001), checking date & time (p=2.31×10^{-08}), calculator (p=1.95×10^{-08}), random activities (p=3.98×10^{-11}), taking class note (p=0.00005)
		PG perform more frequently than UG: None.	None.

Table 4.3: Regions in India

Sl.No.	Region	States (provinces) and *union territories* fall in the region
1.	South	Andhra Pradesh, Karnataka, Kerala, Tamilnadu, Telengana, *Lakshadweep*, *Puduchery*.
2.	North	Haryana, Himachal Pradesh, Jammu & Kashmir, Punjab, Uttar Pradesh, Uttarakhand, *Chandigarh*, *Delhi*.
3.	East	Bihar, Jharkhand, Odisha, West Bengal,, *Andaman & Nicobar Islands*.
4.	West	Goa, Gujarat, Maharastra, Rajasthan, *Dadra & Nagar Havely*, *Daman & Diu*.
5.	Central	Chhattisgarh, Madhya Pradesh.
6.	North-East	Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, Tripura.

transactions, perform instant messaging, perform social networking, play games, read the news, record video, record voice, surf educational sites, take class notes, and use the device as calculator more than the PG students. Significant differences have been found for all these activities other than banking & other financial transactions. There is no activity that is performed more frequently inside the classroom by the PG students compared to the UG students.

C. Ethnicity based Differences

We have analyzed the behavior of the students based on their ethnicity broadly in two ways: comparing the behavior of the students of different regions of India and comparing the behavior of the Indian students with foreign students.

(a) Differences in behavior among the students of various regions of India – At the time of analyzing the behavior of the students of various geographical locations, we have observed that it was very similar while we compared the neighbor states (provinces). This may be the reflection of cultural similarities between the states. Moreover, we got very few students from some of the provinces. Reasons for this may include the small size of the provinces, less number of schools/colleges, problem in the medium of communication, and lack of sufficient infrastructure in the institutes of those states.

We, therefore, were interested in classifying all the provinces and union territories of the country into six regions (based on the culture as well as geographical location) to observe the differences in behaviors based on these regions. Table 4.3 demonstrates the region details whereas Figure 4.11 presents the region-wise distribution of the participants. It may be observed that the participants of different regions were well distributed. The results, after analyzing the behavior on smart-

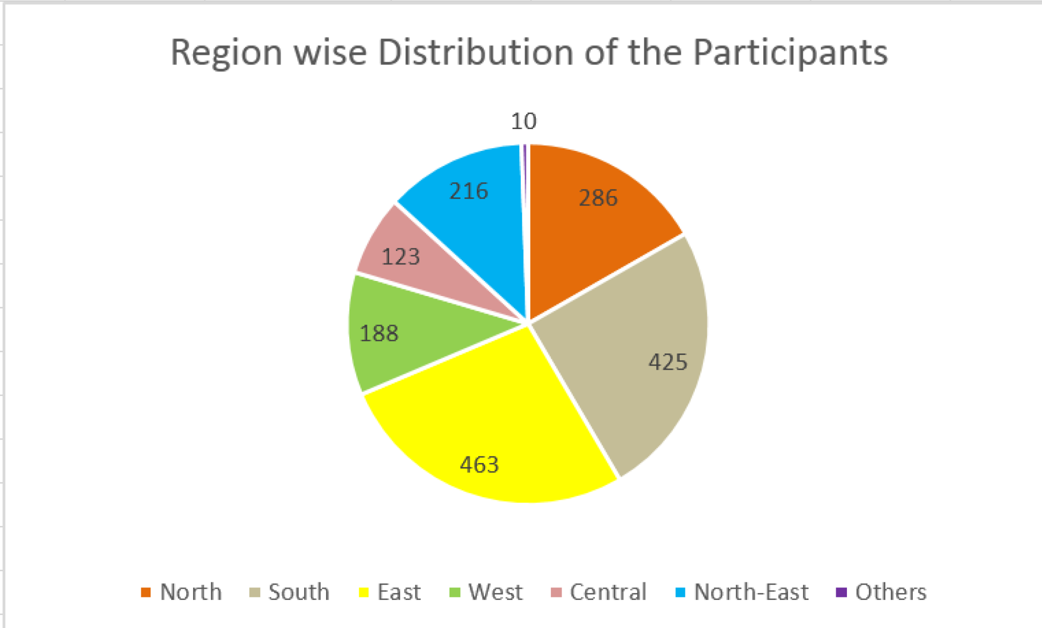


Figure 4.11: Distribution of participants from different regions of the country

phone usage of the students of various regions of India, are presented in Table 4.4.

Table 4.4: Region-wise differences in performing smartphone activities by Indian students

Regions Compared		Differences in frequency of performing the activities *(South refers to south Indian students, East indicates East Indian students, and so on)*	Significance difference has been found while tested statistically *(Pearson's Chi-Square Test with α=0.05)*
South & Central	***In general***	**South perform more frequently than Central:** Playing Game, shopping, banking & other financial transactions, booking cab.	Playing game (p=0.04), booking cab (p=0.04).
		Central perform more frequently than South: Video call, writing mail, taking class note.	Taking class note (p=0.00005).
	Inside classroom	**South perform more frequently than Central:** SMS, playing game, accessing picture gallery, recording voice, accessing weather information, booking cab, random activities.	Playing game (p=0.046), recording voice (p=0.039), accessing weather information (p=0.049), booking cab (p=0.007), random activities (p=0.03).

Regions Compared		**Differences in frequency of performing the activities** ***(South refers to south Indian students, East indicates East Indian students, and so on)***	**Significance difference has been found while tested statistically** ***(Pearson's Chi-Square Test with α=0.05)***
		Central perform more frequently than South: None.	None.
South & East	*In general*	**South perform more frequently than East:** Instant messaging, surfing educational site, reading news, accessing picture gallery, banking & other financial transactions, checking date & time, checking & setting task scheduler, random activities.	Instant messaging (p=2.92×10-11), surfing educational site (p=2.42×10^{-08}), reading news (p=0.001), accessing picture gallery (p=0.003), banking & other financial transactions (p=3.66×10^{-07}), checking date & time (p=0.003), random activities (p=1.02×10^{-06}).
		East perform more frequently than South: Video call, watching live streaming, sharing media, writing/reading blog/microblog, recording video, recording voice, taking class note.	Video call (p=0.000002), watching live streaming (p=0.008), sharing media (p=0.008), writing/reading blog/microblog (p=0.01), recording voice (p=0.01), taking class note (p=0.00003).
	Inside classroom	**South perform more frequently than East:** SMS, playing game, instant messaging, social networking, surfing educational site, writing email, reading news, accessing picture gallery, banking & other financial transaction, recording voice, checking date & time, checking & setting task scheduler, calculator, random activities.	SMS (p=0.000005), playing game (p=5.94×110^{-16}), instant messaging (p=2.08×10^{-12}), social networking (0.00003), surfing educational site (p=2.03×10^{-07}), writing email (p=0.01), reading news (p=3.71×10^{-09}), accessing picture gallery (p=5.37×10^{-08}), banking & other financial transaction (p=0.006), recording voice (p=0.004), checking date & time (p=3.95×10^{-13}), calculator (p=1.15×10^{-08}), random activities (p=1.92×10^{-09}).
		East perform more frequently than South: Capturing picture.	Capturing picture (p=0.04).

Regions Compared		Differences in frequency of performing the activities ***(South refers to south Indian students, East indicates East Indian students, and so on)***	Significance difference has been found while tested statistically ***(Pearson's Chi-Square Test with α=0.05)***
South & North-East	***In general***	**South perform more frequently than North-East:** Voice call, instant messaging, reading news, banking & other financial transaction, checking date & time, checking & setting task scheduler, random activities.	Voice call (p=0.000001), instant messaging (p=2.23×10^{-08}), reading news (p=0.01), banking & other financial transaction (p=1.52×10^{-07}), checking date & time (p=0.003), random activities (p=0.007).
		North-East perform more frequently than South: Video call, SMS, capturing picture, watching live streaming, sharing media, reading/writing blog/microblog, recording video, recording voice, accessing weather information.	Video call (p=0.02), watching live streaming (p=0.01), sharing media (p=0.0001), reading/writing blog/microblog (p=0.01), recording voice (p=0.00003).
	Inside classroom	**South perform more frequently than North-East:** SMS, playing game, instant messaging, social networking, surfing educational site, writing e-mail, reading news, accessing picture gallery, banking & other financial transactions, recording voice, accessing weather information, checking date & time, checking & setting task scheduler, calculator, random activities.	SMS (p=0.002), playing game (p=4.37×10^{-08}), instant messaging (p=3.59×10^{-10}), social networking (p=0.000002), surfing educational site (p=0.003), writing e-mail (p=0.0002), reading news (p=0.00002), accessing picture gallery (p=0.0007), banking & other financial transactions (p=0.02), checking date & time (p=2.48×10^{-14}), checking & setting task scheduler (p=0.000004), calculator (p=1.98×10^{-09}), random activities (p=0.000003).
		North-East perform more frequently than South: Taking class note.	Taking class note (p=6.13×10^{-09}).
South & North	***In general***	**South perform more frequently than North:** SMS, playing game, recording video, accessing weather information, calculator, random activities.	Playing game (p=0.0003), accessing weather information (p=0.01), calculator (p=0.001).

Regions Compared		**Differences in frequency of performing the activities** ***(South refers to south Indian students, East indicates East Indian students, and so on)***	**Significance difference has been found while tested statistically** ***(Pearson's Chi-Square Test with α=0.05)***
		North perform more frequently than South: Social networking.	None.
	Inside classroom	**South perform more frequently than North:** Playing game, reading news, recording video, calculator, random activities.	Playing game (p=0.00005).
		North perform more frequently than South: Social networking.	None.
South & West	*In general*	**South perform more frequently than West:** Video call, SMS, playing game, watching live streaming, banking & other financial transactions, accessing weather information, location based services, booking cab, checking & setting task scheduler, random activities.	Banking & other financial transactions (p=0.0004), location based services (p=0.004), random activities (p=0.047).
		West perform more frequently than South: None.	None.
	Inside classroom	**South perform more frequently than West:** SMS, playing game, accessing picture gallery, banking & other financial transactions, recording video, recording voice, accessing weather information, location based services, booking cab, checking & setting task scheduler, random activities.	SMS (p=0.0007), playing game (p=0.02), accessing picture gallery (p=0.005), banking & other financial transactions (p=0.049), recording video (p=0.02), recording voice (p=0.004), accessing weather information (p=0.03), location based services (p=0.02), booking cab (p=0.003), checking and setting task scheduler (p=0.01).
		West perform more frequently than South: Taking class note.	None.
North & Central	*In general*	**North perform more frequently than Central:** None.	None.

Regions Compared		Differences in frequency of performing the activities ***(South refers to south Indian students, East indicates East Indian students, and so on)***	Significance difference has been found while tested statistically ***(Pearson's Chi-Square Test with α=0.05)***
		Central perform more frequently than North: SMS, capturing picture, accessing picture gallery, accessing weather information.	None.
	Inside classroom	**North perform more frequently than Central:** SMS, writing email, accessing weather information.	None.
		Central perform more frequently than North: Taking class note.	Taking class note (p=0.02).
North & East	***In general***	**North perform more frequently than East:** Voice call, instant messaging, social networking, surfing educational site, reading news, checking date & time, random activities.	Voice call (p=0.01), instant messaging (p=1.1×10-07), social networking (p=0.0008), surfing educational site (p=0.000003), reading news (p=0.0008), checking date & time (p=0.004), random activities (p=0.00003),
		East perform more frequently than North: Video call, SMS, playing game, watching live streaming, sharing media, recording video, recording voice, accessing weather information.	Video call (p=0.01), playing game (p=0.0008), sharing media (p=0.005), recording video (p=0.002), recording voice (p=0.003), accessing weather information ($p=2.48\times10^{-07}$).
	Inside classroom	**North perform more frequently than East:** SMS, playing game, instant messaging, social networking, surfing educational site, writing email, reading news, accessing picture gallery, recording voice, checking date & time, calculator, random activities.	SMS (p=0.001), playing game (p=0.002), instant messaging ($p=1.3\times10^{-12}$), social networking ($p=5.7\times10^{-07}$), surfing educational site (p=0.00005), writing email (p=0.002), reading news (p=0.0005), accessing picture gallery (p=0.004), checking date and time ($p=2.65\times10^{-09}$), calculator (p=0.0006), random activities (p=0.00007).
		East perform more frequently than North: None.	None.

Regions Compared		Differences in frequency of performing the activities *(South refers to south Indian students, East indicates East Indian students, and so on)*	Significance difference has been found while tested statistically *(Pearson's Chi-Square Test with α=0.05)*
North & North-East	*In general*	**North perform more frequently than North-East:** Voice call, instant messaging, social networking, reading news, banking & other financial transactions, checking date & time, checking & setting task scheduler, random activities.	Voice call (0.000001), instant messaging (p=0.00002), social networking (0.006), reading news (009), banking & other financial transaction (0.00002), checking date & time (0.002), taking & setting task scheduler (p=0.03), random activities (p=0.02).
		North-East perform more frequently than North: Video call, SMS, playing game, capturing picture, watching live streaming, sharing media, reading/writing blog/microblog, recording video, recording voice, accessing weather information.	SMS (p=0.01), playing game (p=0.001), capturing picture (0.003), watching live streaming (p=0.007), sharing media (p=0.0004), writing/reading blog/microblog (0.005), recording video (p=0.00008), recording voice (p=0.00001), accessing weather information (0.0002).
	Inside classroom	**North perform more frequently than North-East:** SMS, playing game, instant messaging, social networking, surfing educational site, writing e-mail, reading news, checking date & time, checking & setting task scheduler, calculator, random activities.	SMS (p=0.04), instant messaging (p=8.02$\times 10^{-11}$), social networking (p=6.03$\times 10^{-08}$), surfing educational site (p=0.02), writing e-mail (p=0.00006), reading news (p.01), checking date & time (p=5.53$\times 10^{-11}$), checking & setting task scheduler (p=0.001), calculator (p=0.00004), random activities (p=0.001).
		North-East perform more frequently than North: Listening to music, recording video, taking class note.	Listening to music (p=0.001), taking class note (p=0.002).
North & West	*In general*	**North perform more frequently than West:** Video call.	None.
		West perform more frequently than North: None.	None.
	Inside classroom	**North perform more frequently than West:** SMS.	SMS (p=0.01)

Regions Compared		Differences in frequency of performing the activities ***(South refers to south Indian students, East indicates East Indian students, and so on)***	Significance difference has been found while tested statistically ***(Pearson's Chi-Square Test with α=0.05)***
		West perform more frequently than North: None.	None.
East & Central	***In general***	**East perform more frequently than Central:** Playing game, reading/writing blog/microblog, recording video.	Playing game (p=0.03), recording video (p=0.01).
		Central perform more frequently than East: Voice call, instant messaging, social networking, surfing educational sites, reading news, accessing picture gallery, checking date & time, checking & setting task scheduler, random activities.	Voice call (p=0.02), instant messaging (p=0.0007), social networking (p=0.04), surfing educational sites (p=0.002), reading news (p=0.002).
	Inside classroom	**East perform more frequently than Central:** Writing/reading blog/microblog.	None.
		Central perform more frequently than East: Voice call, playing game, instant messaging, social networking, surfing educational sites, reading news, accessing picture gallery, checking date & time, checking & setting task scheduler, calculator, random activities.	Voice call (p=0.04), playing game (p=0.0003), instant messaging (p=8.3$\times10^{-07}$), social networking (p=0.001), reading news (p=0.005), checking date & time (p=0.000003), calculator (p=0.01).
East & North-East	***In general***	**East perform more frequently than North-East:** Banking & other financial transactions.	Banking & other financial transactions (p=0.049).
		North-East perform more frequently than East: Playing game, capturing picture, recording video.	None.
	Inside classroom	**East perform more frequently than North-East:** Checking and setting task scheduler.	None.

Regions Compared		**Differences in frequency of performing the activities** ***(South refers to south Indian students, East indicates East Indian students, and so on)***	**Significance difference has been found while tested statistically** ***(Pearson's Chi-Square Test with α=0.05)***
		North-East perform more frequently than East: Capturing picture, recording video.	Capturing picture (p=0.02), recording video (p=0.03).
East & West	*In general*	**East perform more frequently than West:** Video call, SMS, watching live streaming, sharing media, reading/writing blog/microblog, recording video, recording voice, accessing weather information, location based services.	Video call (p=2.02×10^{-07}), SMS (p=0.01), watching live streaming (p=9.77×10^{-07}), sharing media (p=0.01), reading/writing blog/microblog (p=0.008), recording video (p=0.04), recording voice (p=0.01), accessing weather information (p=0.0002), location based services (p=2.72×10^{-09}).
		West perform more frequently than East: Instant messaging, social networking, surfing educational site, reading news, checking date & time, random activities.	Instant messaging (p=0.000005), social networking (p=0.03), surfing educational site (p=0.0001), reading news (p=0.001), checking date & time (p=0.0008), random activities (p=0.02).
	Inside classroom	**East perform more frequently than West:** Location based service.	Location based service (p=0.047).
		West perform more frequently than East: Playing game, instant messaging, social networking, surfing educational site, reading news, checking date & time, calculator, random activities.	Playing game (p=0.00003), instant messaging (p=4.75×10^{-08}), social networking (p=0.003), surfing educational site (p=0.003), reading news (p=0.0003), checking date & time (p=1.31×10^{-07}), calculator (p=0.0007), random activities (p=0.0002).
West & Central	*In general*	**West perform more frequently than Central:** Playing game.	None.
		Central perform more frequently than West: Video call, SMS, listening to music, watching live streaming, checking & setting task scheduler.	None.

Regions Compared		**Differences in frequency of performing the activities** ***(South refers to south Indian students, East indicates East Indian students, and so on)***	**Significance difference has been found while tested statistically** ***(Pearson's Chi-Square Test with α=0.05)***
	Inside classroom	**West perform more frequently than Central:** None.	None.
		Central perform more frequently than West: SMS, checking & setting task scheduler, taking class note.	None.
West & North-East	***In general***	**West perform more frequently than North-East:** Voice call, instant messaging, social networking, reading news, banking & other financial transactions, checking date & time, checking & setting task scheduler, random activities.	Voice call (p=0.000003), instant messaging (p=0.0001), reading news (p=0.01), banking & other financial transactions (p=1.07$\times10^{-07}$), checking date & time (p=0.002).
		North-East perform more frequently than West: Video call, SMS, playing game, listening to music, capturing picture, sharing media, reading/writing blog/microblog, recording video, recording voice, accessing weather information.	Video call (p=0.003), SMS (p=0.03), listening to music (p=0.04), capturing picture (p=0.02), sharing media (p=0.04), reading/writing blog/microblog (p=0.01), recording video (p=0.02), recording voice (p=0.000005), accessing weather information (p=0.004).
	Inside classroom	**West perform more frequently than North-East:** Playing game, instant messaging, social networking, surfing educational site, writing email, reading news, checking date & time, checking & setting task scheduler, calculator, random activities.	Playing game (p=0.006), instant messaging (p=1.76$\times10^{-07}$), social networking (p=0.0001), writing email (p=0.02), reading news (p=0.006), checking date & time (p=2.19$\times10^{-09}$), calculator (p=0.0003), random activities (p=0.0002).
		North-East perform more frequently than West: Listening to music, recording video, recording voice, taking class note.	Recording video (p=0.02), taking class note (p=0.005).

Regions Compared		**Differences in frequency of performing the activities** ***(South refers to south Indian students, East indicates East Indian students, and so on)***	**Significance difference has been found while tested statistically** ***(Pearson's Chi-Square Test with α=0.05)***
Central & North-East	***In general***	**Central perform more frequently than North-East:** Voice call, instant messaging, social networking, surfing educational site, writing email, reading news, checking date & time, checking & setting task scheduler, calculator, random activities.	Voice call (p=0.0001), instant messaging (p=0.002), writing email (p=0.008), reading news (p=0.01).
		North-East perform more frequently than Central: Video call, SMS, playing game, sharing media, writing/reading blog/microblog, recording video, recording voice, booking cab.	Recording video (p=0.02).
	Inside classroom	**Central perform more frequently than North-East:** Voice call, SMS, playing game, instant messaging, social networking, surfing educational sites, writing email, reading news, checking date & time, checking & setting task scheduler, calculator, random activities.	Voice call (p=0.02), playing game (p=0.01), instant messaging (p=0.000001), social networking (p=0.00005), writing email (p=0.04), reading news (p=0.03), checking date & time (p=6.63×10^{-08}), checking & setting task scheduler (p=0.003), calculator (p=0.001).
		North-East perform more frequently than Central: Video call, listening to music, sharing media, writing/reading blog/microblog, recording video, recording voice, booking cab.	None.

It may seem that the East Indian students are more sincere compared to the Central Indian students, as they perform non-academic activities (e.g., playing games) in the classroom less than the Central Indian students. However, in general, the East Indian students play more games than the Central Indian students. Nonetheless, the Central Indian students surf educational sites and use the calculator app inside the classroom more than the East Indian students. These indicate that the central Indian students are more fascinating to use their smartphones for both academic and non-academic purposes.

The East Indian students perform many activities more than the West Indian students in general. However, the West Indian students perform these particular activities more than the East Indian students inside the classrooms. Most of these are non-academic. The observation indicates that the East Indian students are more sincere in the classroom. However, the West Indian students surf educational sites more than the East Indian students, both inside and outside the classroom. The West Indian students also use their smartphones as a calculator more than the East Indian students. These indicate that, in general, the West Indian students are more fascinating towards using the smartphone compared to East Indian students.

North-East Indian students perform some non-academic activities (e.g., SMS) more frequently than the North Indian students. However, the same activities are performed by the North Indian students more than the North-East Indian students inside the classroom. Moreover, the North-East Indian students take class notes more than North Indian students. Furthermore, in general, the North Indian students are found to have a tendency of using non-academic activities inside the classroom. A similar observation has been made for South Indian students as well. The observations hint that North and South Indian students are probably less sincere in the classroom. However, research on the topic is required before drawing a conclusion regarding this.

The meta-analysis of the region-wise analyses of the behavior of the students indicates the following.

- South and North Indian students are found to have more inclination to perform non-academic activities inside the classroom compared to the students of any of the other regions.

- The behavior of the North, West, and Central Indian students on performing smartphone activities are similar, both inside the classroom and in general.

- The behavior of the East and North-East Indian students on performing smartphone activities is almost similar.

- North-East Indian students prefer to perform media-related activities, especially ‘video recording’, most.

- East and North-East Indian students seem to be remarkably sincere in the classroom, as they rarely perform non-academic activities inside the classroom.

(b) Differences in behavior between Indian and foreign students – The sole purpose of our study was to analyze the behavior of Indian students in performing smartphone activities. Hence, we have collected data from only Indian students. However, we were also interested to observe if there is any behavioral difference between Indian and foreign students. We, therefore, compared the findings of the contemporary related studies conducted in other countries, with the findings of our studies. The comparison results have been presented in Table 4.5.

We summarize the findings after analyzing the survey results to observe the behavioral differences based on students' academic level, gender, and ethnicity in Table 4.6.

4.5 Discussion

4.5.1 Critical Analysis of Results

In the study, we have found significant gender-based differences for two smartphone activities: playing games and reading news. Male students are found to perform these activities more compared to female students. Further research is required to find out the reasons for this lack of interest of the particular gender in these particular activities, and their interest in other activities like using the phone camera to capture pictures and accessing the picture galleries. One of the probable reasons for playing fewer games by the females might be that the presence of a fewer number of interesting elements for them in most of the popular existing games in the app-stores. Our study also finds that the girls are found to be more sincere, they perform non-academic activities in the classroom less than the boys. It has also been found that the girls perform less financial transactions than the boys. The

Table 4.5: Dissimilarities in smartphone usage behavior between Indian and foreign students

Compared with	Based on the work of	Dissimilarities found in	Details of dissimilarities
Australia, Hong Kong, Portugal, Sweden, and UK	Kukulska-Hulme et al. 2011 [70]	Most frequently performed activity	**In Australia, Hong Kong, Portugal, Sweden, and UK:** SMS. **In India:** Instant messaging.
		Top 3 activities	**In Australia, Hong Kong, Portugal, Sweden, and UK:** SMS, browsing website, and listening to music (in descending order). **In India:** Instant messaging, voice call, and checking date and time (in descending order).
		Least frequently performed activity	**In Australia, Hong Kong, Portugal, Sweden, and UK:** Booking hotel/ticket (movie, travel) is not found in the list of least frequently performed activities. **In India:** Booking hotel/ticket (movie, travel) is the second least frequently performed activities.
Korea	Park, 2014 [100]	Gender bias	**In Korea:** (I) Male are more addictive to SMS than female; (II) males are more addicted to listen to music than females; (III) female are more addictive to social networking apps than male. **In India:** (I) Females are more addicted to SMS than males; (II) addiction of listening to music on smartphone is similar, sometimes (inside classroom) females listen to music more than female; (III) male are more addictive to social networking apps than mail.
Australia	Roberts & Rees, 2014 [110]	Top 3 activities	**In Australia:** Texting (not mentioned whether it is SMS or instant messaging), social networking (Facebook), email (in descending order). **In India:** Instant messaging, voice call, and checking date and time (in descending order).
Florida, US	Schroeder et al., 2016 [114]	Texting	**In Florida:** College students do text more frequently than middle schoolers; however, college students' texting behavior decreases in frequency as they get older. **In India:** School students use SMS more than college students, whereas college students use instant messaging more than school students. However, texting behavior does not decrease in frequency as they get older - PG students use instant messaging more than UG students.

girls may be afraid of the security risks associated with these. At the time of investigating the behavioral differences between college and school students, we have observed that the school students send SMS, share media, and watch live streaming significantly more compared to the college students. On the other hand, the UG & PG students perform social networking and random activities more compared to the school students. Probable reasons for these may include restricted usage of the devices (many parents in India do not allow their school-going children to use smartphones on a regular basis), less connection with the outer world, and lack of openness due to their underage. The college students might also be habituated to use advanced texting apps (IMS – e.g., WhatsApp, WeChat) instead of SMS to communicate with a relatively larger number of friends, relatives, and other professional

Table 4.6: Summary of the analyses on behavioral differences based on the academic level, gender, and ethnicity

Observational aspects	Findings
Gender based differences in performing smartphone activities	I. Boys play games significantly more than girls, both inside the classroom and in general. II. Boys read the news significantly more than girls, both inside the classroom and in general. III. Girls are found to be more sincere in the classroom. They perform less non-academic smartphone activities inside the classroom.
Academic level based differences in performing smartphone activities	I. The school students watch live streaming and share media significantly more than college students. II. The college students perform social networking, and random activities significantly more than school students. III. The UG students watch live streaming, and record voice significantly more than the PG students. IV. The PG students write e-mails and do shopping significantly more than the UG students.
Ethnicity based differences in performing smartphone activities	I. Habits on performing smartphone activities of neighboring states are similar. However, many dissimilarities have been found while compared region-wise (exception: north and central region). II. South and North Indian students are habituated to perform non-academic activities inside the classroom most compared to the students of the other reasons. III. Like neighbor states, the behavior (in performing smartphone activities) may be similar for some neighbor regions as well. E.g., the behavior of the North, West, and Central Indian students are similar. The behavior of the East and North-East Indian students are also found to be similar. IV. Students from a particular region may be very fascinating to perform particular activities (e.g., North-East Indian students are most fascinating about 'recording video', compared to the students from any other regions). V. Differences in behavior on performing smartphone activities have also been found between Indian students and foreign students.

contacts. When we compared the UG students with the PG students to investigate the behavioral differences, it has been found that the PG students send and receive a lot of e-mail and do online shopping significantly more compared to the UG students. Reasons for sending and receiving more mail by PG students may include the requirement of staying in touch as well as formal interactions with a larger number of people for their advanced course, job applications, teaching assistantships, research guidance and collaborations, and so on. When students get older, they may become less dependent on their parents. This is because they sometimes start earning and many times they get fellowships in PG courses. Due to this relative financial prosperity, PG students may have a greater tendency to shop online more than the UG students. If we compare the habits of performing smartphone activities among the school, UG, and PG students, it can be noticed that the school students have a tendency to perform media-related activities (e.g., sharing media from one device to another via Bluetooth or WiFi, live streaming) more than the UG students whereas the UG students perform those activities more than the PG students. The probable reasons for this may include a gradual change in interests with the growing age [114]. When we analyzed the behavioral differences based on ethnicity, we observed that the behavior of students on smartphone activities is similar in the case of neighbor provinces. However, many differences in behavior have been found while compared the same for region wise. Probable reasons may include the cultural differences among the regions of such a big country (with a total area of 3,287,263 square kilometers). The cultural differences may be the reason for the differences in the behaviors of the India and foreign students as well. Other reasons for the country wise dissimilarities in the behavior may include differences in the technological advancements and financial status of the various counties.

4.5.2 Update of Existing Knowledge Base

Although our survey results corroborate some of the earlier findings (e.g., the gender-based difference in playing games, and academic class-based differences in performing media-related activities), the same contradict many aspects of the earlier findings. As per the study of Kukulska-Hulme et al. [70], SMS was the most frequently performed activity on smartphones. The study was conducted in four countries: the UK, Sweden, Portugal, Hong Kong, and Australia. The top three activities they found in their study were sending and receiving SMS, website browsing, and listening to music (presented in descending order). However, the results of our study contradict their study-results both for most frequently performed activities as well as the top three activities. According to our study, instant messaging is the most frequently performed activity and the top three activities are instant messaging, voice call, and checking date and time (presented in descending order). The reasons for these contradictory findings may include the launch of advanced texting service, relatively lower cost for Internet connection, and ethnicity. The dissimilarities have also been found in one of the least frequently performed activities. According to our study, hotel/ticket (movie, travel) booking is the second least frequently performed activity. However, the name of this activity was not found in the list of 'least performed activities' of Kukulska-Hulme et al [70]. The probable reason might be the ethnicity; Indian students may not be technologically and financially updated in that extent to date. Other reasons for this contradiction may include the consideration of only a specific group of participants in their study. They considered only mature students in their study.

Our results also contradict the findings of an Indian study conducted by Jena [61]. As per his finding, the use of smartphones for academic purposes is more than that of non-academic purposes. On contrary, our survey results depict that students use their smartphones largely for non-academic purposes. The reasons for such a difference might be the consideration of some specific participants in his study (only 310 PG students of business management discipline), which might not be a good representation of Indian students in general. Probably because of the same reason (consideration of the participants who are interested in cultural activities only), contradictory findings have been observed in a recent study [94] conducted in India. One such contradiction, for instance, is the difference in frequency of accessing social networking sites (e.g., Facebook, Twitter, and Instagram). In their study, it has been found that accessing social networking sites is the second most frequently performed activity. However, we have found that the same is the fifth most frequently performed smartphone activity by the Indian students.

We have observed some partial contradictions in the findings while compared with that of the study of Park [100]. Although the findings of gender-based differences in performing smartphone activities in his study are similar in various aspects, our findings primarily contradict with three such aspects. Firstly, they found that male students are more addicted to SMS than female students. On the contrary, we have found that female students are more addicted to SMS than male students. Secondly, they found that male students listen to music more than female students. On the other hand, we observed that the frequency of listening to music by the two genders is similar ($\sim$80% at least once/twice a day & $\sim$54% at least several times in a day). Not only that, as per our observation, sometimes (e.g., in the classroom) females are found to listen to music more than males. Thirdly, as per Park, females are more addicted to performing social networking apps. On contrary, we have found that males are more addictive to perform social networking apps than the female. Ethnicity, technological advancement, and changes in the socio-economic status of males and females over time maybe some of the possible reasons for these contradictory findings.

4.5.3 Utilization of Results in Developing Vedinkaksha

We have found fourteen smartphone activities that are frequently performed by the students inside the classroom. Some of these activities are study-related, whereas some of them are not related to the study. We have used this knowledge to build a computational model for Vedinkaksha. By observing the device handling patterns along with the resource utilization patterns and the changes in battery temperature, while the students perform these two groups of activities, we have built a machine learning model. This model, called 'In-Activity', is able to detect whether a student is involved in performing study-related activities or s/he is involved in performing non-study related activities. The details of the model can be found in **Chapter 6** (Section 6.5).

The findings after the academic level, gender, and ethnicity-based analyses of the survey results have not been directly applied in building the computational models for Vedinkaksha, but they helped in understanding many aspects during the development phase of the framework. They signify the importance of conducting the study on Indian students, as the behavior of the students may change based on the ethnicity of the student. These findings indicate that for building behavioral computational models, we should use the behavioral data of the target user group only. Otherwise, the models may misbehave. As the behavior may also be changed based on the academic level of the students, we may think about targeting specific student-group while building a model based on the behavioral data. For instance, if we build a model using the behavioral data of school students, it may not work for college students. It can be noted that we have used the behavioral data of college students (UG and PG students) for building the models which have been incorporated in Vedinkaksha. Also, as the results of the study show some gender-based differences in students' behavior, one may even think about retraining the models if they are applied for a particular gender. Nevertheless, we have developed the computational model using the behavioral data of the students without gender bias.

4.6 Chapter Summary

In this chapter, we have presented a behavioral study that helped us to acquire up-to-date knowledge about the behavior of Indian students on performing smartphone activities. We conducted the study through a systematic survey. We have presented the survey methodology and its result, along with the in-depth analysis of the same to observe the behavioral difference on smartphone usage based on students' gender, academic level, and ethnicity. The up-to-date knowledge about smartphone usage behavior has been utilized to build an AI-based computational model to detect the involvement status of the students, which is one of the key components of our proposed framework for a sensitive classroom system. The details of this model as well as all the other models that have been developed by us to propose Vedinkaksha, and the empirical studies conducted to build and validate these models are presented in the subsequent chapters.

"When you cannot measure, your knowledge is of meager and unsatisfactory kind."

Lord Kelvin (1824 – 1907)

British mathematician & physicist

5

Empirical Study Details

5.1 Introduction

The key components of Vedinkaksha are several novel user-centric computational models, which are presented in the next chapter (**Chapter** 6). We required empirical data to build and validate these models. It can be noted that the empirical data can be collected by observing the users while either they are involved in performing the tasks in the real-world environment, or they are indirectly forced to perform specific tasks in a laboratory environment. We collected the empirical data in the laboratory setting, i.e., through controlled experiments (CEs). We collected the data on our own because there is no existing data set present in the literature, which could be utilized to build and validate the novel models.

We conducted seven controlled experiments for collecting empirical data. The first empirical study was conducted to collect data for building a model to detect the involvement of the students in study-related activities. The second and third studies were conducted to collect data for building models for detecting the affective states of students from their touch and typing patterns on the touchscreen, respectively. In these two studies, we used two novel games for inducing the specific affective states and at the same time for labeling the data without user feedback. We conducted another controlled experiment to verify the induction capability of these two games using EEG data of the users. In the verification process, we assumed that the affective states identified by a commercial EEG device as the ground truth. We established the ground truth through another empirical study which was the fifth controlled experiment. A sixth controlled experiment was conducted to collect the EEG data for validating a process model that detects the affective states of the user from smartphone-user interaction data. Finally, to test the workability of the affect detection model in the target application we conducted the seventh controlled study. The details of all the controlled experiments, which have been conducted to develop the Vedinkaksha framework and its components are presented in the subsequent sections.

5.2 CE1: Data Collection for Building 'In-Activity' Model

We conducted this experiment to collect the sensory data which indicates the behavior of the students on handling the devices, along with the resource utilization data while they perform specific study and non-study related activities. We aimed to collect these data so that we could use them to build a computational model to detect the involvement of the students in study-related activities. The details of the data collection process are presented as follows.

5.2.1 Task Design

A. Identification of Frequently Performed Smartphone Activities

We identified the frequently performed smartphone activities from the results of a survey [125], which have been conducted by us and reported in the previous chapter (**Chapter** 4). It has been found that students perform fourteen activities frequently with their smartphones inside the classroom. The list of the identified fourteen activities is reported in Table 5.1.

B. Categorization of the Activities

It is difficult to categorize a particular activity into either study or non-study related activity. For instance, a student may watch a video for study purposes (e.g., s/he is watching a video lecture) as well as for non-study related purposes (e.g., s/he is watching a movie). However, it is very context-specific. Watching any video

Table 5.1: Frequently performed smartphone activities inside the classroom, along with their categories

Sl. No.	**Activity**	**Category (S=Study, NS=Non-Study related smartphone activity)**
1.	Checking date and time	NS
2.	Running calculator app	S
3.	Performing instant messaging	NS
4.	Surfing educational sites	S
5.	Performing social networking	NS
6.	Taking class note	S
7.	Performing random activities	NS
8.	Sending and receiving SMS	NS
9.	Performing task scheduling	NS
10.	Accessing picture gallery	NS
11.	Reading news	NS
12.	Playing games	NS
13.	Capturing picture	NS
14.	Writing e-mail	NS

while attending a lecture may be considered as non-study related activities, as the student will then miss the lecture being delivered by the instructor. At the same time, finding the meaning of a word or the answer to a question from the Internet might be considered as study-related activities. McCoy [89] conducted a survey to identify the names and frequencies of performing 'non-class' related activities during class-hours using digital devices such as smartphones, tablets, and laptops, along with the level of distractions caused by the activities. Roberts and Rees [110] investigated the usage of mobile devices for 'off task' and 'on task' in the context of classroom learning. The term 'on task' refers to study related activity, whereas the term 'off task' indicates the non-study related activity. We categorized the identified fourteen frequently performed activities into either study or non-study related activities (Table 5.1) based on these works ([89], [110]).

C. Task-set Specification

Among these fourteen activities, seven activities are performed using only native apps (smartphone applications). The seven apps are 'checking date and time', 'calculator', 'instant messaging', 'SMS', 'task scheduler', 'picture gallery' and 'camera'. By observing the system log, it is possible to know the names of the apps and hence its types (study/non-study), without accessing the actual contents. However, the rest of the activities (surfing educational sites, social networking, taking class notes, performing random activities, reading news, playing games, and writing e-mails) can be performed through either app or web browser. In case they are performed through the browser, it is not possible to know the name (and hence, the type of the activity) unless the details of the contents are accessed – which we did not want, as they may contain private data. In our study, we did not record the contents accessed by the students. We collected the readings of the sensors and resource utilization pattern while these seven activities (Sl. No. 4, 5, 6, 7, 11, 12, and 14 of Table 5.1) were performed.

Among these seven activities, 'surfing educational sites' was difficult to specify in the controlled experiment. Someone may simply use a website to know the meaning of a word, whereas someone may explore a particular topic in detail or try to find out the answer to a question in the educational sites. We, therefore, kept all these three types of sub-tasks for the 'surfing educational site' task. The three subtasks were 'finding the meaning of two words', 'finding the answers to two short questions', and 'finding the answer to a conceptual question' – which required the exploration of a topic at a deeper level. For the 'social networking' task, we asked the participants to do the same as they usually do while performing the social networking activities in practice. We instructed the participants to note down the meaning of the two words and answers to the questions asked in task 1 (surfing educational sites) for 'taking class note' activity. We provided a link (URL) where they could write class-notes and save it for future reference. For performing random activities, participants were asked to perform some activities which they usually do in their idle time. We mentioned a few examples such as closing and opening different apps and browsers, checking for attractive apps in the 'Play Store', and so on. To control the 'reading news' activity, the participants were asked to report either the top five news or a brief description of the most interesting news of the day – choice of news-topics was open for them, as per their wish. We instructed the participants to play an online game for controlling the 'playing game' task. We provided the URLs of two popular gaming websites for this. However, they were informed that they could use any other gaming sites as per their choices. For performing the 'writing mail' task, the participants were asked to carefully read an e-mail, which was sent by an experimenter at the starting of the experiment and write a reply to it. Table 5.2 consolidates the set of tasks with their specifications which were supposed to be performed by the participants in the experiment.

Table 5.2: Task set for CE1 (data collection for In-Activity Model)

Sl. No.	**Task**	**Specification of the task**
1.	Surfing educational site	Three subtasks: a . Finding the meaning of two English words. b . Answering two short questions (e.g., to find the name of a pioneer of a particular research field) – does not take a long time to answer. c . Answering a broad question (e.g., to find the differences in two terms) – takes a long time and exploration to answer.
2.	Social networking	Same as they do usually on Facebook/Twitter/Instagram/any other social networking sites (for at least five minutes).
3.	Taking class note	Noting down the meanings of the words and answers to the questions asked in task-1 (surfing the educational sites). A URL was provided to facilitate the task.
4.	Performing random activities	Same as they do usually in their idle time (examples: opening and closing different applications, searching for attractive apps in the play store).
5.	Reading news	Two subtasks: a . Reporting the topic name of the top five news of the day. b . Briefing the most interesting news of the day. For both the subtasks (a. and b.), they were free to choose the news categories (e.g. sport/ politics/ education/ technology).
6.	Playing games	Playing an online game from any gaming sites (for at least five minutes). Some example sites were mentioned.
7.	Writing e-mail	Reading carefully an e-mail sent by an experimenter (at the time of commencement of the experiment) and replying to the mail.

5.2.2 Experimental Setup

We built a special android app for the data collection. The app continuously runs in the background and record sensor and resource utilization readings. We built the app to log the readings of two sensors embedded in the smartphone (accelerometer and gyroscope) along with the memory consumption and battery temperature. This is because we assumed (and confirmed by a pilot study) that using these readings we could differentiate the study and non-study related activities. In the pilot study, we asked five participants (two female, mean age = 25.5 years with SD=3.17) to perform each of the study and non-study related tasks for five minutes, and observed that that the patterns for utilizing some of the resources of the device and the sensory values may be different for study and non-study related activities [127].

We installed the special-purpose app on a smartphone ('Sony Xperia T2 Ultra Dual') to collect data in the controlled experiment. The smartphone had Android 4.3 OS, 1.4 GHz Quad-Core processor, 1 GB RAM, and a 6.00-inch display. Data connection was kept on for the device prior to the experiment to provide Internet access for performing the activities.

5.2.3 Participants

We wrote a formal e-mail to invite students for their voluntary participation in the controlled experiment. In the e-mail, we mentioned the type of activities they had to perform, the place of the experiment, and the expected time required. We also shared the link of a 'Google Sheet' to book the slots for their participation in the experiment, as per their convenience. In addition to inviting students through the mail, we also verbally approached some of the students for participation.

Sixty-five students agreed to participate in our experiment and booked slots for the same. Among them, twenty nine students finally participated in the experiment. In addition to these twenty-nine volunteers, we got another set of ten participants through our verbal approach. We, therefore, got thirty-nine participants in total for this experiment.

Twenty-five of the participating students were males, and the rest of them (14) were females. Fifteen of them were UG students and twenty-four were PG students which included both Masters and Ph.D. students. The participants were from science, engineering, mathematics, and humanities background. The mean age of the participants was 25.15 years with SD (sample) = 4.48. All of the participants were well acquainted with the use of the smartphone.

5.2.4 Participant Selection Strategy

We fixed the participant selection criteria as follows: only those volunteers would be selected who were well accustomed to using smartphones, who were not in severe medical condition (e.g., serious headache, shortness of breath), who did not consume any intoxicating substance in the last six hours before coming to the lab, and who had slept well in the previous night for at least six hours [30]. It may be noted that we followed the same selection strategy for all the experiments mentioned in this thesis. Once selected, each of the participants was asked to sign a consent form, which was also signed by an experimenter.

5.2.5 Procedure of Data Collection

Each volunteer was provided with two printed sheets. The first sheet was a 'consent form' (can be found in **Appendix** A). In this form, the privacy policy as well as the terms and conditions of the controlled experiment were written. The participants (and the experimenter) had to sign after reading and agreeing to it. The experimenter explained the tasks to be performed by the participants and gave sufficient time and scope to clear their doubts and queries, if any, before signing the consent form. The second sheet contained all the tasks and subtasks as well as blank spaces for reporting the results after completing the tasks (applicable for task 1 and 5 of Table 5.2). We did not use the reported results. Only the sensory and resource utilization data of the device were used. However, we designed the experiment in this way so that the participants perform the tasks seriously.

Upon agreeing to participate in the experiment and understanding the task by the participants, we asked them to perform the seven specified activities using the smartphone provided by us. We instructed them to perform the activities one by one, using a mobile browser only. The order of executing the tasks was different for each of the participants. We randomized the order to minimize the effect of confounds. The participants were asked to perform each of the tasks for approximately 4-5 minutes.

During the experiment, one observer was engaged in noting down the starting and the ending time (with all the intervals, if they were done in multiple attempts) of each of the activities performed by each participant. The clock of the observer and of the smartphone were synchronized before the commencement of the experiment.

5.3 CE2: Data Collection for Building 'Touch-Affect' Model

In this experiment, we collected touch data to build and validate a computational model that detects specific affective states of the user from their touch pattern on a small handheld mobile device such as smartphones and tablets. The data collection method also involves the induction of the specific affective states through a novel gaming approach. The data collection strategy (with a brief description of the existing methods) along with the collection of data using a novel game is presented next.

5.3.1 Data Collection Strategy

Affective data can be collected either from a natural source [31], or in the laboratory after artificially inducing particular affects to a group of participants [38] [103] [104]. The natural source means the corresponding affects are elicited naturally, not in the laboratory. For instance, someone is very excited after getting the notification that her/his research paper has been selected as the best paper, and instantly has started to write a mail to her/his mentor to share the good news. If we assume that there is a relation of affect with the typing patterns of the user, the typing data at this moment can be considered as 'affective data from a natural source'. Unfortunately, it is very difficult to collect such data. Moreover, inaccuracy is involved in the data at the time of labeling them. This is because, in most of these cases, data are labeled either by observation-based method or by user-feedback based method. In the observation-based method, inaccuracy is involved as feeling and perceiving the affect may differ [37]. User-feedback based method for labeling data is also problematic many of the times, because users may forget their exact affective state at a particular instance of a time when asked later (which is the common practice) [37]. Other difficulties in the naturalistic data are their limited source, the presence of too much noise, difficult for machine analysis, and copyright issues [31].

Researchers, therefore, artificially induce the required affects to a set of participants in the laboratory through controlled experiments [38] [103] [104]. The participants are then given some specific tasks through which the data are collected. There are many ways to induce specific affects: showing affective audio-visual content, hypnosis, intentionally making an unnecessary delay, taking a stressful interview, fake attack, taking surprise quiz, threatening for traumatic electric shock, and asking to solve an unsolvable puzzle [38] [103] [104] [134]. None of these techniques can guarantee error free data collection. To elaborate, when affective states are induced by showing audiovisual content, users may realize that they are being observed. As a result, they start behaving unnaturally [104]. Sometimes, additional unintended affects (e.g., anxiety with fear, hesitation with hostility and nervousness) may be induced in the attack, threat, and interview methods of induction [104]. Consequently, many a time, fake and inaccurate data are generated. Inaccuracy is also involved at the time of labeling the collected data following these methods. This is because, in most cases, the data are labeled depending upon the user-feedback. Most of the time, after the induction, participants are given some specific tasks through which data are collected and they are asked to report their affective states after and/or before the execution of the tasks. This strategy is followed because it is not possible for a participant to execute the task and report the states at the same time. Affective state at the time of induction, at the time of task execution, and at the time of providing feedback may vary since affect does not stay for long [16] [103] . As a result, inconsistency and inaccuracy in data may occur.

The use of gaming a approach may be a way to overcome these issues. There are many benefits to use a gaming approach to elicit affects. First of all, specific affect can be induced quickly and easily [142]. Secondly, the users can be induced the specific affect without their knowledge. Thirdly, data can be collected systematically without any user feedback. Fourthly, multiple eliciting elements can be incorporated in the game which may help to induce the exact intended affect. Fifthly, multiple types of affective data can be collected through this gaming approach. The types are gameplay (behavioral), objective, and game context [142]. The behavioral type refers to players' actions and preferences in real-time, which includes force applied while touching interactive elements, choosing particular elements among many alternatives, and so on. The objective type indicates the outcomes of elicited affects. Changes in body language, facial expression, voice, and physiological signals like heart rate, EEG are the examples of objective type of data. The game context data is the real-time state of the game. For instance, the mode or level of a game. It is very important as it helps for labeling the other two types of data (i.e., the behavioral and objective data). For example, a gaming mode can specify whether the arousal elicited is due to frustration or positive excitement [142].

We have considered the behavioral and game context data. We have not considered the objective data for building the model because capturing those require additional and expensive equipment. Processing these inputs also involve computationally expensive modules. These may not be supported in a mobile environment for detecting the affective states of smartphone users in real-time. However, we have used the objective data (EEG) later for an additional validation study of the model.

The literature contains some works where behavioral data has been collected through games to build models for detecting the affective states of the users. For instance, Gao et al. [39] collected the behavioral data (players touch pattern on the game screen in different levels of a game) to build a model for detecting the affective states of the players. However, they relied on the feedback from the players for labeling the data, and the labeling was done after the gameplay - it may be hard for the players to recall the particular affective states if they are asked about it later. There is a chance of bias because of the 'primacy and recency memory effect'. In other words, users may forget the times and sequence of the events that occurred in the recent past, as well as the status and level of affective states for the events. Stein et al. [122] also induced and collected the behavioral affective data through a gaming approach for building games with dynamic difficulty levels as per the affective states of the players. However, the collected data were not general interaction data – all of these were explicitly game-related (e.g., the number of enemies killed, the distance between two players). Gilleade and Dix [42], Miller and Mandryk [90], and Vicencio-Moreira et al. [135] worked on identifying types of frustrations (i.e., whether it is good or bad) of the players from their touch pressure during game-play. All these work are focused on a particular application, i.e., on adaptive game design. The input data may not be suitable for identifying the affective states of the users when they interact with general applications. Moreover, the desire of winning the game is the only one stimulator in these games to induce affects. This may limit the capability of inducing specific affect to the users. We, therefore, developed a novel gaming approach, having multiple stimulants, for collecting affective data related to general and basic touch interactions (i.e., not restricted to only game-related features), which is also able to label the data without any feedback from the users.

5.3.2 Experimental Setup

In order to collect the affective data of touchscreen users through the gaming approach, we designed a small game application for Android OS. In this game, a player has to guess a bucket among four buckets where a ball is hidden. If the player guesses

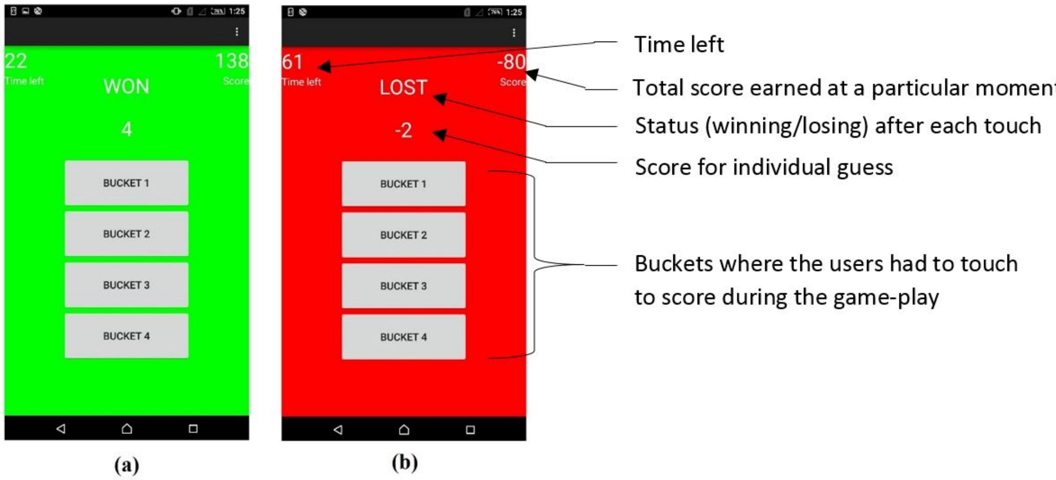


Figure 5.1: Game interface with two possible situations: (a) an example screenshot when a player guesses the right bucket, and hence, wins; (b) an example screenshot when a player guesses a wrong bucket, and hence, loses

correctly, s/he wins and scores four points whereas an incorrect guess results in a penalty of two points. The score is displayed on the top-right of the screen and the time left is displayed on the top-left of the screen. In addition to the score and time limit, two other emotion-inducing elements namely the background music and color have been used as stimulants. When a player wins, the background color becomes green and when a player loses, the background color becomes red [13] [121]. All along during the gameplay, music pieces were played in the background to aid the affect induction process.

The game had two modes: the winning mode and the losing mode. In the winning mode, the player always wins. It does not matter which bucket the player touches. Similarly, in the losing mode, the player always loses. In between the winning and losing modes, we kept a regular mode where the ball is placed randomly into any of the four buckets. The regular mode was fair. The player was informed neither about the strategy nor about the modes of the game.

Figure 5.1(a) is a screenshot when a player wins and Figure 5.1(b) is a screenshot when a player loses. There were two winning and losing modes in the entire duration of the game. The aim of keeping the same mode twice was to observe whether the touch pattern for a particular participant is the same for a specific game-mode. Although we required the touch pattern data for the winning and los-

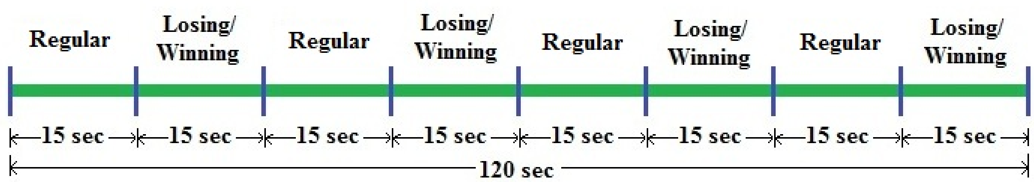


Figure 5.2: Timeline division-strategy of the game for a particular instance

ing modes only, a regular mode was kept at the starting of the game (to accustom the participants with the task), and in between the winning and losing modes (for bringing the participant back into neutral state as well as to avoid overlapping of the two modes). The regular modes were fixed in the timeline of the game. The winning and losing modes were assigned randomly to avoid biases. The timeline division scheme is depicted in Figure 5.2.

The game was of two-minute duration, which comprises of two winning modes, two losing modes, and four regular modes. The duration of each mode was fifteen seconds. We could not keep a longer duration for a particular mode because continuous winning/losing for a longer duration might lead to boredom or user frustration. Most importantly, a longer duration of a mode might help a participant guessing about the mode and the strategy of the game. The purpose of the particular gaming approach might not be served then.

5.3.3 Rationale for Game Design

When a player wins at the time of game playing, s/he feels good. The possible emotions are interest, joy, love, pride, pleasure, admiration, and so on. The emotion here is positive if any. On the other hand, when a player loses, s/he becomes angry, frustrated, guilty, sad, ashamed, disappointed, and so on. All these affective states are considered to be negative. Other than the ease of completing the game in time and earning a higher score in the winning mode, the affective color used in the background (green, which indices positive affect) acts as additional stimulants to induce positive affect. Similarly, in the losing mode, along with the fear of losing score, the red color in the background act as the stimulant for inducing negative

Table 5.3: List of affective music pieces used in the game

Sl. No.	Game mode	Music used
1.	Winning (with high arousal)	Tchaikovsky's "Mazurka from Swan Lake Ballet".
2.	Winning (with low arousal)	Gluck's "Orpheus and Eurydice".
3.	Losing (with high arousal)	Mussorgsky's "Night on Bald Mountain".
4.	Losing (with low arousal)	Marcello's "Adagio from Oboe Concerto in D minor".
5.	Regular or Neutral	Kraftwerk's "Pocket Calculator".

affect. The background colors for the two modes were selected based on earlier works ([13], and [121]). As the third stimulus for affect induction, affective music was played in the background during the game. The music pieces were also chosen from the literature [26] [134]. Thus, by using the three types of stimuli, the winning modes of the games were expected to be able to induce positive emotion, whereas the losing modes were expected to be able to induce negative emotion to the users.

At the same time, we also required to induce two levels of arousal. This is because we wanted to induce four different affective states: positive-high, positive-low, negative-high, and negative low, where 'positive' and 'negative' refer to the valence levels, and 'high' and 'low' indicate the levels of arousal. For this, we have taken the help of affective music. In one of the two winning modes, we have made use of music capable of inducing high arousal with positive valence, and in another winning mode, we have utilized music capable of inducing low arousal with positive valence. The same has been done for the two losing modes as well – one losing mode contained music capable of high arousal with negative valence, and another losing mode contained music capable of inducing low arousal with negative valence. Thus, the games were expected to induce the four specific affective states. In the regular mode, music pieces capable of inducing neutral affects were kept. The affective pieces of music were chosen based on [26] and [134]. In these studies, they have specified the characteristics of music in the context of affect induction capability. For example, Västfjäll [134] has mentioned that music having major-mode, fast-tempo, medium-pitch, firm-rhythm, dissonance-harmony, and high-loudness has the capability to make one excited. They have also mentioned a list of music pieces responsible for positive, neutral, and negative states specifying high and low arousal. Table 5.3 specifies the exact list of music pieces used in the games.

The smartphone, where the game was installed for data collection, was a 'One-Plus One' running on Cyanogen based OS 11 built on top of the Android Marshmallow. The screen size of the phone was 5.5 inches with a resolution of 1080x1920 (401 pixels per inch density i.e. ppi). The smartphone had a quad-core processor (2.5 GHz), and 3 GB of primary memory (RAM).

5.3.4 Participants

We requested the students for participating in the experiment with the same approach as mentioned in 5.2.3 and followed the same strategy to select the participants as mentioned in 5.2.4. In addition to that, we did not collect data from the participants whose finger size was abnormally large or small, as the value of touch-pressure is calculated indirectly based on the area of touch-surface covered by the finger during the touch interaction (in case, the pressure sensor is not available – which is true for most of the existing smartphones). Twenty-nine students voluntarily took part in this study. Out of these twenty-nine volunteers, fourteen were females, and rest fifteen were males. The mean age of the participants was 26.6 years with SD=2.13. All these participants were undergraduate and postgraduate students from engineering, humanities, mathematics, and science disciplines.

5.3.5 Procedure of Data Collection

Once the participants arrived in our laboratory, the rules for the game were described to each participant by an experimenter. At the beginning of each experiment, each of the participants was asked to play the game for familiarization and doubts clearance, if any. Afterward, the actual data collection took place. In this phase, they were asked to play the game for the specified duration. The participants were instructed to earn as much score as possible to get surprise gifts according to the score earned. Although the participation was voluntary, each of the participants was given some gifts as per their scores, and as a token of appreciation (this was done also for all the other experiments mentioned in this thesis).

5.4 CE3: Data Collection for Building 'Type-Affect' Model

In this experiment, we collected the affective data to build and validate another computational model that detects specific affective states of the users from their typing pattern on a mobile touchscreen keyboard. Like in CE2, in this experiment also, the data collection method involves both the induction of the specific affective states and collection of the affective data without user feedback. Affect induction as well as the data collection have been done through a novel game for the same reasons and following a similar strategy as described in the previous experiment (CE2). The details of the experimental setup and data collection procedure are described below.

5.4.1 Experimental Setup

Like in CE2, we developed a special-purpose Android typing game for inducing the specific affective states without the knowledge of the users, as well as for labeling and collecting data without any user feedback. The approach was expected to induce the states minimizing the imitation in emotion (as it was done without users' knowledge) and result in labeled data minimizing the inaccuracy (as the method was unobtrusive), as mentioned in the previous study.

The game had two modes: *fascinating mode* and *dull mode* (conceptually similar to the winning and losing modes of the game that was developed for CE2). While playing the game, a player had to play two fascinating modes and two dull modes. In the fascinating modes, s/he had to type a small paragraph (of 110-120 characters – possible to type for a normal person) whereas in the dull mode s/he had to type a long paragraph (of approximately 350-400 characters – might be impossible to type for a normal person) within the time limit of sixty seconds. Lengths of the paragraphs were decided based on the work of [66], [81], and [93]. Based on the correctness and speed, the score was displayed on the screen in the fascinating modes (players were informed that rewards will be given as per the score). For the correctness of typing, we used a string matching algorithm. The algorithm compares each of the characters of the typed text with each of the characters of the given paragraph. The typing was considered to be accurate when both the texts

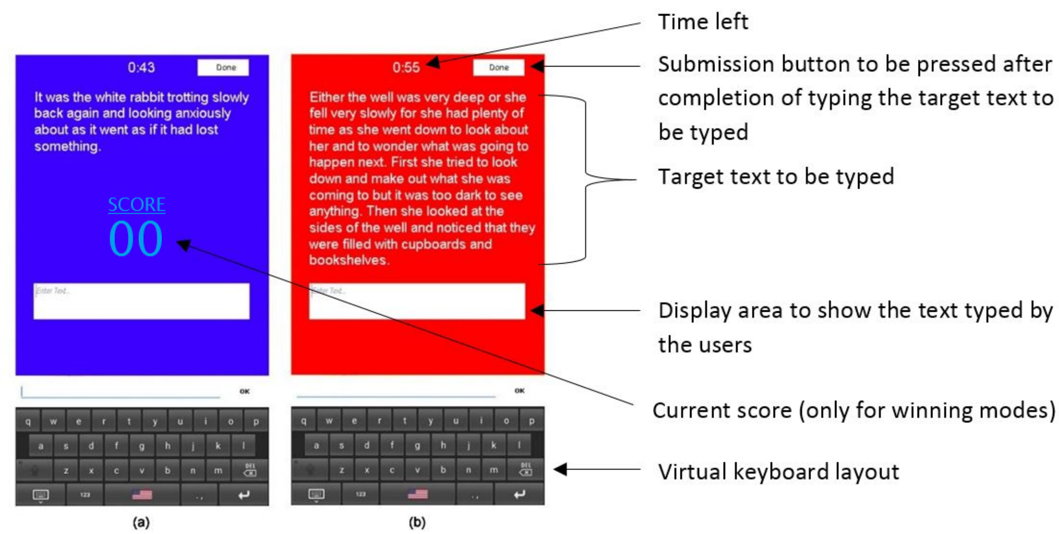


Figure 5.3: Game interface with two possible situations: (a) screenshot when a player plays fascinating mode (b) screenshot when a player plays dull mode of the game

match exactly. Erased characters and backspace keys were not considered for string matching. The score was displayed in the fascinating modes only. In the dull modes, the players were informed that they would be rewarded if and only if they could complete the typing task within the time limit. Two affective colors were kept as background colors in the two modes of the game to aid the induction method. The background color for the fascinating mode was kept blue whereas that of the dull mode was kept red [13] [121]. Figure 5.3 shows the snapshots of the two modes of the game: (a) when a user was playing fascinating mode, and (b) when a user was playing dull mode.

It was expected that the multiple stimulants of the fascinating mode (ease and possibility of typing a short paragraph, keenness of earning a maximum reward, and the blue color in the background) were able to induce positive emotions in the players. The emotions here were assumed to be of positive valence. On the other hand, stimulants used in dull mode (asking for completing an impossible typing job, practically no reward, and the red color in the background) were expected to induce negative emotions in the players. The emotions here were assumed to be of negative valence.

All along during the gameplay, the pieces of background music were played as additional stimulants for the induction method. Music capable of inducing positive valence with high arousal (Tchaikovsky's "Mazurka from Swan Lake Ballet") was kept in one fascinating mode, and music capable of inducing positive valence with low arousal (Gluck's "Orpheus and Eurydice") was kept in another fascinating mode. Similarly, music capable of the negative valance with high arousal (Mussorgsky's "Night on Bald Mountain") was kept in one dull mode, and music capable of inducing negative valence with low arousal (Marcello's "Adagio from Oboe Concerto in D minor") was kept in another dull mode. In between the fascinating and dull modes, we kept a piece of neutral music (Kraftwerk's "Pocket Calculator") to neutralize the induced emotion. As mentioned in CE2, we chose affective music based on Collier [26] and Västfjäll [134].

The modes were placed randomly to avoid bias. In between the two modes score and reward-details along with the rules for the next level of the game were shown to neutralize the affect induced in the last game-mode. Music capable of inducing neutral affect were also kept for this. Transitions of the music in between different modes were done smoothly (the volume of the music was reduced) to avoid the feeling of a jerk. The players were aware of neither the strategy for the induction mechanism nor the modes of the game.

We installed this game application on a smartphone having 5.5 inches screen, 'Android Marshmallow' OS, 2.5 GHz quad-core processor, and 3 GB RAM using which we collected the affective typing data.

5.4.2 Participants

We requested the students for participating in the experiment with the same approach as mentioned in 5.2.3 and followed the same strategy to select the participants as mentioned in 5.2.4. Thirty-three students voluntarily took part in this experiment. Among the thirty-three participants, seventeen were males and sixteen were females. The mean age of the participants was 24.19 with SD=2.91. The participants were undergraduate and postgraduate students of IIT Guwahati, and from engineering, humanities, mathematics, and science disciplines.

5.4.3 Procedure of Data Collection

On arrival, we asked a set of questions to the students to know their current health condition as well as the history of their health for selecting them as participants. As specified in 5.2.4, we did not collect data from the students who were in severe medical condition (e.g., serious headache, shortness of breath), who consumed any intoxicating substance in the last three hours or did not sleep well in last night for at least six hours. In addition to that, some of the volunteers could not participate due to their finger sizes (were relatively large compared to the size of the keys). This was because we wanted to avoid the 'fat finger' problem. Once selected, each participant was asked to sign a 'participant consent form' which was signed by one of the experimenters as well.

After signing the consent form, one experimenter described the rules for the game to each participant. The participants were first asked to play the game for familiarization and clear doubts if any. Afterward, the actual data collection took place. In this phase, they were asked to play the game for the full duration. The participants were instructed to earn as much score as possible so that they can get surprise gifts as per the score earned. Like other experiment mentioned in this thesis, each of the participants was given some gifts as per their scores, and as a token of appreciation.

5.5 CE4: Data Collection for Testing Induction Capability of Games

We have collected the affective data in CE2 and CE3 through the two games we have developed for inducing specific affective states and at the same time collecting the affective data without the knowledge of the users and their feedback. We claimed that the two games are able to induce specific affective states in the players while game-playing. This experiment was conducted for data collection to verify the capability of the games in terms of affect induction.

5.5.1 Experimental Setup

We used the EEG signal of the users for the verification of affect induction capability of the games. The EEG signals were captured by a mobile headset called EPOC+[1] having 14 EEG channels (AF3, AF4, F3, F4, F7, F8, FC5, FC6, O1, O2, P7, P8, T7, and T8). It can collect the EEG signals with a sample rate of 2048/sec, which can be downsampled to 128 SPS. The bandwidth (response frequency) of the device is 0.16 – 43.00 Hz. A proprietary software, called EmotivPRO[2], was used to analyze the EEG signals. The EmotivPRO can store and analyze the raw EEG signals, and derive six affective states ('engagement', 'excitement (arousal)', 'focus', 'interest (valence)', 'relaxation (meditation)', and 'stress') from the raw signals, in real-time. In this study, we required only two of them (valence and arousal). We installed the EmotivPRO software on a laptop having Windows 8.1 Pro 64-bit (6.3, Build 9600) operating system, Intel® Core™ i5-332M 2.6 GHz quad-core processor, and 4 GB RAM. The EPOC+ headset was connected to the laptop via a Bluetooth dongle. The two games, which were developed by us to induce specific affects and unobtrusively collect the affective data, were installed in a smartphone having 6.0 inches display, Android 4.2 Jelly Bean OS, quad-core processor (1.4 GHz), and 1 GB RAM.

5.5.2 Participants

We requested the students for participating in the experiment with the same approach as mentioned in 5.2.3 and followed the same strategy to select the participants as mentioned in 5.2.4. Ten volunteers participated in this experiment. The participants were UG and PG students of IIT Guwahati, and from science, engineering, and humanities disciplines. Five of the participants were male and rest five were female. The mean age of the participants was 25.3 years with SD=3.94.

5.5.3 Procedure

After assuring the non-harmfulness and non-invasiveness of the device, the participants were asked to wear the EPOC+ headset. One experimenter assisted the

[1] https://www.emotiv.com/epoc/
[2] https://www.emotiv.com/emotivpro/

participants to wear the device. The experimenter also verified and confirmed the appropriate positions of the EEG channels on participants' skull. After demonstrating the games by the experimenter and clearing doubts, the participants were asked to play the two games for the entire duration, one by one. During gameplay, the experimenter was engaged to observe the affective states of the participants through the EmotivPRO software. The behavioral affective data were also being collected by the game apps while the participants were engaged in playing the games.

5.6 CE5: Data Collection for Establishing the Ground Truth of Emotion

The EEG signals were used for both validating the affect detection models and verifying the models' applicability in the current context. We assumed the affective states derived from the EEG by EmotivPRO as the ground truth. However, it was required to establish the same. This experiment was conducted to collect the data for verifying the capability of EmotivPRO in the context of detecting the specific affective states. Although an earlier study [30] validated the reliability of an earlier model of the same device (EPOC) in terms of EEG signal capturing capability and consistency of the device, we conducted this study to collect pieces of evidence for establishing the ground truth of the particular affective states we studied.

5.6.1 Experimental Setup

We used EPOC+ headset to capture the EEG signals, and EmotivPRO to analyze the EEG for deriving the affective states as specified in 5.5.1. In addition to the setting consisting of the EEG device and software, we also arranged an audio amplifier and a projector with a large screen to play audiovisuals that are capable of inducing different levels of arousal and valence. We selected these audio-visual clips based on [26] and [134]. Table 5.4 presents the lists of audio-visual contents (downloaded from YouTube) used for this experiment. It may be noted that this list of music was different from the list of music (Table 5.3) used in the two games for eliciting affect. This is because it was not feasible to play a video in the background of the games – the user might be distracted from the tasks in such cases. Most importantly, the

Table 5.4: List of audio-visual used in CE5

Name of the audio-visual	Link (URL)	Used to verify
High Arousal Music Playlist	https://www.youtube.com/watch?v=WHUo4pZLAAA	Arousal levels
The Best Of Eagle Attacks 2018 - Most Amazing Moments Of Wild Animal Fights! Wild Discovery Animals	https://www.youtube.com/watch?v=RB4RCOe-ZEw	Arousal levels
World most Shocking Video that made the whole world cry!	https://www.youtube.com/watch?v=QJxwE7mdGns	Arousal levels
14 Strange Ways of Life the Ancient Egyptians Practiced	https://www.youtube.com/watch?v=GiMbVa6XzTw	Valence levels
What Will Happen to Humans Before 2050?	https://www.youtube.com/watch?v=Cip3LmqQ7Y0	Valence levels
True Facts That Will Shock You	https://www.youtube.com/watch?v=HChCEGR_0lg	Valence levels

purpose of this study was different. In the cases of the games, the music along with the other two elicitors (affective colors and game mode) were used for inducing specific affective states. On the other hand, this experiment was conducted to verify whether the change of the level of the arousal and valence (if any) can be detected by the EmotivPRO. In other words, this experiment was to validate the capability of the EmotivPRO for detecting the specific states correctly so that we can consider them as ground truth. Instead of the audiovisual contents, any other elicitor, which is capable of inducing different levels of arousal and valence, could also be used for this study, keeping the rest of the procedure intact.

5.6.2 Participants

Twenty volunteers participated in this study. Eleven of them were male and rest nine were female. The mean age of the participants was 24.7 years with SD=3.37. The participation selection criteria were the same as specified in 5.2.4.

5.6.3 Procedure

The participants were shown the videos partially (for a maximum of two minutes) on a large projector screen with controlled sound (Figure 5.4). The videos were shown for a short duration because it was observed that the participants were unable to recall their affective states if the same were asked a few minutes later. This corroborates the fact that the obtrusive methods for collecting affective data may contain incorrect information. We recorded the arousal and valence level of the participants by the EmotivPRO while they were watching the videos. At the same time, two experts were also observing and reporting the levels of arousal and valence

Figure 5.4: Study to establish the ground truth of affective states

of the participants by observing their facial expressions and body languages. After each video, we asked the participants to report their levels of arousal and valence either as high or low, at different points of time (approximately after every ten seconds). The videos were replayed for them which helped them to recall their states.

5.7 CE6: EEG Data Collection for Additional Validation Study

During the training and testing of the two affect detection models (Touch-Affect and Type-Affect) with the data collected in CE2 and CE3, we observed high classification accuracy for both the models (reported in **Chapter** 6). In addition to testing the models with the collected data, the models have been further validated using EEG signals of the participants. The details of the data collection process for this additional validation study are reported next.

5.7.1 Experimental Setup

The same mobile EEG headset (EPOC+) and EEG analytic software (EmotivPRO) with the same experimental setup (specific laptop and other equipment), as explained in 5.5.1, have been utilized for this study.

5.7.2 Participants

The study made use of data obtained from a separate group of twelve volunteers (six female, mean age = 24.67 years with SD=2.53). We requested the students for participating in the experiment with the same approach as mentioned in 5.2.3 and followed the same strategy to select the participants as mentioned in 5.2.4.

5.7.3 Procedure

Like 5.5.3, with the consent of the participants, after they were assured about the non-invasiveness of the device, the participants wore the EPOC+ headset with the assistance of an experimenter. The two games which have been developed by us were installed in the smartphones as specified earlier. The clocks (of the laptop where the EmotivPRO was installed and the smartphones where the games were installed) were synchronized before the start of the experiment. After demonstrating the games by the experimenter and clearing doubts, the participants were asked to play the games for the entire duration, one by one. During gameplay, the game apps recorded the affective data, which were required for the two affect detection models, namely the Touch-Affect and Type-Affect. At the same time, the EmotivPRO recorded the arousal and valence level of the participants.

5.8 CE7: Data Collection for Testing Compatibility of Affect Detection Model

We combined the two computational models (Touch-Affect and Type-Affect) to build a process model. The process model called the Smart-Affect can detect the affective states of the smartphone users from their basic touch interaction data (i.e., touch and type data). The Smart-Affect model identifies the nature of input data during user-smartphone interaction, and send the data to a specific computational model to detect the affective state of the user. Although the associated computational models were validated separately (at the time of testing the models with the affective data, and later with the help of EEG signals of the participants), we validated the process model in the target application environment. In other words, we validated the model in a blended learning platform. This experiment was conducted to collect data for verifying the applicability of the combined model in the current context.

5.8.1 Experimental Setup

We considered a blended learning platform called "Avabodhaka" [24] [130] to investigate whether our proposed model works for this application. In Avabodhaka, all the classroom activities including lecture delivery are performed through mobile devices. The system follows a client-server architecture. A local server provides the necessary infrastructure for the classroom activities and the contents for the same. All the smartphones used by the students and the teacher act as clients. The teacher and the students can take part in classroom activities using Avabodhaka after logging in to the system with their credentials. A detailed description of the system is previously presented in **Chapter** 3. Although in the original system, Avabodhaka, clients might use any mobile devices (such as smartphones, tablets, and laptops), we restricted it to only smartphones for this experiment. With Avabodhaka, a background app was also running on the smartphone used by the students. The background app was able to collect the touch and typing data required for our model.

The same EEG setup, i.e., the combination of EPOC+ and EmotivPRO, as specified in 5.5.1, was used to capture and analyze the EEG signals. A background app was installed in a smartphone that was used by the participants to perform the tasks. The clocks of the smartphone where Avabodhaka with the background app was running, and the laptop where the EmotivPRO was installed were synchronized before the commencement of the experiment.

5.8.2 Participants

We performed this study with a separate group of fifteen students (seven female, mean age of 24.9 years with SD=1.85). The request for participating in the experiment and the participant selection strategy were the same as specified in 5.2.3 and 5.2.4, respectively.

5.8.3 Procedure

Sixteen demo lectures were conducted for the study. One of the experimenters, who is a researcher in HCI and a professor of Computer Science and Engineering having almost fifteen years of teaching experience, delivered the lectures using Avabodhaka. The duration of each of the lectures was in between thirty to forty minutes. There were between ten to fifteen students for each lecture. On the first day of the lecture, all the features of the Avabodhaka system were demonstrated to the students by one experimenter. The students were also asked to get acquainted with the system and clarify doubts, if any, regarding the functionality of Avabodhaka. They were also informed that from the second day onward one participant would have to wear the EPOC+ device during the lecture. They were convinced that the device was completely non-invasive and safe to wear. A demonstration on how to wear the device was given to all the volunteers. On the first day, we did not collect any data. From the second day onward, in each lecture, one of the participants wore the EPOC+, which was connected to the EmotivPRO via a Bluetooth dongle, which was running on a laptop. In the fifteen lectures (lecture-2 to lecture-16), we collected both the EEG data (raw EEG signals and derived affective states) as well as the smartphone interaction data from the fifteen participants who wore the EPOC+.

We also extended this study to collect data for verifying the generalizability of the Smart-Affect model. As the inputs of our model are the basic user-smartphone interaction data (touch and type pattern) which are independent of applications, we were interested to know whether the model works in general applications where these basic interaction data are available. The generalizability of the model was verified through a pilot study to justify the fact that although the model has been built with the data collected through game applications, it is independent of application areas and hence works for other applications as well – the blended learning platform is one of them. Through the pilot study, we tested our model for social networking and instant messaging applications (Facebook and WhatsApp). The browser versions of these applications were used for the ease of required data collection. For validation purpose, again the EEG of the participants were used. The same EEG setup, i.e., the combination of EPOC+ and EmotivPRO, as specified in 5.5.1, was used to capture and analyze the EEG signals. A background app was installed in a smartphone that was used by the participants to perform the tasks. The app was able to collect the input data for the Smart-Affect model. The clocks of the smartphone where the background app was installed and the laptop where the EmotivPRO was installed were synchronized before the commencement of the experiment. We conducted this study with nine participants (four female, mean age = 23.11 years with SD=2.31), who were selected with the same criteria as specified in 5.2.4. As an additional criterion, we verified whether the participants were regular users of social networking and instant messaging applications. We asked these participants to wear the EPOC+ and use any of the instant messaging applications of their choice (e.g., WhatsApp, Facebook Messenger) and social networking website (Facebook, Twitter) for fifteen minutes each. For instant messaging, they were instructed to chat with their relatives/friends. For social networking sites, they were instructed to view, like, and post some comments on the 'shared media' and 'post' by their friends and relatives. All of the participants chose the WhatsApp as instant messaging, and Facebook as social networking. We recorded the arousal and valence by the EmotivPRO while the participants were performing the given tasks. At the same time, the background app recorded the touch and typing features that were required for the Smart-Affect model.

5.9 Chapter Summary

All the empirical studies conducted for accomplishing the thesis work have been elaborately presented in this chapter. A summary of all the studies has been presented in Table 5.5. The studies were required to collect data for building the different components of the proposed framework for a sensitive classroom, especially for building

Table 5.5: Summary of the empirical studies

Name	**Purpose**	**No. of Participants**	**Group of Participants**	**Task Set for the Participants**	**Data Collected**
CE1	Data collection for building the In-Activity model.	39	Unique	Performing a set of study and non-study related smartphone activities	Accelerometer reading, gyroscope reading, memory utilization, battery temperature.
CE2	Data collection for building the Touch-Affect model.	29	Unique	Playing a game (the buckets & balls game).	The number of touch events ('action down', 'action up', and 'action move'), the pressure generated for each event.
CE3	Data collection for building Type-Affect model.	33	Unique	Playing a game (the typing game).	Typing speed, backspace key press frequency, maximum text length, touch count, device shake frequency.
CE4	Data collection for testing induction capability of games used in CE2, and CE3.	10	Unique	Playing the two games[0] (the buckets & balls, and the typing game).	Same as the data collected in CE2 and CE3. + EEG data of the participants.
CE5	Data collection for establishing the ground truth of emotion.	20	Unique	Watching audiovisuals[0] (capable of inducing specific affective states).	EEG (raw and derived) data of the participants. + Experts' perception of the specific affective states of the participants. + User feedback on affect induction.
CE6	Data collection for additional validation of the Touch-Affect and Type-Affect model.	12	Unique	Playing the two games[0] (the buckets & balls, and the typing game).	Same as the data collected in CE2 and CE3. + EEG data of the participants.
CE7	Data collection for testing the compatibility of the Smart-Affect model in the target application.	15	Unique	Attending lectures on a blended learning platform.[0]	Same as the data collected in CE2 and CE3. + EEG data of the participants.

several computational models. The studies have been appropriately referred to in the other chapters, where they are associated with. We consolidated all the empirical studies in this chapter with an expectation for improving the readability of the thesis. For the same reason, however, the analysis of these data to build and validate the computational models are presented in the next chapter (**Chapter 6**), where the development processes of all the models have been presented in detail.

[0]indicates that the participants wore a mobile EEG headset during the experiment.

"Let's not forget that the little emotions are the great captains of our lives and we obey them without realizing it."

Vincent Van Gogh (1853 – 1890)

Dutch painter

6

Computational Models

6.1 Introduction

In order to develop Vedinkaksha, we have built four novel computational models, namely Touch-Affect, Type-Affect, Smart-Affect, and In-Activity. Among these four models, the Touch-Affect, Type-Affect, and Smart-Affect models have been built to detect the affective states of the students, whereas the In-Activity model has been built for the detection of involvement of the students in study-related activities. Touch-Affect and Type-Affect are the two computational models that detect the affective states of smartphone users from their touch and typing behavior, respectively. The Smart-Affect is a process model to detect the affective states from basic user-smartphone interaction data, which incorporates the two computation models, namely Touch-Affect and Type-Affect. Based on the availability and type of input data, the process model (Smart-Affect) chooses either of the Touch-Affect and Type-Affect models to compute the affective states of the users for a particular interval of time. If no input data is available for a certain interval, it shows the last identified states. The In-Activity is a separate computation model that detects the involvement of the students in the teaching-learning process from their device handling pattern along with the status and utilization of the resources of the device.

All these models have been built using machine learning techniques. We, therefore, have specified the feature sets for the models, empirically collected data according to the features, analyzed the data for choosing appropriate machine learning algorithms as well as for optimizing the feature sets, and trained and tested the models with the collected data. In addition to testing the models with the collected data, in few cases (in case of affect detection models), we have validated the models with EEG data of the users. We have used the EEG data to verify the workability of the models in the target application as well. The detailed descriptions of the procedure for building and validating the models are presented in this chapter.

6.2 Touch-Affect: Model to Detect Affective States from Touch Interaction Pattern

This is a minimalist model to detect the affective state of the users of a small handheld touchscreen device, such as a smartphone and a tablet. The model detects the affective states of the user from their touch behavior on the touchscreen. The model is expected to be useful as touch is one of the basic user-smartphone (or user-tablet) interactions, and we assumed only such devices for classroom interactions in the proposed framework for a sensitive blended learning platform. The novel minimalist model considers the number of touch events and average pressure created for these events for a specific time interval to detect the affective state of a user in one of the specific four states based on the two levels of arousal and valence.

6.2.1 Affective State Specification

In order to attain the suitability of the model in the target application, we have considered a specific set of affective states based on the Circumplex model of emotion [106]. This is because this particular type of affective states has been found to be most suitable in the context of the classroom environment [141]. According to the Circumplex model of emotional theory, the continuous affective states are discretized based on two levels of arousal and valence. Arousal is the physiological state of being reactive to stimuli. It results in an observable change/action in the physical state of the body which causes us to become alert and ready to move and respond. In general, arousal is the level or amount of physical response. Valance, on the other hand, is the intrinsic attractiveness or aversiveness of an emotion. In other words,

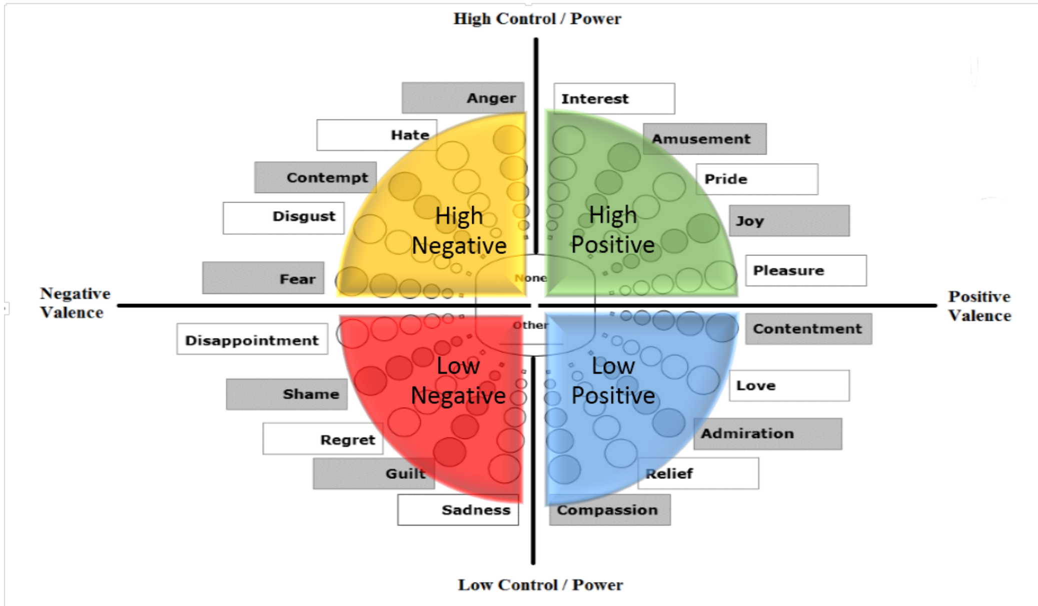


Figure 6.1: The Geneva Emotional Wheel (GEW) and proposed partitions of the wheel into four discrete affective states based on the Circumplex model of affect

it indicates whether or not people are feeling good and interested in staying in a particular state of mind.

We based our idea of four Circumplex states on the Geneva Emotional Wheel (GEW[1]) [111]. The GEW is a theoretically derived and empirically tested instrument, to measure emotional reactions to objects, events, and situations. The GEW contains twenty different emotions with their intensities diverging out as shown in Figure 6.1. The emotions are arranged in a wheel shape with the axes being defined by two major dimensions of emotional experience: valence (x-direction) and arousal (y-direction).

The four quadrants of the GEW are represented as four distinct emotional states based on the Circumplex model of emotion, which are listed below.

1. 'High-Positive' (the top-right quadrant of the GEW),

2. 'High-Negative' (the top-left quadrant of the GEW),

3. 'Low-Negative' (the bottom-left quadrant of the GEW), and

4. 'Low-Positive' (the bottom-right quadrant of the GEW).

[1]http://www.affective-sciences.org/en/gew/

Table 6.1: Features considered for developing the Touch-Affect model

Sl. No.	Name of the feature	Explanation
1.	The number of touch events created per unit of time	An event is any one of the three actions: 'action-down' (putting the finger on the touchscreen), 'action-move' (moving the finger while touching the screen), and 'action-up' (removing the finger from the screen). The touch events are captured by the android API 'onTouchEvent()'.
2.	Average pressure applied on the events	The pressure is calculated indirectly by the android API 'getPressure() '. This API considers the area of contact on the screen during touch to indirectly calculate the pressure applied on the screen. (We, therefore, did not consider the data of those participants whose finger size was abnormally large or small).
3.	Average duration of the touch events	The time gap between two consecutive events.

Here, 'High' and 'Low' indicate the levels of arousal, whereas the 'Positive' and 'Negative' refer to the levels of valence. It can be noted that we have considered this specific set of affective states for all the affect detection models which have been reported in this thesis.

6.2.2 Feature Set Specification

We started by considering the following three features (presented in Table 6.1) as literature hints that those have some relations with the affective states which have been considered in our study [44] [64] [144].

Equations (6.1), (6.2), and (6.3) were used to calculate these feature values.

$$No\_of\_touch\_event(per\_unit\_time) = \sum_{i=1}^{n} E_i \tag{6.1}$$

$$Avg\_pressure(per\_unit\_time) = \frac{1}{m}\sum_{i=1}^{m} P_i \tag{6.2}$$

$$Avg\_duration(per\_unit\_time) = \frac{1}{m}\sum_{i=1}^{m} D_i \tag{6.3}$$

Here, n is the number of touch gesture created per unit time; m is the number of events created per unit time, E_i specifies the number of events generated during the i^{th} gesture, P_i is the pressure generated at the i^{th} event, and D_i is the duration of the i^{th} event.

However, the duration of the touch events was discarded as the literature contains conflicting reports on the relationship between the event duration and the affective states. For instance, the study results of Matsuda et al. [85] imply that the

Table 6.2: Events and pressure generated in the sample data of a particular user in a particular time interval (0 – 3 sec)

Timestamp in millisecond	Event occurred	Pressure generated for the event occurred	Touch interaction activity
1261	Action Down	0.32941177	Tap 1
1309	Action Move	0.32941177	
1324	Action Up	0.32941177	
2783	Action Down	0.3137255	Tap 2
2815	Action Move	0.29803923	
2835	Action Move	0.18823531	
2926	Action Up	0.18823531	

'joy' state is related to a shorter touch duration whereas the 'sad' state is related to a longer touch duration. On the other hand, the study results of Hertenstein et al. [58] indicate the opposite.

We, therefore, finally considered only the two features namely, the number of touch events created per unit of time, and the average pressure applied on these events (Sl. No. 1 and 2 in Table 6.1). These two features are used for detecting affective states as one of the specified four states, using a machine learning algorithm.

6.2.3 Data Analysis

The data collection method for this model has been described in the previous chapter (**Chapter 5**, CE2). During the experiment, while the participants were playing the game, the game app recorded the number of touch events and the pressure generated for each event. Table 6.2 shows a sample of the collected data in CE2.

Using the EEG data of the students, which were collected through a controlled experiment (CE4), we have also validated the capability of the game in terms of inducing specific affective states. In CE4, we recorded the specific affective states (arousal and valence) of the participants through the EmotivPRO. It was observed that the valence (interest) of the players was higher when they were playing the winning and fascinating modes of the games and lower when they were playing the losing and dull modes of the games. The observation indicated that the games were able to induce the two different levels of valence. It was also observed that the arousal (excitement) level was high when the participants were playing a particular winning mode and the same was low when they were playing another winning mode. Similar observations were made for arousals in the two losing modes (in one losing mode, the arousal was high, whereas, in another losing mode, it was low). This

Table 6.3: Average arousal and valence of ten participants while playing the various modes of the games

Participant	**Avg arousal observed**		**Avg arousal observed**		**Avg valence observed**		**Avg arousal observed**	
	in *winning mode* with music capable of inducing *high arousal*	in *winning mode* with music capable of inducing *low arousal*	in *losing mode* with music capable of inducing *high arousal*	in *losing mode* with music capable of inducing *low arousal*	in *winning mode* with music capable of inducing *high arousal*	in *winning mode* with music capable of inducing *low arousal*	in *losing mode* with music capable of inducing *high arousal*	in *losing mode* with music capable of inducing *low arousal*
1.	65	37	67	34	59	58	28	30
2.	70	36	65	28	65	72	30	32
3.	67	38	71	30	62	59	30	31
4.	71	35	73	34	71	62	31	35
5.	64	34	64	39	70	75	38	32
6.	80	39	60	38	79	72	35	34
7.	70	35	69	32	60	58	30	33
8.	66	32	69	29	61	63	31	35
9.	65	29	71	40	66	61	32	29
10.	70	30	60	35	63	70	30	31

was because such eliciting elements were kept in the particular modes (music having the capability of inducing low arousal in one winning and losing mode, and that of inducing high arousal in another winning and losing mode).

Table 6.3 presents the observed average arousal and valence levels (shown by EmotivPRO) of all the participants when they played the different modes of the games. Figure 6.2 illustrates the change of affective states for a particular participant (1) among them while playing the different modes of the games. From this observation, we may conclude that the games are capable of inducing the desired states.

In order to verify the induction capability of the game, we have assumed the affective states identified by EmotivPRO, which were derived from EEG signals captured by EPOC+, as the ground truth. We established the ground truth of the affective states using the data collected in CE5. In CE5, we collected the reports about the affective states of the users from three different sources: from EomotivPRO (derived from EEG signals), from the observers, and from the participants. We compared the affective states found from the three sources and found that the affective states detected by EmotivPRO, reported by the observers, and reported by the participants were similar in 92.08% of observation points. This indicates that the arousal and valence identified by the EmotivPRO can be considered as the ground truth.

Figure 6.2: Change of affective states for a particular participant, while playing different modes of the two games

The above experiments and the analyses of their results indicate that the game is able to induce the desired affective states to the users. Also, we have found that there are some relationships of touch behavior with the affective states, and consequently, we have finalized the feature set for the affect detection model. Nevertheless, we required a suitable machine learning algorithm which could be trained with the collected data to build the model. For this, we plotted the graph for the two features namely, the number of touch events and touch pressure generated after every three-second interval in the winning and the losing modes (Figure 6.3). We chose the three-second time interval for data analysis considering the way humans react to stimuli. Once an emotion is triggered, it takes few moments to respond and the emotion stays for few moments before changing the state (for instance, from high arousal to low arousal and vice versa) [51]. We assume that a three-second interval for observing the change of emotion is sufficient.

While observing the number of events and touch pressure generated on the

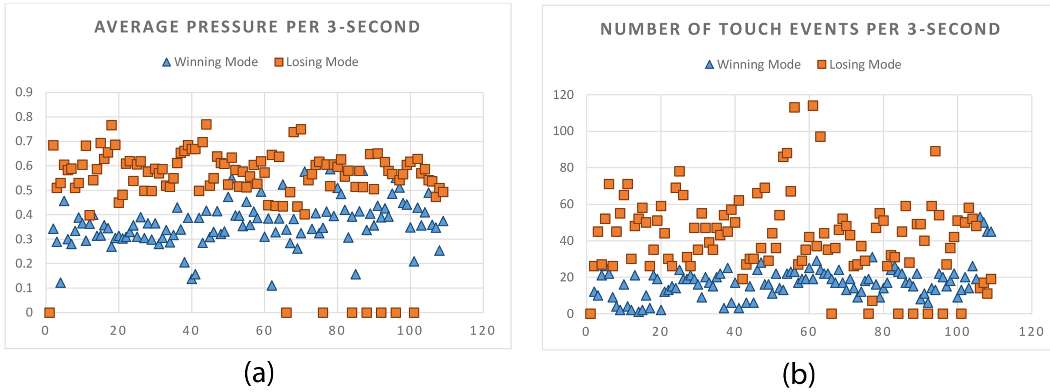


Figure 6.3: Data observation; (a) average touch pressure generated on screen and (b) number of events generated, in winning and losing mode

screen for the winning and losing modes, we found that the data of the two modes were mostly linearly separable (Figure 6.3). Therefore, the discriminative classifier, Support Vector Machine (SVM) with linear kernel might be chosen for classification. We chose the multilevel SVM because the classifications were based on two levels of binary classes – first for valence and then for arousal. Also as per our model, the classes are based on the threshold value for the number of touch events and touch pressure. Therefore, the K-Nearest Neighbour (KNN) could also be explored. As our dataset contained equal distribution of data on the four specified affective states, and the number of features was only two, we thought that the Naïve Bayes could also be a good choice. Additionally, we also explored the Decision Tree (DT). The DT was explored as we used only the two features as input for the classification. Consideration of only the two features would make the tree size smaller and hence would give high classification accuracy.

6.2.4 Training and Testing the Model

We used the cross-validation technique (CV) for training and testing the model. We chose Leave-One-Subject-Out Cross-Validation (LOSOCV) technique for this [52]. We chose this particular CV because Hammerla and Plötz [52] have argued that this CV is best applicable to the type of experiment we conducted. In LOSOCV, instead of random division, the dataset is divided participant wise for the cross-validation. The average classification accuracies, while adapting different machine learning algorithms for the model, are depicted in Figure 6.4.

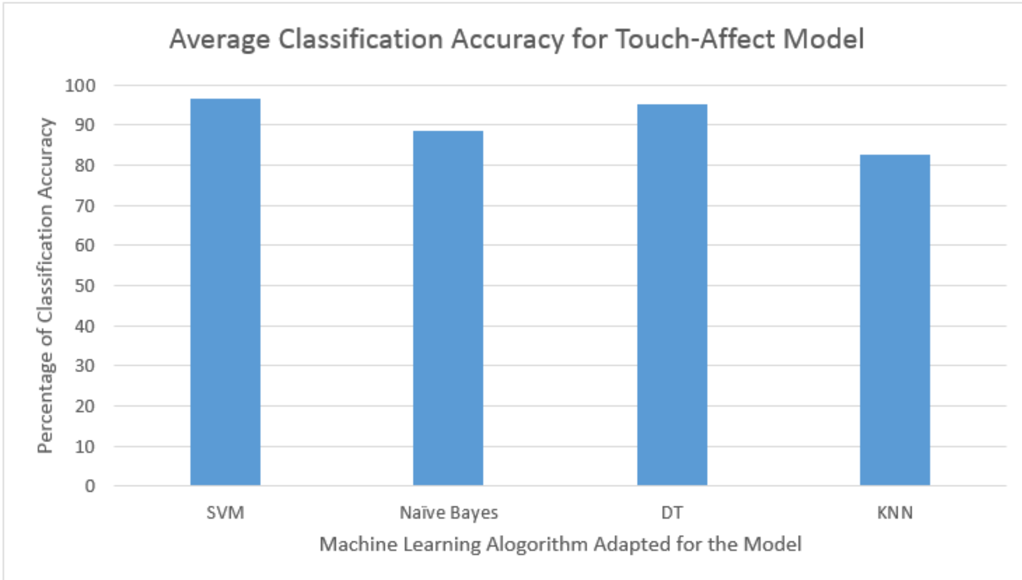


Figure 6.4: Average classification accuracies for the different machine learning algorithm adapted for the Touch-Affect model

Although we got high classification accuracy at the time of testing the model with the affective data collected through the game we have developed, we further extended our experiment to additionally validate the model with the EEG data of the participants. The EEG data were collected through CE5 (which has been reported in the last chapter). During this experiment, EmotivPRO recorded the arousal and valence level of the participants, which were derived from the EEG signals. At the same time, the game app recorded the affective data, as per the definition of the input features of the model. We computed the affective states by the Touch-Affect model and compared those with the affective states detected by the EmotivPRO. We found that the detection of arousal and valence levels by the Touch-Affect models, and by the EmotivPRO were the same in 88.05% of cases. This high similarity additionally validates the Touch-Affect model.

6.2.5 Discussion

Firstly, our proposed method for building an affect detection model addresses the issues in existing methods of affect induction. Using our method, we can induce specific affective states to the users of mobile touchscreen devices without their knowledge. This resolves the problem of imitation in emotion which may present in the methods of inducing emotion through film and music. The real-time unobtrusive data collection method fixes the problem of inaccuracy in data which may exist in the methods of collecting data after inducing emotion through surprise exam and quiz, unsolvable puzzle, the threat of electric shock, sudden attack, and hypnosis. Secondly, our proposed model detects the affective states of the user using the minimum number of indirect input features. Only the two touch features, namely, the number of touch events and the touch pressure, are sufficient for identifying the affective states. The features are inexpensive in terms of computation. This in turn reduces the computational complexity. Moreover, data collection and processing is done in low-cost mobile devices. So, the applicability and adaptability of the model increase. Thirdly, neither the users bother about where the computation is taking place nor the users are required to go to any particular computational setting for the computation. Computing can take place everywhere. Hence, the model is ubiquitous as well. Most importantly, the model has been found to be suitable in our target application, which has been elaborated in 6.4.2.

The gaming approach with multiple stimulants for inducing affective states for collecting affective data in real-time without the knowledge and feedback of its user is novel. We claim the novelty also in building the minimalist model as we could not find any work where the affective states of a touchscreen user have been identified using only two indirect input features with such high accuracy.

With the major advantages of the model, there are some limitations and scope for future research. Our model, in its current state, is able to detect four different affective states: high arousal with positive valence, high arousal with negative valence, low arousal with positive valence, and low arousal with negative valence. However, sometimes we may require to discriminate the emotions more specifically (e.g., distinguishing 'fear' from 'anger', where both these emotions belong to the 'high arousal and negative valence' state) depending upon the application areas (e.g., for movie making industry). Nonetheless, the specific emotional states detected by our model have been found to be best suitable for our target application, i.e., the sensitive classroom system [141]. Therefore, the model seems to be perfect for the current context. Nevertheless, one can further explore the idea to induce the affective states in a more specific way and build a model to detect them in a comprehensive way, if the target applications demand. It may be required to modify the game for that purpose.

In this work, we have chosen three seconds as an interval to observe affective states of the users. Our model worked with high accuracy considering this time interval. However, the choice of an appropriate time interval is not an easy task. There may be one way to figure it out in this context. We could choose different time intervals (for instance, from one to twenty seconds) and examine the differences in the feature attributes in high and low arousal. The time interval with a significant difference could be chosen as the appropriate one. However, there are a few problems. First, we need to increase the duration of the game modes, particularly the winning and losing modes, for this method to work. It is expected that the continuous winning or losing over a longer duration may give some hints to the players (like 'there is something wrong in the game'). This consequently may affect the purpose of the gaming approach for emotion induction and data collection. Another problem of choosing the appropriate time interval is that the same may differ depending upon the context, the feature set to be observed, and the task defined for emotion induction. This opens a new research issue for future researchers.

6.3 Type-Affect: Model to Detect Affective States from Typing Pattern on Virtual Keyboard

Other than the general touch, there is another basic interaction in the case of user-smartphone (or user-tablet) interactions, which is the typing on a virtual keyboard on the touchscreen. In our target application, the users (students) are expected to have this interaction many a time (for making a query, adding a class note, and taking an examination). Although the typing in the virtual keyboard is also performed through touching the screen, this touch pattern is different from the general touch pattern. In the case of general touch interaction, users generally interact through the tap, scroll, swipe, or multi-touch gesture. On the other hand, in the case of typing interaction, users only tap. Moreover, the frequency of tap in case of typing is much higher compared to that of general touch interaction. A single computation model may not be sufficient for detecting the affective states of

Table 6.4: Features considered for developing the Type-Affect model

Sl. No.	Name of the feature	Explanation
1.	Typing speed	The number of characters typed by the user per second.
2.	Backspace keypress frequency	The number of backspace key pressed by the user per second.
3.	Special symbol frequency	The number of special symbols (e.g., astrological symbol, emoji, and smiley) used by the user per second.
4.	Maximum text length	The total number of keys pressed without pressing the backspace key in a second.
5.	Erased text length	The number of characters erased in a second.
6.	Touch count	The number of keys pressed including the backspace key, in a second.
7.	Device shake frequency	The number of times the device is shaken (displaced above a threshold value) in a second.

the user form the two types of touch interaction behaviors. Therefore, we propose another computational model, called Type-Affect that detects the affective state of a user in one of the four specific states from her/his typing pattern on the virtual keyboard.

6.3.1 Feature Set Specification

In order to build the Type-Affect model, we initially considered the following seven features (presented in Table 6.4) as they have some relations with affect and emotion [72] [76].

Equations (6.4) – (6.10) were used to calculate these feature values.

In the equations, C is the number of effectively typed characters (without considering the erased character), s indicates second, C_b is the backspace key, C_{Sp} is the special character, C_e is the erased character, len indicates length, max indicates maximum, $substr$ is the symbol for substring, and D_{th+} means the actual acceleration[2] occurs towards an arbitrary direction of the device above a threshold value.

[2] $actual\ acceleration = \sqrt{x^2 + y^2 + z^2}$; x, y, and z are the scalar component of the X, Y, and Z axis of the accelerometer sensor, respectively.

$$Typing\_speed = \frac{1}{s}\sum_{i=1}^{s} C \tag{6.4}$$

$$Backspace\_key\_press\_frequency = \frac{1}{s}\sum_{i=1}^{s} C_b \tag{6.5}$$

$$Special\_symbol\_frequency = \frac{1}{s}\sum_{i=1}^{s} C_{Sp} \tag{6.6}$$

$$Maximum\_text\_length = max(len(substr)), substr \notin C_b \tag{6.7}$$

$$Erased\_text\_length = \sum_{i=s-1}^{s} C_e \tag{6.8}$$

$$Touch\_count = \sum_{i=s-1}^{s} C + 2\sum_{i=s-1}^{s} C_b \tag{6.9}$$

$$Device\_shake\_frequency = \sum_{i=s-1}^{s} D_{th+} \tag{6.10}$$

Among the seven features, the 'Special symbol frequency' was discarded as it is not used in the target application. This is because the special symbols like emoji and smiley are sometimes used just for the sake of formalities, and sometimes even for hiding the actual affective states. The 'Erased text length' was also ignored because it is not possible to know the exact number of characters are erased during typing with a virtual keyboard – multiple characters can be deleted with a single backspace key press. We, therefore, proceeded with the remaining five features (Sl. No. 1, 2, 4, 6, and 7 in Table 6.4) for building and validating the model for the detection of affective states of the users from their typing pattern on the virtual keyboard on smartphone and tablet screen in the educational setting.

Among these five features, we required a threshold value for the 'device shake frequency' feature. Otherwise, it was difficult to determine how much 'acceleration' (along an arbitrary axis) of the device could be considered as a shake. The absolute shake value is obtained when a trigger[3] is detected. The trigger refers to the automatic detection of the acceleration. The android version we worked on (Marshmallow 6.0.1), checks for a trigger every 0.02 seconds. The absolute shake value is calculated as the actual acceleration from the readings of x, y and, z axes of the

[3]https://developer.android.com/reference/android/hardware/TriggerEvent.html

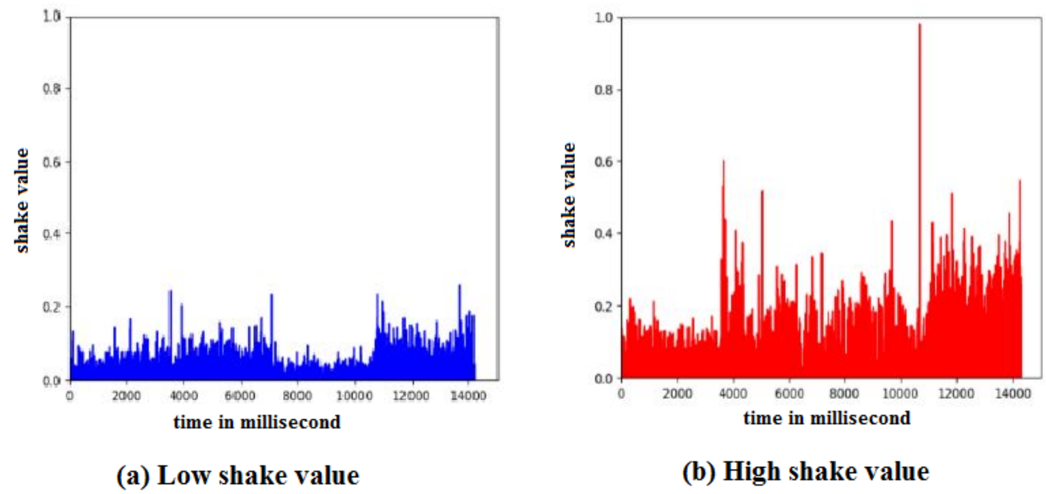


Figure 6.5: Low and high shake value

accelerometer sensor.

We calculated the threshold value through a pilot study with a separate group of ten volunteers (five female, mean age = 25.1 years with SD = 3.14). In the pilot study, all the participants were asked to play the same game (which was developed for data collection for this model) twice per level. Once they were asked to type the target text, keeping the device as still as possible. A second time, the users were allowed to handle the device as they usually do it while typing.

Figure 6.5 represents the plots obtained corresponding to low and high shake. It clearly indicates that the shake values in the usual typing occasion are much higher than the same on the first occasion where the device was kept still as much as possible while typing. This indicates that the data can be naturally clustered into two separate groups. Therefore, we calculated the shake threshold as the mean of the two cluster centers. Cluster centers were decided by taking the means of the shake values in the individual clusters. The threshold value was found to be 0.04.

6.3.2 Data Analysis

The affective data collection as per the specified features for the model, through a novel game, has been presented in the previous chapter (**Chapter** 5, Section 5.4), where all the controlled experiments have been put together. The inducing capability of the game has also been tested in the same way as described in 6.2.3.

Table 6.5: Sample of collected data for Type-Affect model

Timestamp	Typing speed	Backspace keypress frequency	Maximum text length	Touch count	Device shake frequency
2017-11-03-18.21.25.	0.428571	0	0	0	0
2017-11-03-18.21.26.	0.571429	0	1	1	0
2017-11-03-18.21.27.	0.142857	0	1	1	286
2017-11-03-18.21.28.	1.857143	0	1	1	97
2017-11-03-18.21.29.	1.428571	0	0	0	122
2017-11-03-18.21.30.	2	0	1	1	72
2017-11-03-18.21.31.	1.571429	0	1	1	51
2017-11-03-18.21.32.	1.571429	1	0	1	63
2017-11-03-18.21.33.	1.166667	1	1	2	65
2017-11-03-18.21.34.	1	0	4	4	0
2017-11-03-18.21.35.	1.857143	0	2	2	56
2017-11-03-18.21.36.	2.285714	4	0	4	0

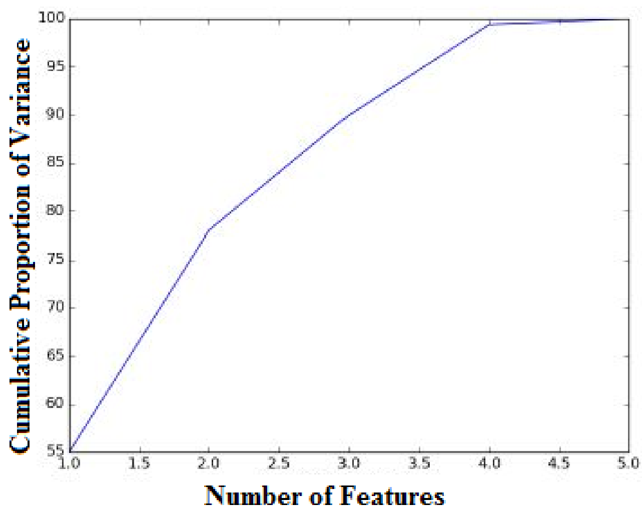


Figure 6.6: Result after applying PCA

A sample of the collected data can be found in Table 6.5.

We analyzed the collected data (a) to minimize and finalize the number of features, (b) to decide the appropriate time slice for which the final set of selected features should be taken together before applying the classifier, and (c) to decide the machine learning algorithm(s) to be adapted for the model. The detailed analyses are described below.

(a) Feature Reduction – We used the Principle Component Analysis (PCA) to identify the number of features that could be rejected. We plotted a graph for the cumulative proportion of variance against the number of features (Figure 6.6).

It was observed that the curve became almost parallel to the x-axis after four features. Hence, we decided that one of the five features could be rejected. Therefore, our final set of features could be of size four. We applied the 'f-regression technique' to identify the feature which could be discarded among the five features. The f-regression technique calculates the correlation between two variables. It selects the features one by one, which has a high correlation with the output variable (valence and arousal level, in our case), and at the same time very less dependency on the already selected features. We observed that except the 'backspace key press frequency', for all the other four features: 'typing speed', 'touch count', 'maximum text length', and 'device shake frequency' had high correlations with valence and arousal level. Thus backspace key press frequency was rejected from the final set of features.

(b) Identification of Appropriate Time Slice – We also identified the appropriate time slice for which all the features should be taken together before applying the classifier. We empirically found out the appropriate time slice by statistically testing the following hypothesis for each time slice from one second to ten seconds.

The null hypothesis corresponding to each feature was:

H_0**:** *For feature X; the mean of feature values of X for low arousal is the same as the mean of feature values X for high arousal.*

Hence, the alternative hypothesis for each feature was:

H_1**:** *For feature X; the mean of feature values of X for low arousal is different from the mean of feature values X for high arousal.*

We applied the t-Test for each feature for all the time slices to find out the appropriate time slice for which the hypothesis for all remaining features was rejected. In a similar way, the statistical test was done for valence as well. The results are presented as follows (Table 6.6).

The result (presented in Table 6.6) showed that the window size of 7 seconds had the minimum required significance level ($p < 0.05$) to reject all the null-hypothesis. Moreover, it was the only window size for which null hypothesis for all the features got rejected. We, therefor, chose 7 seconds as an interval, i.e., in every seven-second, our model should predict the affective states based on the four features.

Table 6.6: Hypothesis testing for different time slices

Time Slice	Minimum significance level to reject all null hypothesis using t-Test	
	p-value for valence	**p-value for arousal**
1 second	0.09	0.08
2 second	0.08	0.09
3 second	0.08	0.08
4 second	0.14	0.13
5 second	0.13	0.15
6 second	0.16	0.14
7 second	0.04	0.04
8 second	0.16	0.17
9 second	0.08	0.12
10 second	0.18	0.20

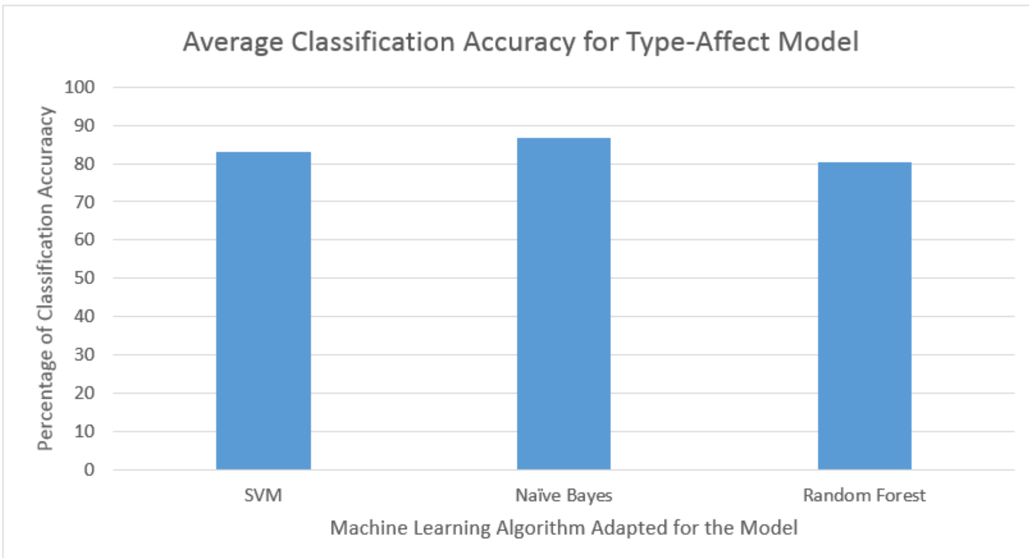


Figure 6.7: Average classification accuracies for the different machine learning algorithm adapted for the Type-Affect model

(c) Choice of the ML algorithms – For this model also, we used the multilevel Support Vector Machine (SVM) and Naïve Bayes classifiers due to the same reasons as described in the case of the Touch-Affect model. However, as the number of features used in this model is four, we thought of exploring the Random Forest (RF) algorithm instead of the DT in this case.

6.3.3 Training and Testing the Model

Like the Touch-Affect model, for the same reason, we used the LOSOCV for training and testing the model. The average classification accuracies, while adapting different machine learning algorithms for the model, are depicted in Figure 6.7.

Although we got high classification accuracy at the time of testing the model with the affective data collected through the game we have developed, like in the case of the Touch-Affect model, we additionally validated the Type-Affect model with the EEG data in the same way we did it for the Touch-Affect model as described in 6.2.4.

6.3.4 Discussion

As specified earlier, like the Touch-Affect model, in this case also the data collection method was purely unobtrusive, which helps to reduce the inaccuracy in emotional data. The model detects the specific affective states of the users of small hand-held touchscreen devices without their knowledge. This minimizes the imitation in emotion. Moreover, no extra hardware and sensors are required to capture the emotional data in our model. Only typing patterns and dynamics are sufficient to identify the users' affective states. The model is therefore expected to be suitable for the particular blended learning platform (as specified in 6.4.2), which we have considered as the basis of the Vedinkaksha framework. We claim our model to be ubiquitous as the affective states are identified without the knowledge of the user by a device that stays always with the user, and it does not matter where the user is. The number of features to be used in the model has been minimized by excluding the unnecessary and application-specific features used by earlier works [76] [77]. This makes the model simple, generalized, and inexpensive in terms of computation. This in turn increases the adaptability of the model in the target applications as well as also in various other applications. Most of the related works (e.g., Lee et al [76], and [77]) were validated based on the users' feedback which may not be always reliable [104]. Our additional validation studies (consist of EEG signals of the participants) demonstrate that our proposed affect induction method can induce the intended specific affective states. These studies also validate that our proposed minimalist model can detect users' affective states with high accuracy.

Despite the novelty and strengths of the model, there is still scope for future work. Our model fits best in our target application area (e.g., in education [141]), as it is able to detect only the four affective states based on the Circumplex model. However, as mentioned while discussing the Touch-Affect model, one may intend to refine this model for identifying users' affect and emotion in a more comprehensive way for some other applications, if required.

6.4 Smart-Affect: Model to Detect Affective States from Basic Smartphone Interaction

We have built and validated two computational models (Touch-Affect and Type-Affect) for identifying the affective states of the users from their behavioral patterns on basic interactions (touch and typing) on small handheld touchscreen devices such as smartphones and tablets. The models have been found to be suitable and useful in our target application. However, predicting the type of interaction (whether it is touch or typing interaction) beforehand is not possible. In other words, it is not certain that a student will be involved in performing a particular type of interaction for a specific time in the assumed blended learning platform. This leads to the problem of applying the two computational models independently in the Vedinkaksha framework for real-time affect detection. In order to address this problem, we have proposed a process model called Smart-Affect, which detects the affective states of the students in real-time in the assumed educational setting. The process model checks the availability and type of interaction for a certain interval on a regular basis. In a particular instance of time, based on the type of the user-smartphone interaction data, the Smart-Affect model chooses either of the two computation models, i.e., either Touch-Affect or Type-Affect model, to compute the affective state of the students. In case, the process model finds none of the two types of interaction data at a particular instance, it shows the latest state of the student. The details of the model are presented as follows.

6.4.1 Model Description

The process model first checks for the presence of touch interaction data. If such data are available, it further checks the type of the available touch data. In other words, it identifies the source of the touch data, i.e., whether the data are generated through typing on a virtual keyboard or through general touch interaction on other portions of the screen. This can be identified through monitoring the system log for the status of the keyboard application. If it is identified that the keyboard is being

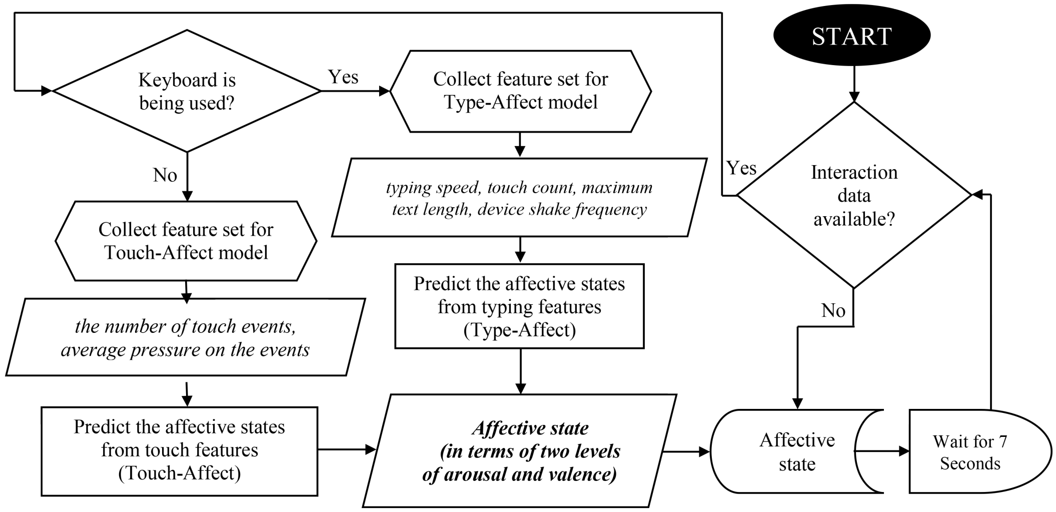


Figure 6.8: Working principle of the Smart-Affect model to detect the affective states of the user from basic smartphone interaction data

used, typing features are collected and are sent to the Type-Affect model, which is responsible for detecting the affective states from the typing behavior of the users. Otherwise, the touch features are collected and sent to the Touch-Affect model, which detects the affective states of the users from their general touch behavior. If the process model recognizes that there is no interaction data available, it shows the last identified states, which is stored in temporary memory. It may be noted here that the process model checks the affective states after every seven seconds. This is because we have assumed that once a particular emotion is induced, it takes some time to react, and stays for some time (although the duration is very short). We chose a seven-second time slice in particular because during the development of the Type-Affect model we found it as the most appropriate time window for detecting the affective states from the typing pattern of the students (explained in 6.3.2***(b)***). Moreover, from the literature, we have found that the type of affective states, which have been considered by us, stay for ±5 seconds [16] [103]. Although the Type-Affect model considers the typing data for the entire seven-second window, the Touch-Affect model considers the touch interaction pattern for the last three seconds of the time window while detecting the state in real-time; because that particular model was trained in such a way. The overall process model, in the form of a flow diagram, is shown in Figure 6.8.

6.4.2 Model Validation in Target Application

We conducted a validation study for the model in our target application with the help of EEG data of fifteen participants. The details of the data collection process for the validation study have been reported as CE7 in Chapter 5 (Empirical Study Details). In the study, we collected both the EEG data (raw EEG signals and derived affective states) as well as the smartphone interaction data from the participants. Later, we detected the affective states of each of the students using our model and compared the states with the sates detected by the EmotivPRO at each instant. It has been found that in 85.2% of cases the states were similar. This high similarity indicates that our model is able to detect the affective states of the students of a blended learning environment. It may be noted here that the EmotivPRO calculates the states in a hundred-point rating scale, whereas our model detects these in two classes ('High/Low' for the arousal and 'Positive/Negative' for the valence). Therefore, we considered the arousal levels 'more than fifty' detected by the EmotivPRO as 'High' and 'less than or equal to fifty' as 'Low' (which was validated by a self-report based study). Similar consideration was made for comparing the valence levels. Table 6.7 represents a sample of the comparison study for the similarity calculation (data for one minute of a particular student has been shown as an example). The result of this validation study also indicates that the two computational models, namely the Touch-Affect and Type-Affect, are useful in the target application area, as they are the two main components of the Smart-Affect affect process model.

In addition to validating the model for the teaching-learning application, we were also interested in validating the model for the other applications to test the generalizability of the model. For this, we collected the touch and typing data, and at the same time the emotional states of the students detected by EmotivPRO while students were performing social networking and instant messaging applications (details of the data collection method can be found in Chapter 5, Section 5.8). We detected the affective states by applying the Smart-Affect model and compared those with the affective states detected by EmotivPRO at each instant. It has been observed that in 84.90% of cases, the detections by our model and EmotivPRO were the same. This high similarity indicates that the model can be applied in other appli-

Table 6.7: Sample of the result of the comparison study for validating the Samrt-Affect model in a blended learning environment

Time instant	**Affective state detected by our proposed process model** ***('Positive' and 'Negative' indicate the level of valence, whereas 'High' and 'Low' refer to the level of arousal)***	**Affective state detected by the EmotivPRO** ***(EmotivPRO detects the states in a hundred point rating scale)***	**Remark**
t1 (10:21:15)	Positive High	Valence (78), Arousal (65)	Similar
t2 (10:21:22)	Positive Low	Valence (70), Arousal (63)	Not Similar
t3 (10:21:29)	Positive Low	Valence (74), Arousal (40)	Similar
t4 (10:21:36)	Positive Low	Valence (65), Arousal (30)	Similar
t5 (10:21:43)	Positive High	Valence (70), Arousal (64)	Similar
t6 (10:21:50)	Positive Low	Valence (71), Arousal (42)	Similar
t7 (10:21:57)	Negative Low	Valence (40), Arousal (31)	Similar
t8 (10:22:04)	Negative Low	Valence (30), Arousal (48)	Similar
t9 (10:22:11)	Negative High	Valence (31), Arousal (61)	Similar
t10 (10:22:18)	Negative High	Valence (20), Arousal (65)	Similar
t11 (10:22:25)	Negative Low	Valence (30), Arousal (55)	Not Similar
t12 (10:22:32)	Negative Low	Valence (32), Arousal (46)	Similar

cation areas where the required touch interaction data are available. This particular study is not directly related to this thesis but is very important, as its result hints that although the model is trained with the affective data collected through games, it works in other applications where the basic user-interaction data are available as input. Our assumed blended learning platform is one such application.

6.4.3 Discussion

The primary functionality of the proposed process model is to combine the two novel affective computational models so that they can be adapted in the Vedinkaksha framework. It helps the framework in choosing a particular computational model, among the two, to detect the affective state of a student for a particular instance of time. As it is a mare process model, we majorly focused on reporting the process description instead of the details of the theory analysis which has been done while presenting the two computational models: Touch-Affect and Type-Affect. This is because they are the most crucial components of the process model and our primary contributions in developing affect-detection models for the proposed framework. Nevertheless, the process model plays a vital role, as only through it the two affective computational models can effectively work in the Vedinkaksha framework.

We validated the model for target application empirically through a controlled experiment. Sixteen demo lectures were delivered for this purpose. However, in the future, the effectiveness of the model should also be observed in the classroom systems when real lectures will be delivered. Nonetheless, we believe that the model will definitely work there as we considered the behavioral data of the students, which incorporates only the basic user-smartphone interaction (touch and typing) pattern, for the detection of the affective states. The result of the generalizability study is the basis of such a belief. This is because there we tested and validated the model not only for the target application but also for two other general applications (instant messaging and social networking) where these basic interaction data are available. Nevertheless, we intend to verify the effectiveness of the model at the time of practicing a sensitive classroom system, which has been built under the framework of Vedinkaksha and has been reported in the next chapter (**Chapter** 7).

6.5 In-Activity: Model to Detect Involvement in Study-related Activities

In addition to the three affect detection models, we have built and validated another computation model to address the research issue related to the detection of students' involvement in classroom activities. Assuming a blended learning platform in the BYOD paradigm ([130] [24]) as the basis of the Vedinkaksha framework, we define students' involvement in the teaching-learning process as follows.

> *A student is said to be involved in the teaching-learning process if s/he performs any study-related activity with her/his smartphone; if the student performs any non-study related activity using the smartphone, s/he is considered to be not involved.*

To build and validate the model for detecting this involvement, we have pursued the following methodology. We have first identified all the possible activities, which are performed frequently by the students with their smartphones inside the classroom. We have identified the activities through a survey conducted by us throughout India [125]. After identifying the frequent activities, we have categorized them into two groups namely, the 'study related' and 'non-study related' activities. The categorization has been done based on the literature [89] [110]. Utilizing the knowledge of the frequently performed activities and their categorization, we have designed a set of tasks for a controlled experiment to collect device handling and resource utilization data while students perform the two groups of activities. This is because we observed different patterns of these data for the two groups of activities, which was confirmed by a pilot study as well [127]. We have collected these data from thirty-nine students through a controlled experiment (CE1). The data were labeled and analyzed to finalize the set of features for a machine learning model, as well as for validating the same. The details are described next.

6.5.1 Choice of Data as Feature Value

By observing the students in performing different smartphone activities, we hypothesized that students' interaction behavior while performing study-related activities is different from that of non-study related activities. We observed through a pilot study that the device handling patterns for these two groups of activities are different [127]. In the study, we asked five UG and PG students (two female, mean age = 25.5 years with SD = 3.17) to perform fourteen frequently performed activities one by one, each for approximately five minutes. While the participants were performing the tasks, a background app was logging the sensory and resource utilization data. An observer was engaged in noting down the starting and ending time of the individual tasks during the experiment. The clocks of the observer and the smartphone through which the data were collected were synchronized before the commencement of the experiment. After manually labeling the collected data and plotting them in a graph, we have observed that some of the students tilt and shake the device more for non-study related activities compared to study-related activities (Figure 6.9 and Figure 6.10).

We also observed that memory utilization may differ during the performance of the two groups of activities (Figure 6.11). This is because most of the non-study related applications (e.g., games, social networking) require higher memory consumption. The battery temperatures have also been found different for the executions of the study and non-study related smartphone activities. We, therefore, thought of collecting the values of embedded sensors such as accelerometer and gyroscope (through which it is possible to systematically observe the movement of the device), along with the amount of primary memory consumption and the tempera-

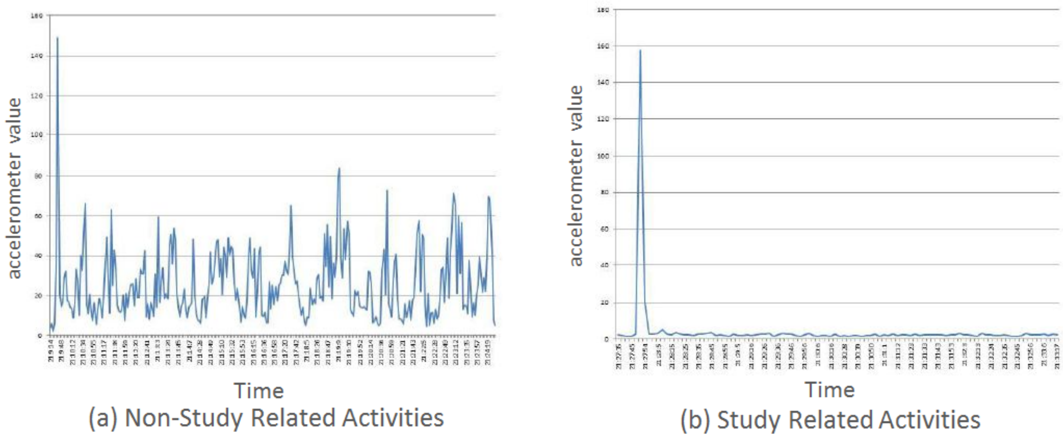


Figure 6.9: Accelerometer value for study and non-study related activities (for Participant-I)

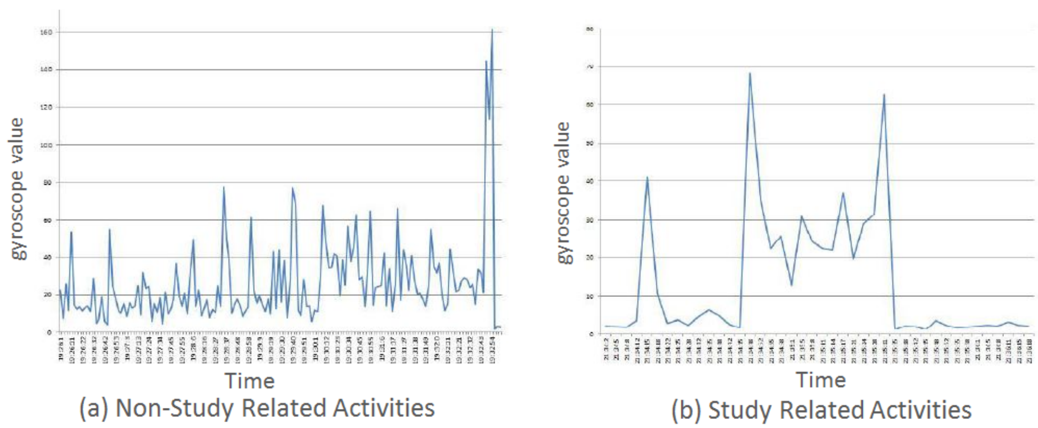


Figure 6.10: Gyroscope value for study and non-study related activities (for Participant-II)

ture of the battery for these two groups of activities. Analyzing all these values we could define the appropriate set of features, using which a machine learning-based algorithm could classify students' activity in one of the study-related and non-study related activities without accessing the actual contents. We collected the values of accelerometer, raw accelerometer, gyroscope, memory consumption, and battery temperature for this purpose. The data collection process has been presented in the previous chapter (**Chapter** 5, Section 5.2) where all the controlled experiments have been presented together.

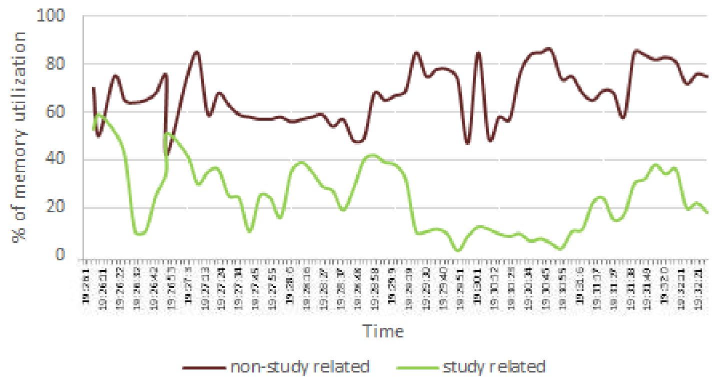


Figure 6.11: Memory utilization for study and non-study related activities (for Participant-V)

6.5.2 Data Analysis

We analyzed the data for defining the final set of features to be used in a machine learning algorithm, and consequently for selecting appropriate classifier(s) to be used for the model. For the analyses, we required to label and preprocess the collected data, as the values logged by the background app (sample of which has been shown in Table 6.8) might not be directly used as the features. For example, it is not possible to capture the actual acceleration in an arbitrary direction by logging the sensor values. However, it is possible to capture the scalar components along X, Y, and Z axes of the acceleration, from which we can calculate the actual acceleration. This is true for capturing the rotational speed through the logs of the gyroscope sensor as well. The labeling, preprocessing and further analysis of the data for building the model is described next.

A. Data Labeling

We manually labeled the data as study-related and non-study related activities based on the categorization of McCoy [89], and Roberts and Rees [110]. It may be noted that the participants were not continuously involved in the tasks in the experiment. We assigned some additional tasks (e.g., reporting search results) so that they perform the actual tasks seriously, as discussed during the description of the task design for CE1. Therefore, the exact times when the participants were

Table 6.8: A sample of collected data from a particular participant

Timestamp	**Accelerometer (X)**	**Accelerometer (Y)**	**Accelerometer (Z)**	**Gyroscope (X)**	**Gyroscope (Y)**	**Gyroscope (Z)**	**Raw Acceleration (X)**	**Raw Acceleration (Y)**	**Raw Acceleration (Z)**	**Battery Temperature**	**Memory Used**
2018-07-05-15.28.35.	0.7086792	5.171478	8.868118	0.000992	-0.10462	-0.07498	0.051935613	0.01650095	0.48681736	29	565
2018-07-05-15.28.36.	0.8236084	5.803543	7.603988	-0.02244	-0.05563	-0.01426	0.108957	-0.05315	-0.44991	29	566
2018-07-05-15.28.37.	1.1109161	5.535400	9.653427	-0.06504	0.073257	0.072021	-0.00269	-0.04849	-0.28530	29	566
2018-07-05-15.28.38.	0.9768371	5.765243	7.776367	-0.04693	0.008285	-0.01638	-0.04231	-0.02284	-0.23331	29	566
2018-07-05-15.28.39.	1.0726013	5.516235	8.65744	0.129898	0.307617	0.021957	0.185187	0.020753	0.120959	29	566
2018-07-05-15.28.40.	0.7278442	5.669464	8.12114	0.156524	0.238388	-0.22731	0.535698	0.140076	0.449176	30.8	567
2018-07-05-15.28.41.	0.9193725	5.669464	8.159439	-0.01072	0.005081	0.002777	0.058650	0.012304	0.114509	30.8	567
2018-07-05-15.28.42.	1.1683655	5.516235	8.944733	-0.21205	0.417343	-0.01106	-0.01846	-0.06129	0.322383	30.8	567
2018-07-05-15.28.43.	1.4556732	5.899307	8.274368	0.087280	-0.10250	0.056045	-0.05573	0.070951	0.377250	30.8	565
2018-07-05-15.28.44.	1.3215942	5.803543	7.795517	0.018035	-0.00982	0.001724	-0.04241	-0.00708	-0.05065	30.8	566

on-the-task were required for the labeling. The starting and ending times of each of the tasks and subtasks, which were manually noted down by the observers at the time of the controlled experiment, helped us to remove the portion of the logged data when the participants were off-the-task.

At the time of data labeling, we observed that three of the participants did not complete the majority of the tasks. We also observed that, for two of the users, the data were not properly recorded, although they performed all the tasks completely (this might be because of unintentional closing of the background app, which was responsible for recording the data). These five users' data were not considered for further analyses. Thus, in total, data of 34 participants were used for building and validating the model.

Each of the thirty-four participants performed every activity for 3-4 minutes on average, which resulted in data for approximately twenty-five minutes for individual participant. We recorded the data at each second. This way, we got 48,160 data points. Among these data points, 18,522 were of study related, and the rest (29,638) were of non-study related. The number of the study-related data points was less compared to that of the non-study related data points because in the task-set the number of the study related activities was less than that of the non-study related activities.

B. Feature Set Specification

Our observation on the behavior of the students on performing smartphone activities indicated that students' behavior is different while performing study-related activities from that of non-study related activities. We also observed that the resource utilization for performing the two types of activities may not be the same. The observations were affirmed by a pilot study [127]. In order to capture the movement, rotation, resource utilization, and change of temperature of the device, we defined a set of features to be used in a machine learning algorithm.

While analyzing the data collected for building the model, we observed that none of the sensory values was able to differentiate study or non-study related activities, unless we consider a time window. This is because we assumed that the device of the student will be shaken more, or resource will be utilized more when a

Table 6.9: Features considered for developing In-Activity model

Sl. No.	**Name of the feature**	**Explanation**
1.	T_{avg}	Average battery temperature of the device for a particular window (in $^{\circ}C$).
2.	M_{avg}	Average memory (RAM) used by the applications for a particular window (in MB).
3.	A_{tot}	Sum of acceleration of the device along an arbitrary direction (in m/s^2).
4.	$A(R)_{tot}$	Sum of raw acceleration (without considering the gravity = 9.8 m/s^2) of the device along an arbitrary direction (in m/s^2).
5.	R_{tot}	Sum of the rotation speed of the device along arbitrary axes (in m/s).

student will perform non-study related activities, compared to that of study-related activities within a certain amount of time. Therefore, a time window was mandatory to define the set of features.

We considered a small window size of seven seconds. This is because we wanted to detect the involvement by observing the behavioral and resource utilization data for a small time window. This will help in adapting the model in real-time applications and systems where time constraint is an important factor. Moreover, especially in our target application, it might not be possible to wait for a long time to get the involvement status, as another component for defining the mental state does not stay for a long time (emotion stays for five to seven seconds only [16] [103]). Nonetheless, we have explored larger windows as well (of up to 60 seconds) to examine whether consideration of a larger window provides better accuracy.

As only the values of the sensors are not sufficient for qualifying as features (as those do not capture the pattern of the user behavior in this context), we defined the features for the particular window size as per Table 6.9.

Among these, the $A(R)_{tot}$ was found to be redundant in the current context due to the consideration of the A_{tot} as a feature. This was because we ultimately wanted to compare the summation of the acceleration along arbitrary axes for study-related activities with that of the non-study related activities. In other words, we were working with the difference of acceleration for the two types of activities. It is therefore immaterial, in the current context, whether the gravity is included in the acceleration or not – if included, it is included for both types of activities. The difference in acceleration for the two types of activities will be the same whether gravity is considered or not. We, therefore, considered T_{avg}, M_{avg}, A_{tot}, and R_{tot} as the final set of features for the model. The features were computed using equations (6.11) – (6.14).

$$T_{avg} = \frac{1}{w}\sum_{i=1}^{w} t_i \tag{6.11}$$

$$M_{avg} = \frac{1}{w}\sum_{i=1}^{w} m_i \tag{6.12}$$

$$A_{tot} = \sum_{i=1}^{w} \sqrt{x(a)_i{}^2 + y(a)_i{}^2 + z(a)_i{}^2} \tag{6.13}$$

$$R_{tot} = \sum_{i=1}^{w} \sqrt{x(g)_i{}^2 + y(g)_i{}^2 + z(g)_i{}^2} \tag{6.14}$$

Where w is the window size, t is the temperature of the battery, and m is the memory consumed at a particular instant of time. The *x(a)*, *y(a)*, and *z(a)* indicate the scalar components along the X, Y, and Z axes, respectively, of the actual acceleration toward an arbitrary direction. Similarly, the *x(g)*, *y(g)*, and *z(g)* indicate the scalar components along the X, Y, and Z axes, respectively, of the actual rotation speed about an arbitrary axis. We can get the acceleration and rotation speed values from the accelerometer and gyroscope sensors, which are embedded in almost every smartphone nowadays.

C. Data Preprocessing

We calculated the feature values for the different window sizes as per equation (6.11), (6.12), (6.13), and (6.14), in an overlapping manner (shifting the window by one sec). The overlapping window was considered to capture the change in feature values irrespective of time, particularly for the features responsible for capturing the movement and/or rotation. This is because it is almost impossible to predict the time when a particular user will move/rotate the device for performing a particular activity.

6.5.3 Building and Validating the Model

A. Feature Analysis

After calculating the feature values using equations (6.11) – (6.14), we observed for patterns in the feature values. As it was very difficult to observe the data patterns

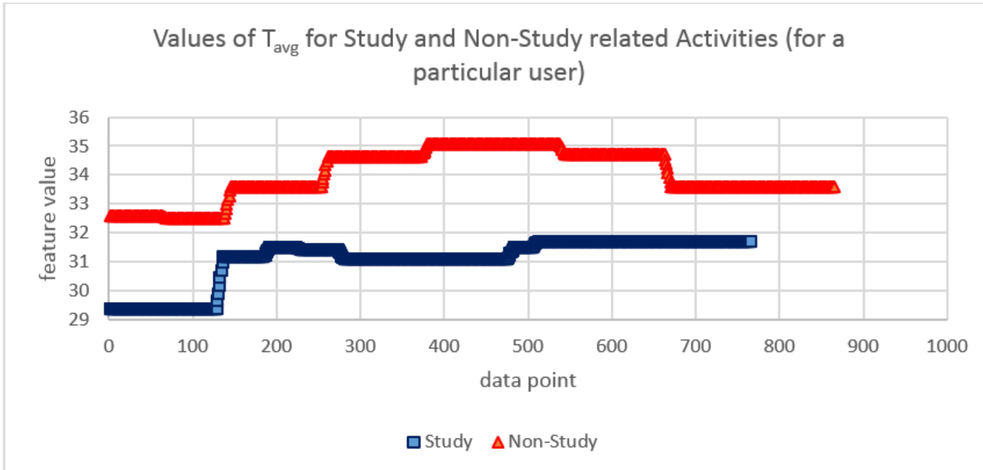


Figure 6.12: An example where a particular feature value (T_{avg}, in this example) for study and non-study related activities are fully distinguishable.

after plotting the 48,160 data points for the four features, we observed the data of each user separately, to have an overall idea about the data. We observed that the feature values for the study and non-study related activities are fully distinguishable in many cases. Figure 6.12 represents the values of a particular feature (T_{avg}, here) for both study and non-study related activities (for a particular user) where the feature values of study and non-study related activities are fully distinguishable. We also found numerous cases, where the feature values were mostly distinguishable. Figure 6.13 is an example of such cases. In very few cases, we found that the feature values were indistinguishable for the two categories of activities. Figure 6.14 is one of the examples where the values of the particular features (M_{avg}, in this case) for the two categories are not distinguishable. In all these figures (scatterplots), triangle shapes (colored in red) indicate the feature values for non-study related activities; whereas square shapes (colored in blue) represent the feature values for study-related activities.

We considered a small, medium, and large window to observe the effect of changing the window size on the feature values for both the study and non-study related activities. The summary of the observations has been reported in Table 6.10 and Table 6.11.

Table 6.10 shows the number of users (and percentage) classified into the study and non-study related categories based on a particular feature. For example, data

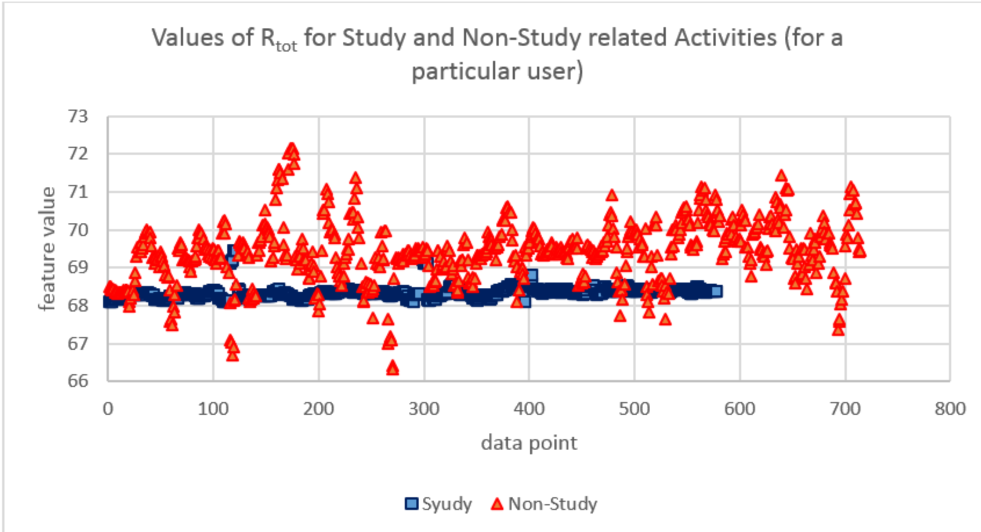


Figure 6.13: An example where a particular feature value (R_{tot}, in this example) for study and non-study related activities are mostly distinguishable.

of 24 out of the 34 participants (i.e., 70.58% of the users) can be classified into the study and non-study groups based on the T_{avg} feature only, if we consider a 7 sec window. Similarly, for the same window size, data of 9 (26.47%) participants can be classified into either of the two groups based on M_{avg} feature only. The entries in the table indicate the roles of the individual features for the classification for various window sizes. For example, most of the users can be classified based on the T_{avg} feature, irrespective of the window sizes. However, the feature is not able to classify all the users. Consideration of the other features is important - when one feature is unable to classify, another may help in that. This is because all the other features have notable roles for the classification. It can further be observed that an increase in window size is not always helpful for every feature. For instance, considering a seven-second window, R_{tot} feature can classify 32.35% of users into two groups. With the increase in the window size from seven to twenty-one, it is able to classify 35.29% of users. However, with the consideration of a sixty-second window, it is able to classify 23.53% of users.

Each entry of Table 6.11 indicates the number (and percentage) of the participants whose data can be classified into either of the two groups based on the number of features mentioned in the column-headings. For example, the value 00

[0]Total = Fully + Mostly.

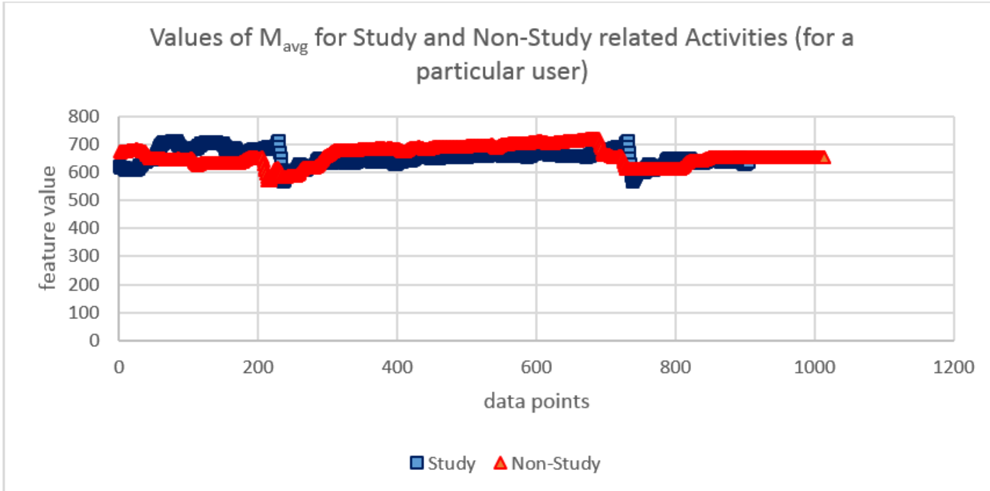


Figure 6.14: An example where a particular feature value (M_{avg}, in this example) for study and non-study related activities are not distinguishable.

Table 6.10: Summary of the observation of the patterns of the data of individual users

Window Size	**Features**	**The number (and %) of users can be classified based on the feature value**		
		Total[0]	**Fully**	**Mostly**
7 second	T_{avg}	24 (70.58%)	24 (70.58%)	—
	M_{avg}	09 (26.47%)	07 (20.59%)	02 (05.88%)
	A_{tot}	07 (20.59%)	03 (08.82%)	04 (11.76%)
	R_{tot}	11 (32.35%)	02 (05.88%)	09 (26.47%)
21 second	T_{avg}	28 (82.35%)	21 (61.76%)	07 (20.59%)
	M_{avg}	08 (26.53%)	06 (17.65%)	02 (05.88%)
	A_{tot}	09 (26.47%)	02 (05.88%)	07 (20.59%)
	R_{tot}	12 (35.29%)	02 (05.58%)	10 (29.41%)
60 second	T_{avg}	28 (82.35%)	21 (61.76%)	07 (20.59%)
	M_{avg}	10 (29.41%)	06 (17.65%)	04 (11.76%)
	A_{tot}	12 (35.29%)	02 (05.88%)	10 (29.41%)
	R_{tot}	08 (23.53%)	01 (02.94%)	07 (20.59%)

in the table (for the 7-sec window) indicates that none of the users can be classified into one of the two classes using all the four features applied individually. The value 07, of the very next cell in the same row, indicates that the data of seven out of the thirty-four participants can be classified into either of the groups based on three of the four features applied individually. However, the values of one feature could not be used to distinguish the group of activities. In other words, we can use any of the three features individually to categorize only seven users at the most. The value 00 in the first column signifies that consideration of all the four features is important – none of the users are there who can be classified based on any of the four features. Similarly, the entry 04 (last column of the first row) indicates that there are four

Table 6.11: Analysis of the importance of features and user behavior

Window Size	The number (and %) of users' data can be distinguished into the two classes based on				
	4 Features	**3 Features**	**2 Features**	**1 Feature**	**No Feature**
7 second	00 (00.00%)	07 (20.59%)	07 (20.59%)	16 (47.06%)	04 (11.76%)
21 second	00 (00.00%)	07 (20.59%)	12 (35.29%)	04 (11.76%)	01 (02.94%)[0]
60 second	00 (00.00%)	07 (20.59%)	09 (26.47%)	17 (50.00%)	01 (02.94%)[0]

participants whose data cannot be classified based on any of the features.

However, we observed that consideration of a larger window may help the classification to some extent. For example, if we consider a window of 21 seconds, one feature (M_{avg}) may help to classify three of these four users' data into one of the categories. We further observed that increasing the window size after a certain value may not lead to any gain. For example, we found a negligible effect of increasing the window to 60 seconds from 21 seconds.

We conducted t-Test (two-sample assuming equal variances), considering α=0.05, to determine the statistical significance of the differences found in the feature values for study and non-study related activities. For conducting the test, we extrapolated the data (for each student) for study-related tasks to make it equal with that of non-study related tasks. This is because the number of data for study-related activities was less than that of non-study related activities in the dataset. We assumed the following null hypotheses for conducting the t-Test.

$H1_0$: There is no difference in the T_{avg} values for study and non-study related activities.

$H2_0$: There is no difference in the M_{avg} values for study and non-study related activities.

$H3_0$: There is no difference in the A_{tot} values for study and non-study related activities.

$H4_0$: There is no difference in the R_{tot} values for study and non-study related activities.

The test results show that the null hypothesis can be rejected in every case (with p=0.00 for T_{avg}, M_{avg}, A_{tot}, and R_{tot}). These indicate that the differences are significant for all the features for study and non-study related activities.

[0]Considering the 21 and 60-sec windows, 3 of 4 users' data were mostly distinguishable into either of the two groups, which were fully indistinguishable for the 7-sec window.

The observations related to the data and feature analyses indicated that the four features can be used to determine the student's involvement in study or non-study related activities. Along with this, the choice of an appropriate classifier was important as our proposed model is a machine learning algorithm trained with empirical data collected through a controlled experiment. It may be noted that the proposed model classifies the users' activities into two groups (study/non-study related activities). In other words, the classification is binary. We observed during feature analyses that the data of the two groups were linearly separable in most of the cases. We, therefore, chose to use the state-of-the-art binary classifier namely, the SVM with linear kernel for our model. We also observed that users' data could be classified based on either one, two, or three feature(s) as indicated in Table 6.11. Hence, we also explored the RF algorithm for our model.

B. Model validation

We used a cross-validation technique to train and test the model. Like the Touch-Affect and Type-Affect models, in this case, also we used the LOSOCV (Leave-Subject-One-Out-Cross-Validation), sometimes called LOOCV (Leave-One-Out-Cross-Validation), as the model has been built with the behavioral data of human users.

Figure 6.15 shows the results of the cross-validation study. The figure depicts classification accuracies for various window sizes. We have shown the highest and average classification accuracies for each of the windows considered for computing the feature values. For example, the light shaded portion of the topmost double-shaded bar indicates the classification accuracy of the Random Forest classifier considering a 60 second time window; the darkly shaded portion of the same indicates the classification accuracy of the SVM considering a 60 second time window; and so on. The results indicate that the model can detect the involvement with very high accuracy (93.69%). At the same time, it can also be observed that the classification accuracies of the two classifiers explored (the SVM and Random Forest) are almost similar, for each window size. Hence, we may conclude that both the classifiers are equally applicable to the model.

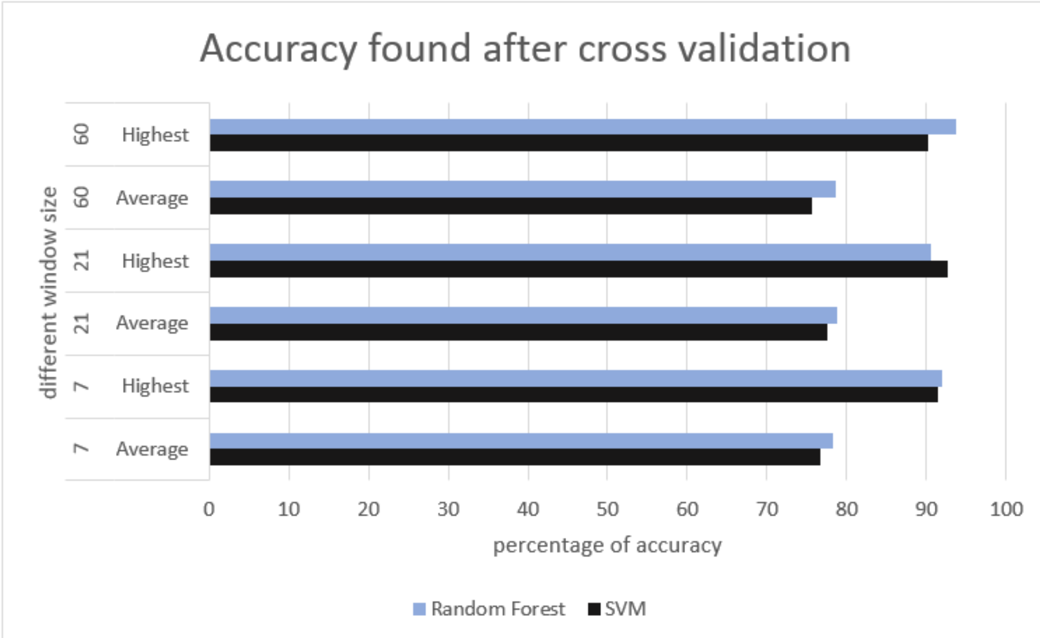


Figure 6.15: Classification accuracy considering various window sizes.

6.5.4 Discussion

Our study results show that the proposed model can detect the 'involvement' with high accuracy. This addresses the problem of inconsistency in detection (which may exist in the manual observation method), and imitation in involvement (which may be present in the self-report method). Being an automated measurement method, the model also addresses the other two important issues of the manual methods: a cumbersome technique to detect the involvement, and the cost of additional manpower. In addition, the model addresses issues in the existing automated measurement methods as well. Some of these methods (e.g., [141], [115]) require expensive and wearable equipment, sensors, and probes to detect involvement. The users may not agree to wear these. The users also might be aware of the fact that they are being monitored. As a consequence, the users might not act naturally – imitation in involvement may occur. Our proposed model identifies the involvement without the knowledge of the user. It also does not require any additional and wearable sensors, equipment and probes to detect involvement. It, therefore, solves the issues of lack of acceptability and imitation in the involvement. Few existing works (e.g., [132]) access private data of the students to detect involvement, which may not be ethical. Our proposed model does not require any privacy-sensitive data. Some of the works (e.g., [138], [141]) also require a special arrangement and fixed set up to detect the involvement. External factors including the ambient light, as well as users' position and gesture, may affect the accuracy in this case. Our model does not require any such fixed setup. The model detects the involvement from resource utilization data and values of the sensors which are embedded in the smartphones, which are part of the blended learning platform. As we propose the model for the BYOD paradigm and the detection is done by smartphone-based on four simple features, the computational and monetary cost is expected to be lesser. This might help to increase the adaptability of the model. As the students are expected to be unaware of the computations for involvement detection, and it does not require any particular computing setup, the model can be termed as ubiquitous as well.

It may be noted that the activities performed in our controlled experiment (i.e., in the task set) were decided based on a survey. The frequently performed activities found in the survey was for the conventional classroom (not for a blended learning platform) where students sometimes use their smartphones as per their wish, whereas we are proposing the model for blended learning platforms. We expect that the activities will be performed in the blended learning platform as well, as the devices will be allowed officially there. However, there may be a possibility of performing new or alternative activities in the blended learning platform and/or in the future. For example, students used to send SMSs as one of the most frequently performed activities earlier, although instant messaging has become more popular nowadays. Nonetheless, the changes in activities might not have any major effect on the overall approach for detecting the involvement. This is because our proposed model classifies the activities into either of the two umbrella-categories namely, study and non-study related categories. It does not identify a specific activity. If any changes in the activities in any of the umbrella categories are found, data can be collected for the new activities and labeled (with the category) following the same approach proposed by us. The feature values can also be computed for those newly collected data based on the equations that we have defined. Retraining the model with these data to update the model is expected to be not cumbersome. In summary, even if some new activities are performed by the students in the blended learning platform (or in the future), the overall approach for detecting the involvement will remain the same – only some minor refinement in the form of retraining the model with updated training data may be required.

We expect that the model will work fine in the blended learning platform for detecting the involvement of the students, where the classroom activities are performed through smartphones. In other words, the model is expected to be useful in the classrooms where the academic policy allows smartphones for classroom teaching. The target student groups are likely to be the undergraduate and postgraduate students of the blended learning platform. The model might not be adapted for teaching the school students as the smartphones are not allowed in many school campus (till 12^{th} grades). However, the model can be adapted in a synchronous distance learning platform (an e-learning platform where both the teacher and students remain live for teaching and learning, but in different geographical locations) to know the involvement of the students, if devices like the smartphones or tablets are used for the pedagogical purpose. In this case, students irrespective of their academic classes might be benefited. Most importantly, the model can be utilized for building a sensitive classroom where students' mental state is predicted based on many important aspects where involvement is one of them (other aspects are affective state, cognitive load, and level of participation in classroom activities).

Although our proposed model is expected to be able to address many challenges present in the stat-of-the-art, there are few limitations and scope for future works. It may be noted that the average accuracies are not very high. This might be because of the presence of data of some users in the training and testing set, for whom the values of all the four features for both the study and non-study related activities were similar (indistinguishable), as observed at the time of data analysis. There may be two probable reasons for this: either it is very difficult to predict the behavior of some users, or the specific users did not participate in the experiment seriously – it might be the case that they were not in a good health condition or with heavy cognitive load. It can also be observed that the accuracy increased when the window size was increased to 21 seconds from 7 seconds. However, the accuracy did not increase much while the window size was increased to 60 seconds. This might be because of the characteristics of the data set, and the type and number of features used for building the model. Nevertheless, we have considered only three window sizes: a smaller window (of 7 sec), a medium window (of 21 sec), and a larger window (of 60 sec). This was because of the context of the applications – we cannot wait for a very long time for the result in some real-time applications. Our target application is one such example. In Vedinkaksha, the mental states of the students are defined and predicted frequently, based on the involvement along with emotion and activity. However, the emotion that is one of the major components for defining the mental state, does not stay for a long time (stays for five to seven seconds [16] [103]. Therefore, identifying the involvement within a short period of time is very important in the current context. Nonetheless, there is a scope for future studies to find out the appropriate window size for the betterment of the model.

We have defined the involvement in the context of a blended learning platform, where it has been assumed that all the classroom activities are performed using smartphones. As per the definition, if a student performs any study-related activity using her/his smartphone inside the classroom, s/he is considered to be involved in the teaching-learning activities. On the other hand, when the student is found to be busy in performing any non-study related activity, s/he is considered as not involved. However, in practice, a student may sometimes communicate/discuss verbally with the teacher present in the blended classroom. In that case, the student can be considered as involved, if the communication/discussion is related to the topic(s) being covered by the teacher. If the student talks about something, which is not related to the current topic, s/he may not be considered as involved. Our model, in its current state, is unable to detect this involvement. Nevertheless, the model is able to detect the involvement and non-involvement when they are unambiguous. When a student is performing some study-related activities with her/his smartphone, s/he is definitely involved. Similarly, when a student is performing some non-study related activities using a smartphone, s/he can be considered as definitely not involved. Nonetheless, future researchers may think about identifying the particular involvement which cannot be detected by our proposed model currently.

In order to design the task-set for the study, we required to identify the frequently performed smartphone activities. We considered a particular survey on smartphone activities [125] for this. This is in spite of the presence of other similar surveys. One of the reasons for such consideration, as mentioned earlier, was that the particular survey was the most recent one – the behavior of the students might change with the passage of time [125] [78]. However, this was not the only reason. There were two more important reasons. One was that the survey was conducted, by ourselves especially for building this model, on the Indian students. This is important as our proposed model has been built using the data collected from the students who are Indians. Behaviors and preferences may change because of the ethnicity of the students as well [125] [124]. Another reason for the consideration of the particular survey was that it covered a large number of participants (1711), compared to the other surveys. The students who participated in the survey were from every part of India. These indicate that the participants might be a good representation of the Indian students. Other surveys were done based on a limited number of participants and restricted to particular geographical locations (e.g., [110] conducted a similar survey on 99 students of Bond University in Australia, Müller et al. [91] conducted their survey on 176 students of the USA, Lee et al. [78] collected data from 95 Korean students for knowing the behavior of the students on the smartphone). Since the behavior of the students on performing smartphone activities may change with the change of their ethnicity, it may be advisable to retrain the model with the data of the students belonging to a particular country where the model will be deployed. One can also make the model adaptive, which might also solve the issue [143] [63] [4].

Currently, we have defined four features for detecting the 'involvement'. The model is able to detect the involvement with a maximum accuracy of 93.69%. However, further exploration is required on the possibility of improvement in accuracy with the addition of new features. In other words, a future study can also be conducted for finding out the optimal feature set to be used for achieving higher accuracy. Also, studies may be conducted to determine the suitability of classifiers other than the SVM and the Random Forest, which we have used, to improve the model performance further.

Table 6.12: Summary of the models reported in this chapter and have been incorporated in Vedinkaksha

Sl. No.	Model Name	Purpose	Feature set used
1.	Touch-Affect	Detection of affective states from touch patterns on small handheld touchscreen devices such as smartphones and tablets.	{the number of touch events, average pressure applied on the events}
2.	Type-Affect	Detection of affective states from typing patterns on the virtual keyboard of small handheld touchscreen devices such as smartphones and tablets.	{typing speed, touch count, maximum text length, device shake frequency}
3.	Smart-Affect	Detection of affective states from basic user-smartphone interaction data, which combine Touch-Affect and Type-Affect models.	{the number of touch events, average pressure applied on the events} OR {typing speed, touch count, maximum text length, device shake frequency}
4.	In-Activity	Detection of involvement in study-related activities in a blended learning platform.	{average battery temperature, average memory consumption, the sum of accelerations, the sum of rotation speed}

6.6 Chapter Summary

In this chapter, we have presented four novel models for detecting the affective states and involvement of the students in the blended learning platform, which are essential for the Vedinkaksha framework. The models have been summarized in Table 6.12. The models have been found to be very accurate and suitable for the target application. The incorporation of these models to build a sensitive classroom system (following the proposed framework) has been discussed in the next and last chapter (**Chapter 7**) which also concludes the thesis with a discussion on future research scope.

"One looks back with appreciation to the brilliant teachers, but with gratitude to those who touched our human feelings. The curriculum is so much necessary raw material, but warmth is the vital element for the growing plant and for the soul of the child."

Carl Jung (1875 – 1961)
Swiss psychiatrist and psychoanalyst

7 Conclusion and Future Scope

We have proposed a computational framework for a sensitive blended learning platform, which detects the mental states of the students based on their affective states, involvement, and level of classroom activities. To detect the affective states and involvement in a blended learning platform, we have built several computational models. The models are found to be able to detect the states with high accuracy and are compatible in a blended learning system. We have also implemented a high-fidelity working prototype for a sensitive classroom system following the proposed framework to validate the same. This chapter concludes the thesis with the description of incorporating the models to implement the sensitive classroom system following the framework, along with a discussion on the limitation and avenues for future research.

7.1 A Sensitive Classroom System

We have implemented a working prototype for a sensitive classroom system based on our proposed framework for a sensitive blended learning platform called Vedinkaksha. The system is described as follows.

7.1.1 System Design

We have designed the system following the proposed framework that has been presented in **Chapter** 3. The system provides interfaces for all the essential classroom activities and at the same time, detects the mental states of the students. The system utilizes the models (which have been developed by us and are reported in the previous chapter) as the key components for computing the mental states of the students. In addition to the mental state detection, it also computes the understanding and learning levels of the students in real-time. The system also has a provision for visualizing the computed states in a simpler form, so that those can easily be understood by the teacher with minimal effort during the lecture hour. The Level-1 DFD of the system is shown in Figure 7.1 (the detailed design, i.e., Level-0 to Level-3 DFDs, has been reported as **Appendix** B).

7.1.2 Implementation Details

We implemented the system in a client-server architecture. An Android app (called "vedinkakSa") has been built for the clients, which provides the interfaces for all the essential classroom activities, and at the same time, collects the input data for sensitive modules of the system. The server is responsible for computing sensitive information and pass the information, along with the teaching and examination materials, to the authentic users. The server also takes some smart actions such as alerting and/or motivating the students, alerting the teacher, and augment the classroom with identified states for visualizing the classroom. The system has broadly six modules namely, the *student's mental state detection*, *classroom visualization*, *alert generation*, *understanding and learning level detection*, *exam & quiz conduction*, and *live classroom management*. Some of these modules provide the system features for availing input data for the sensitive modules which compute the sensitive information by utilizing the models we have developed, whereas some of them take some smart actions based on the identified information. Through the course of offering the sensitivity in this way, the overall system also provides features for performing all essential classroom activities which may help to attract the students to perform the activities through these features, so that the system may avail sufficient input

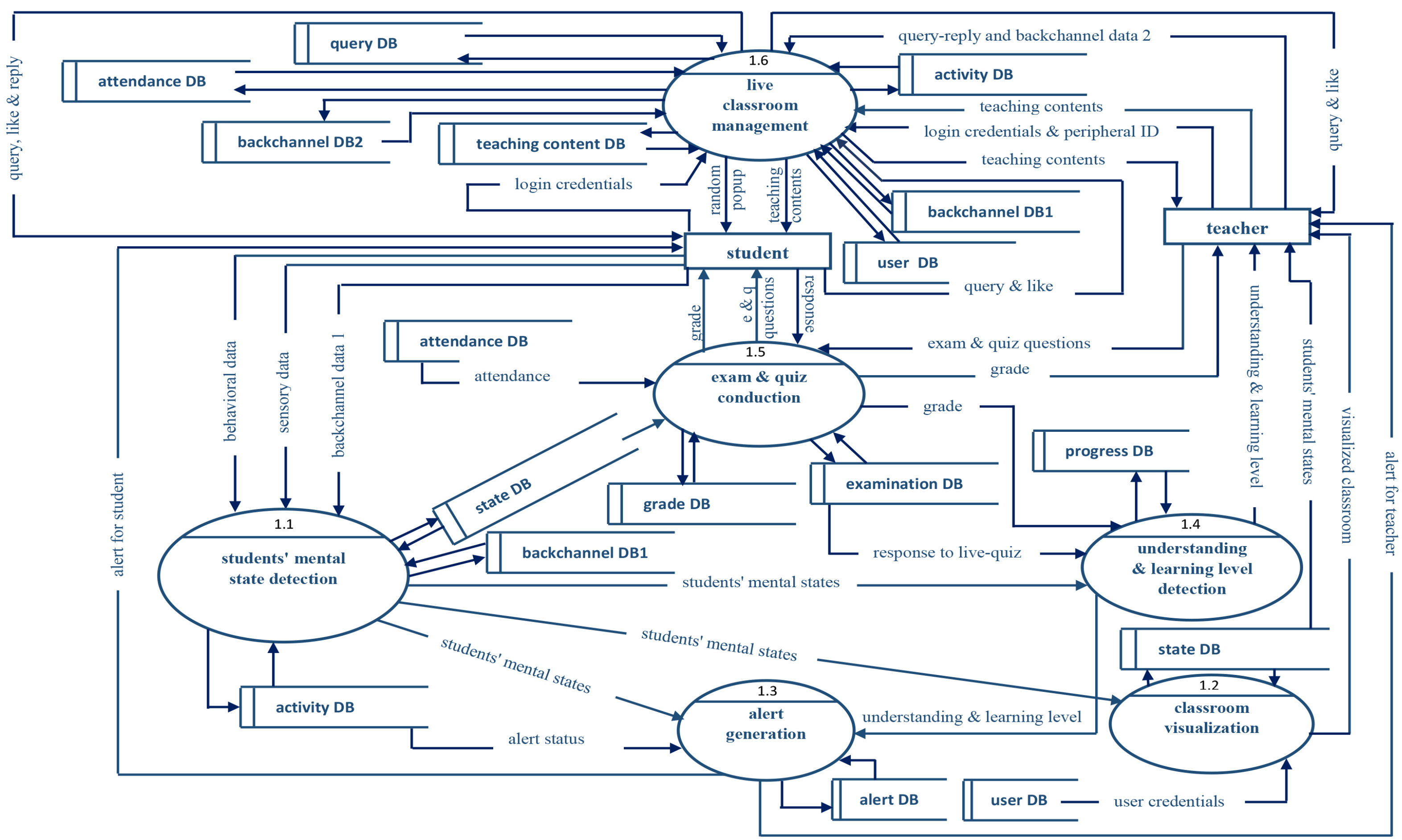


Figure 7.1: Level-1 DFD of a sensitive blended learning platform under the umbrella of Vedinkaksha

data for computing the sensitive information.

A. System Features for Availing Input Data for Sensitive Modules

All the input data which are required for the sensitive modules of the system are availed through two modules namely the 'live classroom management', and 'exam & quiz conduction'. The details of the input data collection through these modules and the purpose of the same are presented next.

(a) Live classroom management – This module is responsible for providing the interfaces for lecture delivery and classroom interactions. This has been decomposed into four submodules namely, *login management*, *lecture delivery*, *query management*, and *activity management*.

Using the login management feature of the system, a user can securely enter into a live classroom session. The teacher is required to provide her/his unique identification number (ID) along with a password. In addition to the user ID and password, the students are required to scan a Quick Response (QR) codes during the login process, which are pasted on the desks in the classroom and contains the seating position. The QR codes are pasted on the desks in the classroom and contain the seating position. Also, during the sign-up process, each user is required to upload a profile picture. Through the log-in module, the system can avail of some user credentials (user ID, profile picture, seating position) which are used by the other modules of the system, both for computing the sensitive information as well as for utilizing the same. To specify, the user IDs of the students are required for all the other modules. Whereas their user IDs and profile pictures along with the QR codes are required for the 'classroom visualization module'. Similarly, the user ID of the teacher is essential for live classroom activity management and an exam or quiz conduction. It is also required for passing sensitive information to the specific teacher. Screenshots for the login interface and main menu after successful login are shown in Figure 7.2.

The lecture delivery panel is required for the live delivery of the lecture materials. It delivers the current slide to each of the user's device. In other words, it appropriately synchronizes the slides so that every student gets the slides live, which is currently being displayed on the teacher's device. The lecture delivery panel does

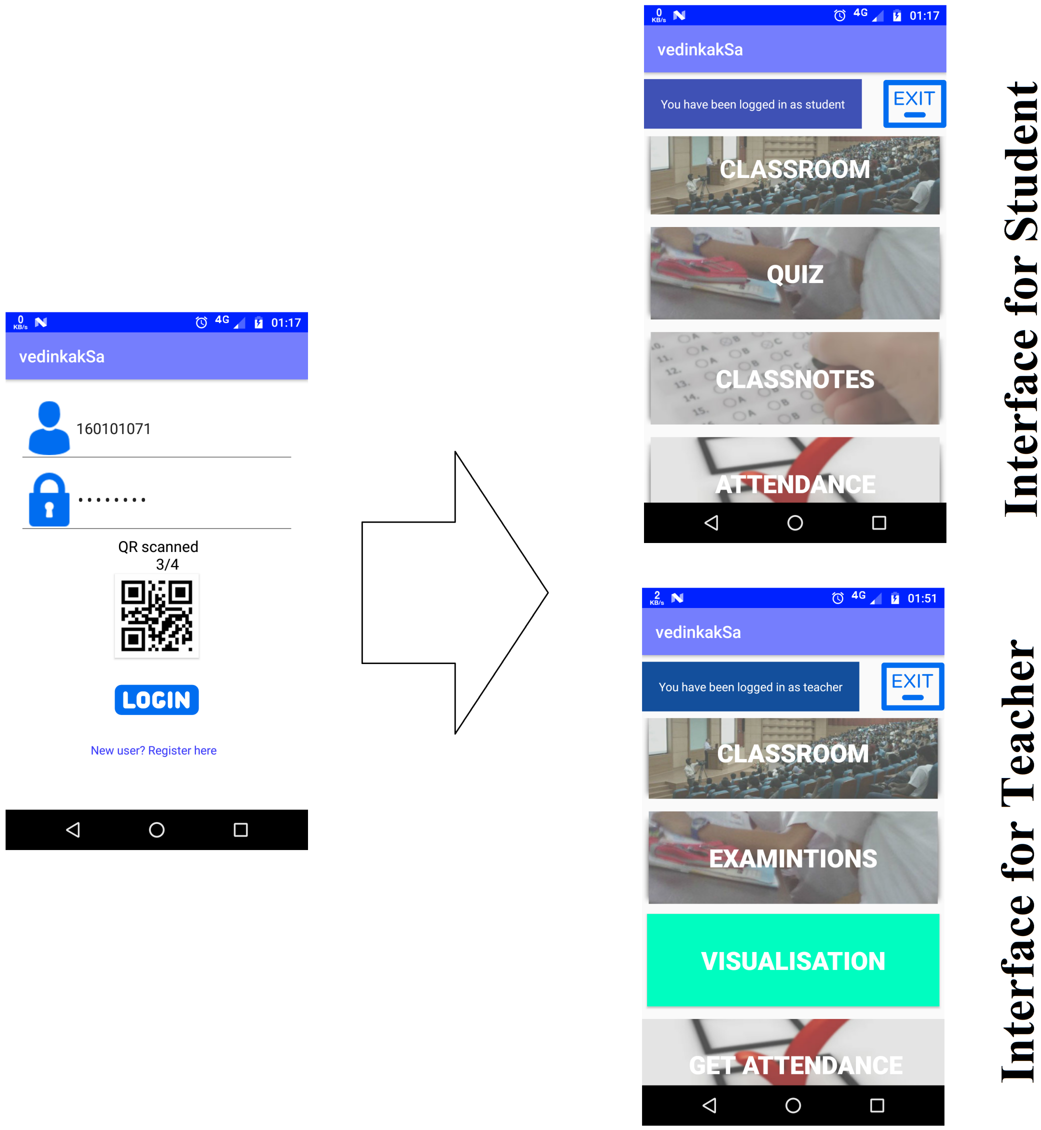


Figure 7.2: Login page with the main menu for both the students and the teacher.

not provide any input for the sensitive modules but its integrated panel (i.e., the query panel) does so. It can be noted that keeping the query panel and the lecture delivery panel on the same screen is beneficial [130].

Through the query panel, all the queries and 'likes' are managed. During the lecture period, students can make queries to clear their doubts. If a student has the same query which has already been asked by another student, the student can 'like' the particular query instead of retyping the same. The 'like' feature helps to prioritize the queries and to avoid the redundancy in queries. The queries are displayed to all the devices in sorted order based on the number of 'likes', as shown

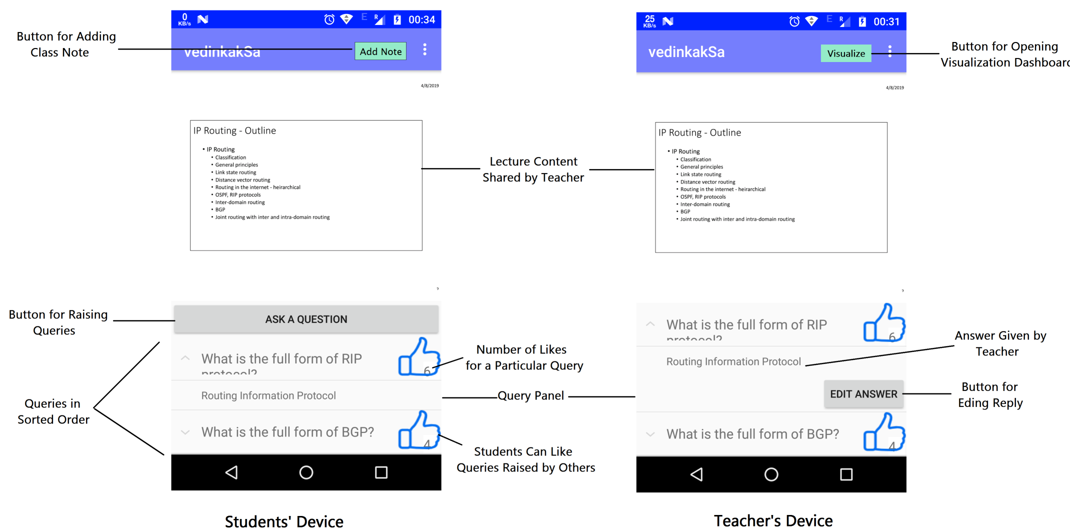


Figure 7.3: Query panel along with the lecture delivery panel for both the students and teacher.

in Figure 7.3. The queries can either be replied by the teacher through the query panel or be clarified verbally, as per the wish of the teacher. In this way, the system can have touch and typing data of the students, which are utilized as the inputs of the affect-detection model. As asking of questions directly or indirectly (by liking a query) are one of the classroom activities, through this feature the system can have the activity logs as inputs for defining activity level as well.

The 'activity management' module stores the backchannel data in the 'backchannel database' and identifies important activities from them in the context of the blended learning system. The activities include the number of queries asked by a student and the time of asking the queries, voting the queries through 'likes', times of responding the random popups, class note taking status (time and duration), and alert sending information (when are they sent, and whether they are acted on or not). These activity details are stored in a database called 'activity database'. The 'activity management' module consists of four submodules namely, the *attendance management*, *random popup management*, *class-note management*, and *activity details updating*. The 'attendance management' module records attendance by observing the login status and duration along with the approval from the teacher and responses from students. The 'random popup' management module generates pop-up messages in random intervals to check the *attentiveness* and send

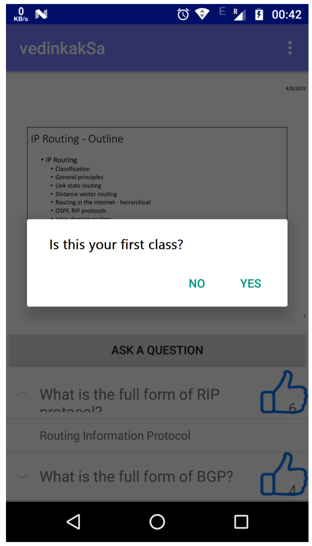


Figure 7.4: Random popup to observe the attentiveness of the students.

their responses to the 'activity details updating' process to store them in the 'activity DB'. The 'random popups' automatically disappear after five seconds within which the students are supposed to respond (Figure 7.4). In addition to updating the response history of the random popups, the 'activity details updating' process also updates the status of other activities such as responses for 'live-quizzes', query posting and 'liking' history, and class note-taking history. These are required for defining the level of classroom activities which is one of the considerations for defining the mental states of the students. The class note-taking facility is provided by the 'class-note management' process. It provides a scratchpad that allows the students to take notes during the lecture (Figure 7.5). The class notes are saved offline in the student's device and organized according to the lecture slides for which they have been taken. The notes can thus be viewed as and when required. Class-note taking is one of the classroom activities, and hence it can be utilized by the system as an input for defining the level of activities. Students are required to type on a scratchpad to add a note. The system, therefore, can have typing data through this feature, which can be used as an alternative source for the inputs of the Type-Affect model. This is because, many a time, a student may not post a query but may note down some important information for future reference.

(b) Exam & quiz conduction – The 'exam and quiz conduction' module conducts all the examinations and quizzes along with their evaluations and the grade

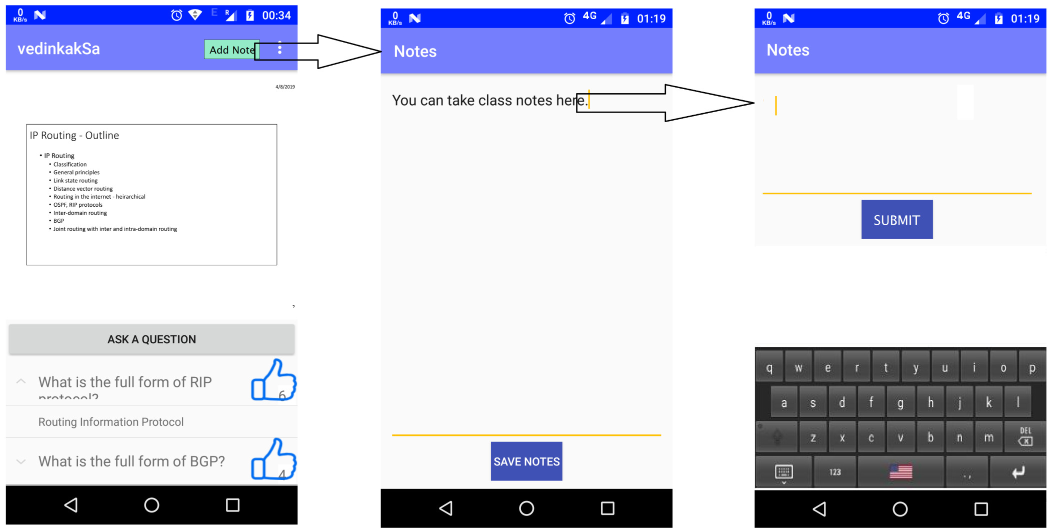


Figure 7.5: Interface for taking class notes for future reference.

processing task. It also manages the homework and assignment submission and evaluation. All these tasks are executed through four processes namely, *the exam & quiz processing*, *grade processing*, *live-quiz processing*, and *HW & assignment processing*. The 'exam & quiz processing' submodule provides interfaces for taking the examination, adding questions for quizzes and exams, viewing the existing questions and quizzes, commencement of new quizzes and exam and their automatic evaluations. When a teacher adds a question along with its answer, it is stored in a database called 'examination database' in the server. Currently, the system supports three types of questions: the *MCQ* (multiple options are provided among which students choose one), *exact string match* (no option is given but the answer should be an unambiguous string), *partial string match* (stored keywords are matched with the words in the answer-string). After selecting a set of existing questions (or a predefined quiz), when the teacher starts an exam/quiz, the questions are got displaced to students' devices. When the students answer to the questions, the responses are saved in the examination database. The responses are then get evaluated based on the previously stored answers given by the teacher, and the scores earned by the individual students are sent to the grade processing submodules. The screenshots of a few interfaces for the 'exam & quiz conduction' module have been shown in Figure 7.6. Also, students can submit their homework and assignment through the 'HW & assignment processing' submodule, which are stored in a server database called

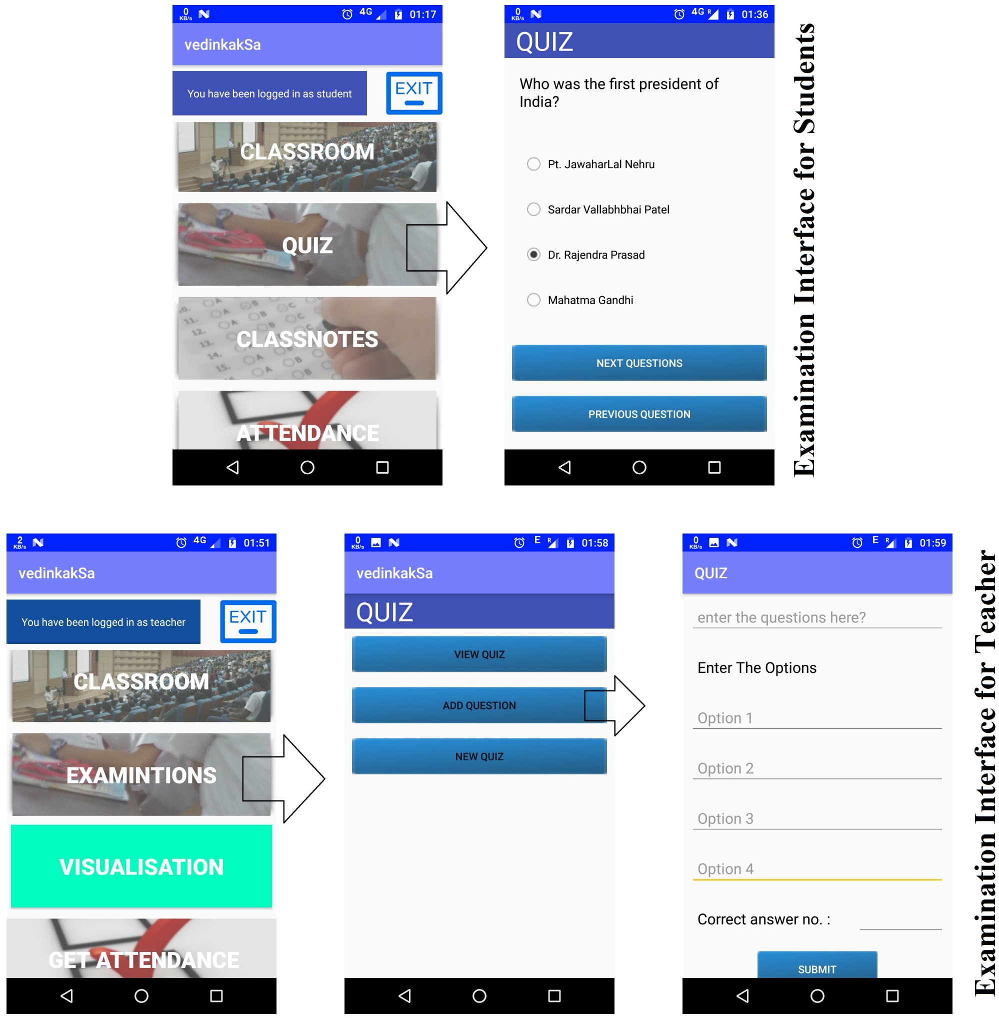


Figure 7.6: Screenshots of examination panel of the sensitive classroom system

'HW & assignment database'. Currently, the automatic evaluation of homework and assignment has not been implemented but the teacher can download the assignment and upload the scores after their manual evaluation. These scores are also sent to the grade processing submodule. The grade processing submodule determines the grades of the students based on a grade processing scheme considering the scores in quiz, exam, and assignment along with the percentage of attendance. In other words, with the help of 'exam & quiz processing' and 'HW & assignment processing' modules, the grade processing modules provide the grade information to the 'understanding and learning level detection' module of the system for computing the learning level of the students. The grades are also stored in 'grade database', and are sent to individual students as well as the teacher. Other than these tradi-

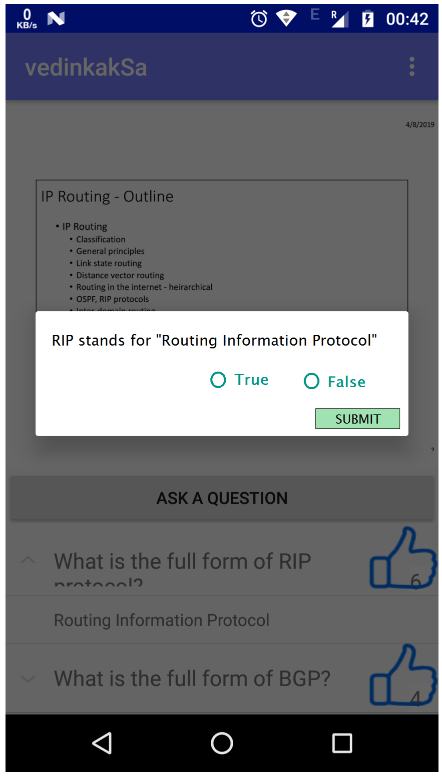


Figure 7.7: Example of a 'live-quiz'-question in the sensitive classroom system

tional quizzes, exams, and assignments, our system has a special type of quiz called 'live-quiz'. During the lecture, the teacher can open the 'live-quiz' panel and ask MCQ questions, which are appeared in the students' device as a popup dialog box as shown in Figure 7.7. The responses of these questions are also stored in the 'examination database'. The correctness of the responses to these questions defines the understanding level of the students, which are computed in the 'understanding and learning level detection' module of the system.

B. Implementing the Sensitivity of the System

The sensitivity of the system has been achieved through the implementation of two system modules, namely the *students' mental state detection*, and *understanding & learning level detection*. The details of the modules are presented as follows.

(a) Students' mental state detection – This module is the heart of the sensitive classroom system, as it detects or senses the mental states of the students. The mental state detection is done with the help of four submodules namely, *involve-*

Table 7.1: Defining mental states based on students' involvements, affective states, and activity levels

Sl. No.	Involvement	Activity Level	Affective State		Mental State	Meaning
			Arousal	Valence		
1.	Not Involved	×	×	×	S1	Not Engaged
2.	Involved	High	High	Positive	S2	Ideal State
3.	Involved	High	High	Negative	S3	Frustrated
4.	Involved	High	Low	Positive	S4	Good State
5.	Involved	High	Low	Negative	S5	Getting No Interest
6.	Involved	Low	High	Positive	S6	Showing Off
7.	Involved	Low	High	Negative	S7	Not Understanding
8.	Involved	Low	Low	Positive	S8	Understanding But Lazy
9.	Involved	Low	Low	Negative	S9	Feeling Shy

ment detection, *affective state detection*, *activity level detection*, and *state mapping*. We have utilized the novel models, which have been developed by us, for implementing the sensitivity of the system. The 'involvement detection' and 'affective state detection' submodules are the two machine learning-based computational models (the 'In-Activity' and 'Smart-Affect' models), which detect students' involvement in study-related activities and their affective states, respectively. The states are stored in the 'state database'. The 'activity level detection' submodule detects the level of classroom activities performed by a student within a certain period of time. It accesses the backchannel and activity databases to observe if any of the classroom activities (among query posting, query liking, class note-taking, responding to the 'random popups' & 'live-quizes', and acting on the alert signals) has been performed within a certain interval. In the current implementation, it considers five minutes as the interval. If a student performs at least one of the specified activities within the five-minute duration, it defines the activity level of the student as 'high'. Otherwise, the activity level of the students is said to be 'low'. Based on the identified activity levels along with the affective states and involvements of the students, the 'state mapping' submodules defines the mental states of the students in one of the nine states, S1 to S9, as per the specific state mapping scheme shown in Table 7.1, which has been described in details earlier in **Chapter 3**.

(b) Understanding & learning level detection – This module is composed of two submodules namely, the *understanding level detection*, and *learning level detection*. The term 'understanding', in the current context, refers to the short-term grasping of the lecture materials in the live classroom session. Whereas, the term 'learning' indicates the long-term learning outcomes throughout the course. The 'understanding level detection' submodule accesses two databases for getting inputs for defining the level of understanding. It accesses the examination database for the responses of 'live-quizzes' as they are supposed to directly indicate the understanding of the current lecture topic, and at the same time, the 'state' database for knowing the number and IDs of the students who are in particular mental states (S2, S4, S5, S8, and S9) as they indirectly indicate that the students belong to these states are understanding. The process specifies the level of understanding concerning the whole class so that the teacher can be informed about it when the maximum of the students is unable to understand the teaching content. In that case, s/he may tailor her/his teaching content and/or change the way of lecture delivery for an efficient teaching process. In the current implementation, the module specifies the understanding level of the students as 'low', if less than fifty percent of the students are in 'understanding' state. The module computes the level of understanding on a regular interval (five minutes), and update the information in the 'progress database'. The 'learning level identification' submodule accesses the 'grade database' to have the grades earned by the students, using which it computes the learning level of the students. For this computation, it also considers the attendance and understanding level history, which are accessed from the 'attendance database' and 'progress database', respectively. Once the learning level is computed, they are stored in the 'progress database', and may be accessed by the academic administrator(s).

C. Utilizing the Sensitive Information

The system might be underutilized if the identified sensitive information is not presented in an understandable manner, or the system does not take some smart actions based on the identified states for the improvement of the pedagogy. In order to utilize the sensitive information, we implemented two modules namely *classroom visualization* and *alert generation*. The details of these two modules are described as follows.

(a) Classroom visualization – We implemented a visualization scheme to present the identified states to the teacher so that s/he can easily understand them

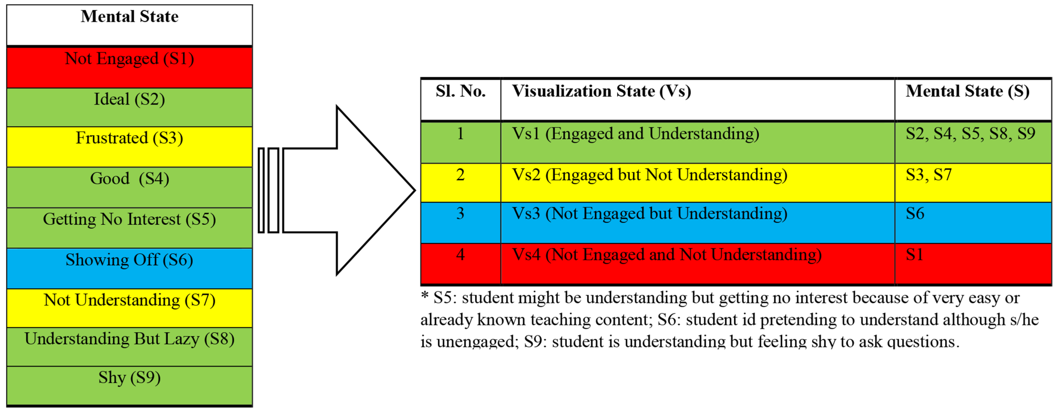


Figure 7.8: Mental-state to visualization-state mapping scheme.

and visualize the whole classroom with minimal effort. For this purpose, we further mapped these nine states into four super states – it would be difficult for a teacher to understand and remember the states if all the nine states are shown together. Moreover, few of the identified mental states can be grouped together as they belong to a similar category in the context of the learning process. For example, S2, S4, S5, S8, and S9 can be grouped together for classroom visualization because, in each of the cases, students are supposed to be engaged and understanding the teaching contents. Similarly, S3 and S7 can be put together in another group, as in these cases students are engaged but not understanding. The nine mental states (S1-S9) are mapped to four *visualization states* (Vs1-Vs4) as per the scheme depicted in Figure 7.8 based on students' engagement and understanding. The specific color codes for the visualization states are used to augment the dashboard for a visualized classroom. It is expected that the used color codes will help to increase usability in terms of learnability, memorability, aesthetics, and user satisfaction.

The classroom visualization module is composed of two submodules namely, the *visualization state mapping* and *layout sketching & augmentation*. The 'visualization state mapping' process executes the 'mental states' to 'visualization states' mapping as per the scheme in Figure 7.8 and sends the visualization state information to the 'layout sketching & augmentation' process. It also updates the state database with the visualization states for future reference. The 'layout sketching & augmentation' process accesses the user database to have students' credentials. With the user ID, this process requires two more information: profile pictures of the students which are

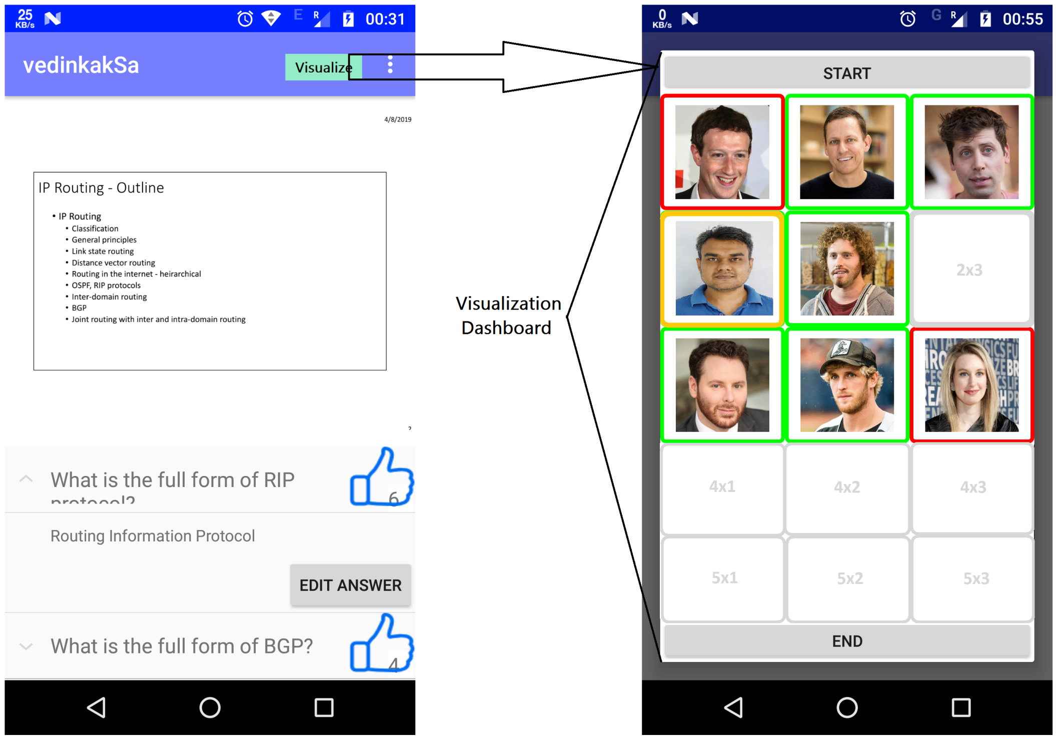


Figure 7.9: Screenshot of visualization dashboard.

uploaded during signing-up, and seating positions that are provided by the students during logging in into the system through scanning QR codes pasted in the classroom desks. QR codes have to be scanned by the students in each classroom session as the entries for them in the 'user database' get truncated when the teacher logs out, assuming that each student may not sit in a particular seat in the classroom for every lecture. After getting the seating position of the students and their profile pictures, the 'layout sketching & augmentation' process sketches a classroom layout on the visualization dashboard, and augments the states for individual students with the help of the user ID, once the visualization state information is received from the 'visualization state mapping' process. The pictures of the students, in the visualization dashboard, are augmented with the bounded box having four different colors. A green-colored bounded box indicates that the particular student is in Vs1 (engaged and understanding) state. Similarly, the yellow, blue, and red-colored bounded boxes indicate that the corresponding students are in Vs2 (Engaged but Not Understanding), Vs3 (Not Engaged but Understanding – Shoeing Off), and Vs4 (Not Engaged and Not Understanding) states, respectively. A representative screenshot for the classroom visualization dashboard is shown in Figure 7.9.

(b) Alert generation – This module is also composed of two submodules namely, the *alert generation for student*, and *alert generation for teacher*. The 'alert generation for student' submodule accesses the 'state database' to know the mental states of the students. If a student belongs to S1 ('Not Involved') or S6 (Showing Off) state, this process sends a popup message with device vibration to alert her/him. In case a student is found to be in S8 (understanding but lazy) or in S9 (Feeling Shy) state, this process sends a popup message having some motivational text so that s/he would get encouraged in participating in classroom activities, particularly asking queries and taking part in discussions. The 'user database' is accessed for identifying the students to whom the alerts are sent. All the alerts generated by the module are stored in the 'alert database'.

The 'alert generation for teacher' submodule has not been implemented in the current prototype of the system but has been thought in the following way. After accessing the 'state database' and 'progress database', if it has been found that more than a certain number of students are in S1 ('Not Engaged'), S3 (Frustrated), S5 ('Getting No Interest'), S6 (Showing Off), or in S7 (Not Understanding) state, it would generate an alert signal to a peripheral device belong to the teacher. A peripheral device has been considered for sending alert signals to the teacher with the expectation for a minimum interruption in lecture delivery. Along with the alert signals, it would also send the reasons for which the alerts have been sent so that the teacher can take action(s) accordingly. For instance, if the alerts are sent because of having the majority of the students in S1 or S6 state, the teacher can warn the students. Similarly, if the teacher knows that the alert has been sent because of having the majority of the students in S3, S5, or S5 state, s/he may tailor her/his teaching content and/or way of lecture delivery. Nevertheless, the 'majority' should be defined with a threshold value for the process to work. Currently, we have decided *fifty percent* as the threshold value. In other words, if more than fifty percent of the students are in S1, S3, S5, S6, or S7 state, the teacher should be alerted. However, the threshold value may be changed depending upon the opinions of the academic policymakers of the institute where the system will be deployed.

7.2 Summary of Thesis

Our primary focus for the thesis work was to propose a framework for a sensitive blended learning platform, which can automatically sense the mental states of the students in real-time. The mental states of the students, in terms of learning progress, are defined based on their affective states, involvements, and levels of classroom activities. Computational models are required for automatic detection of the affective states and involvements. As the existing computational models are seemed not to be readily adaptable in the current context, we developed four novel models namely, the 'Touch-Affect', 'Type-Affect', 'Smart-Affect', and 'In-Activity', which are suitable for a blended learning platform. The 'Touch-Affect' model is able to detect the affective state of the users from their touch pattern on the touchscreen of small handheld devices such as smartphones and tablets. The 'Type-Affect' model detects the affective states of the users from their typing pattern on the virtual keyboard of these devices. The 'Smart-Affect' is a process model, which combines the 'Touch-Affect' and 'Type-Affect' models for the detection of user's affective state from basic user-smartphone interaction behavior, i.e., touch and typing behavior in the blended learning platform. The 'In-Activity' model detects whether the students are involved in the study-related activity or not. The models have been built and validated with behavioral data of the students which have been collected through empirical studies. We addressed many research challenges and proposed novel ideas and methodology during the period of data collection as well as model building and validating. A behavioral study on smartphone usage has also been conducted by us for this purpose. Finally, we have proposed a framework that utilizes these models to detect the mental states of the students in a blended learning platform and take some smart actions based on the identified states in real-time to improve the teaching-learning process. To validate the proposed framework, we have built a working prototype of a sensitive classroom system following the proposed framework and observed that the system can detect the mental states of the students in either of a predefined set of states and take some live actions based on the identified states. The novelty of every contribution to cross each of the milestones to reach the final goal of the thesis is summarized as follows.

(a) **Touch-Affect model:** This is a minimalist model to detect the affective states of the user from their touch pattern on small handheld touchscreen devices. The machine learning-based model takes the number of touch events and average pressure created for these events for a specific time interval to detect the affective state of a user in one of the specific four states based on the two levels of arousal and valence. The model is capable of detecting the specific states with 96.65% accuracy. Our model is claimed to be novel as we could not find any work where the affective states of a touchscreen user has been identified using only two behavioral input features with such high accuracy. During the development period of the model, we also proposed a novel gaming approach for collecting affective data. Following this approach, it is possible to induce specific affective states without the knowledge of the users, and automatically label the data without any user-feedback, which consequently minimizes the imitation and error in the training dataset.

(b) **Type-Affect model:** This is another machine learning-based model to detect the affective states of the user from their typing behavior on the virtual keyboard of the touchscreen devices. We have built this minimalist model after reducing the unnecessary and application-specific features. Taking four basic typing features, the model can detect the affective states of the user with high accuracy (86.6%), and hence, suitable for the current context with many other applications. The model is found to be novel, and hence an original contribution to the thesis.

(c) **Smart-Affect model:** This model combines the Touch-Affect and Type-Affect models so that they can be utilized in all those applications, where both the general touch and typing through the virtual keyboard on the touchscreen are available as inputs but we are not sure when a particular type of input is available. It chooses one of the novel computational models based on the types of the available input data for the detection of the affective states. The model has been empirically validated for the target application using the EEG data of the students. We have found that none of the existing models can be readily adopted in the current context for detecting the specific affective states from basic user-smartphone interaction behavior. We, therefore, claim this model also novel.

(d) **In-Activity model:** This model also considers only the behavioral data along with the resource utilization pattern but no privacy-sensitive data for the detection of students' involvement in study-related activities. With the best of our knowledge, there is no such work in the literature where students' involvement has been detected in this way. Therefore, the model has been found to be novel. We have also conducted a study on the smartphone usage behavior of the students throughout India, as it was essential for the model development process but there was no up-to-date study in the literature. The study along with the novel model is therefore also original research contributions of this thesis.

(e) **Vedinkaksha framework:** Utilizing all these four novel models (Touch-Affect, Type-Affect, Smart-Affect, and In-Activity) and keeping the pedagogical and psychological aspects in the teaching-learning context in minds, we have developed a computational framework for a sensitive blended learning platform, called Vedinkaksha. A classroom system built as per the framework is supposed to be able to detect the mental states of the students along with their levels of understanding and learning and to take some smart actions based on the identified states. This was the primary goal of the thesis, and hence, the main contribution of the same.

(f) **Sensitive classroom system:** We have designed a sensitive classroom system following the Vedinkaksha framework, and implemented high-fidelity a working prototype for the same. In its present state, the prototype has been observed to be capable of identifying the mental states and at the same time, taking some smart actions such as alerting and/or motivating unengaged and shy students and sending a visualized classroom to the teacher based on the identified states in real-time. This is one more contribution to the thesis. While proposing the framework for a sensitive classroom system as well as implementing the same, we have chosen a particular blended learning setting. It can be noted that the choice was not random. Through an empirical study, we have proven the efficiency of the particular platform in terms of increasing classroom interactions and learning outcomes. The study result also indicated that the students were inclined to accept the specific classroom setting for their regular classes. This study can be considered as an additional contribution of the thesis which also extends the existing work.

While developing the Vedinkaksha framework, in addition to addressing the research issues associated with computation, we have also studied and discussed with many other researchers of multidisciplinary areas related to the thesis work. The areas include education, learning technology, psychology, behavioral science, and cognitive science. For instance, at the time of defining the mental states of the students by considering their affective states, involvements, and level of classroom activities (Table 7.1), we had discussions with educationalists and educational psychologists on it. These discussions and studies have helped us also in solving a pedagogical issue in the current context as follows. At the time of developing the query management module for the Vedinkaksha, we faced a challenge in displaying the queries in limited space on the smartphone screen – it would be difficult to access the queries if all of them are displayed together in the query panel. This is because in a lecture session, a huge number of queries may be generated, and it may be cumbersome for the teacher to find out the relevant query among them at a particular moment. This is instead of even the queries are arranged in a sorted order based on their priority. We solved this issue by displaying only those queries which are related to the current slide being displayed in the lecture delivery panel. In this case, also, the queries are displayed in sorted order based on their priorities. Although there is an option for both the teacher and students to access all the queries if they are interested. In this way, the discussions with the multidisciplinary researchers, and the studies of such research works have smoothened the way of achieving the goal through the research contributions.

Along with the contributions and achievements we have identified a few limitations, further investigation of which may strengthen our work, and at the same time, may open up new directions for future research. We discuss the identified shortcomings and future research scope as follows.

7.3 Future Research Scopes

We have developed several computational models and utilized them for developing a computational framework for a sensitive classroom system. During developing the framework, we have assumed a particular classroom setting called blended learning platform as the background. This is one of the ways but may not be the only way to develop such a framework. One can utilize the models in the synchronous distance learning platform, where the teacher and students interact through smartphone/tablet for the classroom activities. That may be another way of developing a sensitive learning platform, keeping the basic idea intact. In that case, although the students may miss the advantages of f2f classroom, at least the sensitivity of the system can be achieved with the benefit of e-learning. Nonetheless, some research issues may arise in that case, which can be identified and addressed by future researchers.

We conducted all the empirical studies, for developing the framework as well as the computational models associated with it, on UG and PG students. At the same time, through a survey, we have found that the behavior of the students may change based on their academic levels. The models, and hence the system developed by utilizing them following the proposed framework, may therefore be best suitable for the students of UG and PG level. If we wish to use the system for the students of a different academic level, the models may be retrained with the behavioral data of the students who belong to that particular level. Although, in that case, the overall framework should remain intact.

In order to propose the Vedinkaksha framework, we mentioned that the client devices may either be smartphones or tablets, as we think the models would work for both these devices because they have similar characteristics in terms of hardware and sensors embedded with them. However, we have conducted all the empirical researches with only smartphone users. This is because smartphones are the most common devices used by the students nowadays – the popularity of tablets has been reduced in recent times. Nevertheless, it may further be investigated if changes in devices may affect accuracy. This is because changes in size and weight may have an effect on the device handling pattern, as well as touch and typing patterns, which are the inputs of the computational models. For example, a user may use her/his ten fingers for typing on the virtual keyboard of a tablet (as many tablets come with larger displays), which is very odd to assume in case of typing on the virtual keyboard of a smartphone. These may lead to differences in behavioral data patterns depending upon the types of interactive devices. Therefore, in that case, a refinement of the associated model may be required through retraining the model with up-to-date behavioral data, keeping the other aspects intact. Moreover, in the future, today's smartphones may even change in terms of size, weight, hardware, and software features. In that case, also, the retraining of the models may be required. Making the models adaptive may be another way to tackle this.

The affective state detection model detects the affective states of the students after every seven-second interval, the involvement detection model also can detect the involvement of the students considering a time window of seven seconds (although consideration of a twenty-one-second window may enhance the performance of the model). However, in the current proposal, the activity level is defined considering a five-minute time-window, assuming that an active student should perform at least one of the specified classroom activities within the five-minute duration. As the mental states are defined based on students' affective states, involvements, and activity-levels, we have considered a five-minute time-window to detect the mental state, as within these five minutes all its components can provide inputs. Alternatively, we could even consider a seven-second window to detect the mental states considering the last identified affective states, involvement, and activity levels. Nonetheless, deciding an appropriate time slice for defining these states may lead to a new research problem, a solution of which may help to strengthen our work.

Currently, we have discussed a simple classroom visualization technique and implemented the same in the working prototype. However, new and more efficient visualization techniques e.g., multilevel visualization, or visualization through live camera may be thought, which itself may open up new research problems.

In the current proposal, we have assumed the main devices for sending alert signals to students, whereas, a peripheral device for sending an alert signal to the teacher. We have also specified and discussed the criteria for generating and sending alert signals. However, it is required to do further research on identifying proper scheme on time, type, duration, frequency, and way of sending the alerts, particularly for peripheral one, so that it causes a minimal interruption in lecture delivery and other essential classroom activities.

After developing the first prototype of a sensitive blended learning platform, we have primarily tested the sensitivity of the system in terms of model's output, detection of mental states as per the predefined mapping for the same according to affect, involvement and activity. We have also tested the visualization scheme followed by us. However, the sensitivity of the system should rigorously be tested with the feedback from the students and teacher once the system will be practically deployed for teaching and learning activities. Furthermore, the usability of the whole system (in terms of learnability, memorability, effectiveness, efficiency, and user satisfaction) should also be tested empirically once it will be deployed for academic purposes in real-time. These may also lead to open new research avenues for multidisciplinary researchers in the future.

We have built several novel computational models for the development of Vedinkaksha. The models have been developed keeping the classroom context in mind and found to be suitable for the same. However, there may be scopes for applying some of the models, particularly the affect detection models, in several other applications as well. This is because the models use the basic user-smartphone interaction data as inputs for detecting the states, which are independent of application areas. However, there could be a requirement of minor refinement of the models before applying them in those applications, which also may lead to open up new research directions.

Bibliography

[1] N. Aggarwal. A study on decline trend in the market share of apple iphone.

[2] A. A. Al-Qahtani and S. E. Higgins. Effects of traditional, blended and e-learning on students' achievement in higher education. *Journal of computer assisted learning*, 29(3):220–234, 2013.

[3] O. AlZoubi, S. K. D'Mello, and R. A. Calvo. Detecting naturalistic expressions of nonbasic affect using physiological signals. *IEEE Transactions on Affective Computing*, 3(3):298–310, 2012.

[4] C. Anagnostopoulos and S. Hadjiefthymiades. Intelligent trajectory classification for improved movement prediction. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(10):1301–1314, 2014.

[5] J. R. Anderson, C. F. Boyle, and B. J. Reiser. Intelligent tutoring systems. *Science*, 228(4698):456–462, 1985.

[6] J. J. Appleton, S. L. Christenson, D. Kim, and A. L. Reschly. Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of school psychology*, 44(5):427–445, 2006.

[7] L. Archambault and K. Crippen. Examining tpack among k-12 online distance educators in the united states. *Contemporary issues in technology and teacher education*, 9(1):71–88, 2009.

[8] I. Arroyo, K. Ferguson, J. Johns, T. Dragon, H. Meheranian, D. Fisher, A. Barto, S. Mahadevan, and B. P. Woolf. Repairing disengagement with non-invasive interventions. In *AIED*, volume 2007, pages 195–202, 2007.

[9] T. Ashwin and R. M. R. Guddeti. Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks. *Education and Information Technologies*, 25(2):1387–1415, 2020.

[10] R. Ballagas, M. Rohs, J. G. Sheridan, and J. Borchers. Byod: Bring your own device. In *Proceedings of the Workshop on Ubiquitous Display Environments, Ubicomp*, volume 2004, 2004.

[11] M. Barak, S. Waks, and Y. Doppelt. Majoring in technology studies at high school and fostering learning. *Learning Environments Research*, 3(2):135–158, 2000.

[12] L. Barkhuus. Why everyone loves to text message: social management with sms. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*, pages 324–325, 2005.

[13] L. Bartram, A. Patra, and M. Stone. Affective color in visualization. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 1364–1374, 2017.

[14] S. Battin-Pearson, M. D. Newcomb, R. D. Abbott, K. G. Hill, R. F. Catalano, and J. D. Hawkins. Predictors of early high school dropout: A test of five theories. *Journal of educational psychology*, 92(3):568, 2000.

[15] G. Bauer and P. Lukowicz. Can smartphones detect stress-related changes in the behaviour of individuals? In *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, pages 423–426. IEEE, 2012.

[16] C. Beedie, P. Terry, and A. Lane. Distinctions between emotion and mood. *Cognition & Emotion*, 19(6):847–878, 2005.

[17] C. Boonroungrut and T. T. Oo. Exploring classroom emotion with cloud-based facial recognizer in the chinese beginning class: A preliminary study. *International Journal of Instruction*, 12(1):947–958, 2019.

[18] N. Bosch and S. D'Mello. Automatic detection of mind wandering from video in the lab and in the classroom. *IEEE Transactions on Affective Computing*, 2019.

[19] A. F. Botelho, R. S. Baker, and N. T. Heffernan. Improving sensor-free affect detection using deep learning. In *International Conference on Artificial Intelligence in Education*, pages 40–51. Springer, 2017.

[20] K. Bowen and M. D. Pistilli. Student preferences for mobile app usage. *Research Bulletin)(Louisville, CO: EDUCAUSE Center for Applied Research, forthcoming), available from http://www. educause. edu/ecar*, 2012.

[21] P. E. Cairns and A. L. Cox. *Research methods for human-computer interaction.* Cambridge University Press, 2008.

[22] E. Chapman. Alternative approaches to assessing student engagement rates. *Practical assessment, research & evaluation*, 8(13):1–10, 2003.

[23] C.-M. Chen, S.-H. Hsu, Y.-L. Li, and C.-J. Peng. Personalized intelligent m-learning system for supporting effective english learning. In *2006 IEEE International Conference on Systems, Man and Cybernetics*, volume 6, pages 4898–4903. IEEE, 2006.

[24] N. Choudhury, V. Tamarapalli, and S. Bhattacharya. An ict-based system to improve the learning experience in a large classroom. In *2015 IEEE Seventh International Conference on Technology for Education (T4E)*, pages 27–30. IEEE, 2015.

[25] M. Ciman, K. Wac, and O. Gaggi. isensestress: Assessing stress through human-smartphone interaction analysis. In *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pages 84–91. IEEE, 2015.

[26] G. L. Collier. Beyond valence and activity in the emotional connotations of music. *Psychology of Music*, 35(1):110–131, 2007.

[27] C. Conati. Probabilistic assessment of user's emotions in educational games. *Applied artificial intelligence*, 16(7-8):555–575, 2002.

[28] A. T. Corbett, K. R. Koedinger, and J. R. Anderson. Intelligent tutoring systems. In *Handbook of human-computer interaction*, pages 849–874. Elsevier, 1997.

[29] R. S. d Baker, S. M. Gowda, M. Wixon, J. Kalka, A. Z. Wagner, A. Salvi, V. Aleven, G. W. Kusbit, J. Ocumpaugh, and L. Rossi. Towards sensor-free affect detection in cognitive tutor algebra. *International Educational Data Mining Society*, 2012.

[30] S. Debener, F. Minow, R. Emkes, K. Gandras, and M. De Vos. How about taking a low-cost, small, and wireless eeg for a walk? *Psychophysiology*, 49(11):1617–1621, 2012.

[31] E. Douglas-Cowie, R. Cowie, I. Sneddon, C. Cox, O. Lowry, M. Mcrorie, J.-C. Martin, L. Devillers, S. Abrilian, A. Batliner, N. Amir, and K. Karpouzis. The humaine database: Addressing the collection and annotation of naturalistic and induced emotional data. In *International conference on affective computing and intelligent interaction*, pages 488–500. Springer, 2007.

[32] T. Dragon, I. Arroyo, B. P. Woolf, W. Burleson, R. El Kaliouby, and H. Eydgahi. Viewing student affect and learning through classroom observation and physical sensors. In *International Conference on Intelligent Tutoring Systems*, pages 29–39. Springer, 2008.

[33] C. Epp, M. Lippold, and R. L. Mandryk. Identifying emotional states using keystroke dynamics. In *Proceedings of the sigchi conference on human factors in computing systems*, pages 715–724, 2011.

[34] M. Eslahi, M. V. Naseri, H. Hashim, N. Tahir, and E. H. M. Saad. Byod: Current state and security challenges. In *2014 IEEE Symposium on Computer Applications and Industrial Electronics (ISCAIE)*, pages 189–192. IEEE, 2014.

[35] J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris. School engagement: Potential of the concept, state of the evidence. *Review of educational research*, 74(1):59–109, 2004.

[36] J. Fu, T. Ge, M. Li, and X. Hu. Affective computation of students' behaviors under classroom scenes. In *NeuroManagement and Intelligent Computing Method on Multimodal Interaction*, pages 1–6. 2019.

[37] A. Gabrielsson. Emotion perceived and emotion felt: Same or different? *Musicae scientiae*, 5(1_suppl):123–147, 2001.

[38] Y. Gaffary, D. A. G. Jáuregui, J.-C. Martin, and M. Ammi. Gestural and postural reactions to stressful event: design of a haptic stressful stimulus. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 988–992. IEEE, 2015.

[39] Y. Gao, N. Bianchi-Berthouze, and H. Meng. What does touch tell us about emotions in touchscreen-based gameplay? *ACM Transactions on Computer-Human Interaction (TOCHI)*, 19(4):1–30, 2012.

[40] T. Garcia and P. R. Pintrich. Assessing students' motivation and learning strategies in the classroom context: The motivated strategies for learning questionnaire. In *Alternatives in assessment of achievements, learning processes and prior knowledge*, pages 319–339. Springer, 1996.

[41] A. S. Gertner and K. VanLehn. Andes: A coached problem solving environment for physics. In *International conference on intelligent tutoring systems*, pages 133–142. Springer, 2000.

[42] K. M. Gilleade and A. Dix. Using frustration in the design of adaptive videogames. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pages 228–232, 2004.

[43] N. Gitinabard, Y. Xu, S. Heckman, T. Barnes, and C. F. Lynch. How widely can prediction models be generalized? performance prediction in blended courses. *IEEE Transactions on Learning Technologies*, 12(2):184–197, 2019.

[44] D. Glowinski, N. Dael, A. Camurri, G. Volpe, M. Mortillaro, and K. Scherer. Toward a minimal representation of affective gestures. *IEEE Transactions on Affective Computing*, 2(2):106–118, 2011.

[45] J. D. Gobert, M. A. Sao Pedro, R. S. Baker, E. Toto, and O. Montalvo. Leveraging educational data mining for real-time performance assessment of scientific inquiry skills within microworlds. *Journal of Educational Data Mining*, 4(1):104–143, 2012.

[46] D. González-Gómez, J. S. Jeong, D. A. Rodríguez, and F. Canada-Canada. Performance and perception in the flipped learning model: an initial approach to evaluate the effectiveness of a new teaching methodology in a general science classroom. *Journal of Science Education and Technology*, 25(3):450–459, 2016.

[47] S. Govaerts, A. Holzer, B. Kocher, A. Vozniuk, B. Garbinato, and D. Gillet. Blending digital and face-to-face interaction using a co-located social media app in class. *IEEE Transactions on Learning Technologies*, 11(4):478–492, 2018.

[48] A. C. Graesser, K. VanLehn, C. P. Rosé, P. W. Jordan, and D. Harter. Intelligent tutoring systems with conversational dialogue. *AI magazine*, 22(4):39–39, 2001.

[49] A. C. Graesser, K. Wiemer-Hastings, P. Wiemer-Hastings, R. Kreuz, T. R. Group, et al. Autotutor: A simulation of a human tutor. *Cognitive Systems Research*, 1(1):35–51, 1999.

[50] B. Grawemeyer, M. Mavrikis, W. Holmes, S. Gutiérrez-Santos, M. Wiedmann, and N. Rummel. Affective learning: improving engagement and enhancing learning with affect-aware feedback. *User Modeling and User-Adapted Interaction*, 27(1):119–158, 2017.

[51] J. J. Gross. Emotion and emotion regulation. *Handbook of personality: Theory and research*, 2:525–552, 1999.

[52] N. Y. Hammerla and T. Plötz. Let's (not) stick together: pairwise similarity biases cross-validation in activity recognition. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, pages 1041–1051, 2015.

[53] S. Han and Y. J. Yi. How does the smartphone usage of college students affect academic performance? *Journal of Computer Assisted Learning*, 35(1):13–22, 2019.

[54] N. Harada, M. Kimura, T. Yamamoto, and Y. Miyake. System for measuring teacher–student communication in the classroom using smartphone accelerometer sensors. In *International Conference on Human-Computer Interaction*, pages 309–318. Springer, 2017.

[55] R. L. Hazlett. Measuring emotional valence during interactive experiences: boys at video game play. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1023–1026, 2006.

[56] J. A. Healey and R. W. Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on intelligent transportation systems*, 6(2):156–166, 2005.

[57] N. T. Heffernan and C. L. Heffernan. The assistments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, 24(4):470–497, 2014.

[58] M. J. Hertenstein, R. Holmes, M. McCullough, and D. Keltner. The communication of emotion via touch. *Emotion*, 9(4):566, 2009.

[59] N. Hoic-Bozic, V. Mornar, and I. Boticki. A blended learning approach to course design and implementation. *IEEE transactions on education*, 52(1):19–30, 2008.

[60] E. Y. Huang, S. W. Lin, and T. K. Huang. What type of learning style leads to online participation in the mixed-mode e-learning environment? a study of software usage instruction. *Computers & Education*, 58(1):338–349, 2012.

[61] R. Jena. The impact and penetration of smartphone usage in student's life. *Global Journal of Business Management*, 8(1):29–35, 2014.

[62] S. R. Jimerson, E. Campos, and J. L. Greif. Toward an understanding of definitions and measures of school engagement and related terms. *The California School Psychologist*, 8(1):7–27, 2003.

[63] Z. Jin, Y. Sun, and A. C. Cheng. Predicting cardiovascular disease from real-time electrocardiographic monitoring: An adaptive machine learning approach on a cell phone. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6889–6892. IEEE, 2009.

[64] N. Kambouropoulos and P. K. Staiger. Personality and responses to appetitive and aversive stimuli: the joint influence of behavioural approach and behavioural inhibition systems. *Personality and Individual Differences*, 37(6):1153–1165, 2004.

[65] A. Kapoor and R. W. Picard. Multimodal affect recognition in learning environments. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 677–682, 2005.

[66] P. Khanna and M. Sasikumar. Recognising emotions from keyboard stroke pattern. *International journal of computer applications*, 11(9):1–5, 2010.

[67] Y. Kim, T. Soyata, and R. F. Behnagh. Towards emotionally aware ai smart classroom: Current issues and directions for engineering and education. *IEEE Access*, 6:5308–5331, 2018.

[68] V. Kolodyazhniy, S. D. Kreibig, J. J. Gross, W. T. Roth, and F. H. Wilhelm. An affective computing approach to physiological emotion specificity: Toward subject-independent and stimulus-independent classification of film-induced emotions. *Psychophysiology*, 48(7):908–922, 2011.

[69] S. G. Koolagudi and K. S. Rao. Exploring speech features for classifying emotions along valence dimension. In *International Conference on Pattern Recognition and Machine Intelligence*, pages 537–542. Springer, 2009.

[70] A. Kukulska-Hulme, J. Pettit, L. Bradley, A. A. Carvalho, A. Herrington, D. M. Kennedy, and A. Walker. Mature students using mobile devices in life

and learning. *International Journal of Mobile and Blended Learning (IJMBL)*, 3(1):18–52, 2011.

[71] M. Kwet and P. Prinsloo. The 'smart'classroom: a new frontier in the age of the smart university. *Teaching in Higher Education*, 25(4):510–526, 2020.

[72] P. J. Lang, M. K. Greenwald, M. M. Bradley, and A. O. Hamm. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3):261–273, 1993.

[73] R. W. Larson and M. H. Richards. Boredom in the middle school years: Blaming schools versus blaming students. *American journal of education*, 99(4):418–443, 1991.

[74] J. Lazar, J. H. Feng, and H. Hochheiser. *Research methods in human-computer interaction*. Morgan Kaufmann, 2017.

[75] C. Lee and S.-J. Lee. Prevalence and predictors of smartphone addiction proneness among korean adolescents. *Children and Youth Services Review*, 2017.

[76] H. Lee, Y. S. Choi, S. Lee, and I. Park. Towards unobtrusive emotion recognition for affective social communication. In *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, pages 260–264. IEEE, 2012.

[77] P.-M. Lee, W.-H. Tsui, and T.-C. Hsiao. The influence of emotion on keyboard typing: an experimental study using visual stimuli. *Biomedical engineering online*, 13(1):1–12, 2014.

[78] U. Lee, J. Lee, M. Ko, C. Lee, Y. Kim, S. Yang, K. Yatani, G. Gweon, K.-M. Chung, and J. Song. Hooked on smartphones: an exploratory study on smartphone overuse among college students. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 2327–2336, 2014.

[79] A. Lesgold, S. Lajoie, M. Bunzo, and G. Eggan. Sherlock: A coached practice environment for an electronics troubleshooting. *Computer assisted instruction and intelligent tutoring systems: Shared goals and complementary approaches*, pages 201–232, 1992.

[80] J. Li, D. Shi, P. Tumnark, and H. Xu. A system for real-time intervention in negative emotional contagion in a smart classroom deployed under edge computing service infrastructure. *Peer-to-Peer Networking and Applications*, pages 1–14, 2020.

[81] Y. M. Lim, A. Ayesh, and M. Stacey. The effects of typing demand on emotional stress, mouse and keystroke behaviours. In *Science and Information Conference*, pages 209–225. Springer, 2014.

[82] Y. M. Lim, A. Ayesh, and M. Stacey. Continuous stress monitoring under varied demands using unobtrusive devices. *International Journal of Human–Computer Interaction*, 36(4):326–340, 2020.

[83] G. Loewenstein and J. S. Lerner. The role of affect in decision making. *Handbook of affective science*, 619(642):3, 2003.

[84] H.-R. Lv, Z.-L. Lin, W.-J. Yin, and J. Dong. Emotion recognition based on pressure sensor keyboards. In *2008 IEEE International Conference on Multimedia and Expo*, pages 1089–1092. IEEE, 2008.

[85] Y. Matsuda, I. Sakuma, Y. Jimbo, E. Kobayashi, T. Arafune, and T. Isomura. Emotional communication in finger braille. *Advances in Human-Computer Interaction*, 2010, 2010.

[86] G. Matthews, S. E. Campbell, S. Falconer, L. A. Joyner, J. Huggins, K. Gilliland, R. Grier, and J. S. Warm. Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion*, 2(4):315, 2002.

[87] G. Matthews, L. Joyner, K. Gilliland, S. Campbell, S. Falconer, and J. Huggins. Validation of a comprehensive stress state questionnaire: Towards a state big three. *Personality psychology in Europe*, 7:335–350, 1999.

[88] M. M. McCaslin and T. L. Good. *Listening in classrooms.* HarperCollins, 1996.

[89] B. R. McCoy. Digital distractions in the classroom phase ii: Student classroom use of digital devices for non-class related purposes. pages 1–44, 2016.

[90] M. K. Miller and R. L. Mandryk. Differentiating in-game frustration from at-game frustration using touch pressure. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*, pages 225–234, 2016.

[91] H. Müller, J. L. Gove, J. S. Webb, and A. Cheang. Understanding and comparing smartphone and tablet use: Insights from a large-scale diary study. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, pages 427–436, 2015.

[92] I. R. Murray and J. L. Arnott. Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America*, 93(2):1097–1108, 1993.

[93] A. N. H. Nahin, J. M. Alam, H. Mahmud, and K. Hasan. Identifying emotion by keystroke dynamics and text pattern analysis. *Behaviour & Information Technology*, 33(9):987–996, 2014.

[94] J. K. Nayak. Relationship among smartphone usage, addiction, academic performance and the moderating role of gender: A study of higher education students in india. *Computers & Education*, 123:164–173, 2018.

[95] P. Nicholson. A history of e-learning. In *Computers and education*, pages 1–11. Springer, 2007.

[96] A.-M. Nortvig, A. K. Petersen, and S. H. Balle. A literature review of the factors influencing e-learning and blended learning in relation to learning outcome, student satisfaction and engagement. *Electronic Journal of e-Learning*, 16(1):46–55, 2018.

[97] H. S. Nwana. Intelligent tutoring systems: an overview. *Artificial Intelligence Review*, 4(4):251–277, 1990.

[98] R. T. Osguthorpe and C. R. Graham. Blended learning environments: Definitions and directions. *Quarterly review of distance education*, 4(3):227–33, 2003.

[99] L. Paquette, R. S. Baker, M. A. Sao Pedro, J. D. Gobert, L. Rossi, A. Nakama, and Z. Kauffman-Rogoff. Sensor-free affect detection for a simulation-based science inquiry learning environment. In *International conference on intelligent tutoring systems*, pages 1–10. Springer, 2014.

[100] N. Park and H. Lee. Gender difference in social networking on smartphones: A case study of korean college student smartphone users. *International Telecommunications Policy Review*, 21(2):1–18, 2014.

[101] N. Park and H. Lee. Nature of youth smartphone addiction in korea. *Journal of Communication Research*, 2014.

[102] J. Parsons and L. Taylor. *Student Engagement: What do we know and what should we do?* University of Alberta, 2012.

[103] P. Philippot. Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition and emotion*, 7(2):171–193, 1993.

[104] J. Polivy. On the induction of emotion in the laboratory: Discrete moods or multiple affect states? *Journal of Personality and Social Psychology*, 41(4):803, 1981.

[105] K. Porayska-Pomsta, M. Mavrikis, S. D'Mello, C. Conati, and R. S. Baker. Knowledge elicitation methods for affect modelling in education. *International Journal of Artificial Intelligence in Education*, 22(3):107–140, 2013.

[106] J. Posner, J. A. Russell, and B. S. Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(3):715, 2005.

[107] K. S. Rao, T. P. Kumar, K. Anusha, B. Leela, I. Bhavana, and S. Gowtham. Emotion recognition from speech. *International Journal of Computer Science and Information Technologies*, 3(2):3603–3607, 2012.

[108] A. L. Reschly and S. L. Christenson. Prediction of dropout among students with mild disabilities: A case for the inclusion of student engagement variables. *Remedial and Special Education*, 27(5):276–292, 2006.

[109] S. Ritter, J. R. Anderson, K. R. Koedinger, and A. Corbett. Cognitive tutor: Applied research in mathematics education. *Psychonomic bulletin & review*, 14(2):249–255, 2007.

[110] N. Roberts and M. Rees. Student use of mobile devices in university lectures. *Australasian Journal of Educational Technology*, 30(4), 2014.

[111] V. Sacharin, K. Schlegel, and K. R. Scherer. Geneva emotion wheel rating study. *Available at http://archive-ouverte.unige.ch/unige:97849*, 2012.

[112] A. Sano and R. W. Picard. Stress recognition using wearable sensors and mobile phones. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 671–676. IEEE, 2013.

[113] C. Savvaki, A. Leonidis, G. Paparoulis, M. Antona, and C. Stephanidis. Designing a technology–augmented school desk for the future classroom. In *International Conference on Human-Computer Interaction*, pages 681–685. Springer, 2013.

[114] B. L. Schroeder, D. E. Whitmer, S. K. Bailey, and V. K. Sims. Individual differences in middle school and college students' texting. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 60, pages 1215–1219. SAGE Publications Sage CA: Los Angeles, CA, 2016.

[115] S. Senthil and W. M. Lin. Measuring students' engagement using wireless heart rate sensors. In *2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)*, pages 699–704. IEEE, 2017.

[116] S. Shah, J. N. Teja, and S. Bhattacharya. Towards affective touch interaction: predicting mobile user emotion from finger strokes. *Journal of Interaction Science*, 3(1):6, 2015.

[117] S. Shahyad, S. Pakdaman, M. Hiedary, M. Miri, M. Asadi, A. Nasri, and A. S. Alipour. A comparison of motivation, frequency and content of sms messages sent in boys and girls high school student. *Procedia-Social and Behavioral Sciences*, 15:895–898, 2011.

[118] D. J. Shernoff, M. Csikszentmihalyi, B. Schneider, and E. S. Shernoff. Student engagement in high school classrooms from the perspective of flow theory. In *Applications of flow in human development and education*, pages 475–494. Springer, 2014.

[119] E. A. Skinner and M. J. Belmont. Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of educational psychology*, 85(4):571, 1993.

[120] E. A. Skinner, J. G. Wellborn, and J. P. Connell. What it takes to do well in school and whether i've got it: A process model of perceived control and children's engagement and achievement in school. *Journal of educational psychology*, 82(1):22, 1990.

[121] M. V. Sokolova and A. Fernández-Caballero. A review on the role of color and light in affective computing. *Applied Sciences*, 5(3):275–293, 2015.

[122] A. Stein, Y. Yotam, R. Puzis, G. Shani, and M. Taieb-Maimon. Eeg-triggered dynamic difficulty adjustment for multiplayer games. *Entertainment computing*, 25:14–25, 2018.

[123] A. P. Sweet, J. T. Guthrie, and M. N. Ng. Teachers' perceptions and students' literacy motivations. reading research report no. 69. *National Reading Research Center*, 1996.

[124] T. M. Swindle, W. L. Ward, L. Whiteside-Mansell, P. Bokony, and D. Pettit. Technology use and interest among low-income parents of young children: differences by age group and ethnicity. *Journal of Nutrition Education and Behavior*, 46(6):484–490, 2014.

[125] S. Tikadar and S. Bhattacharya. How do they use their smartphones: A study on smartphone usage by indian students. In *IFIP Conference on Human-Computer Interaction*, pages 132–151. Springer, 2019.

[126] S. Tikadar and S. Bhattacharya. A novel method to build and validate an affective state prediction model from touch-typing. In *IFIP Conference on Human-Computer Interaction*, pages 99–119. Springer, 2019.

[127] S. Tikadar and S. Bhattacharya. Predicting students' involvement in blended learning environment. In *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, volume 2161, pages 76–78. IEEE, 2019.

[128] S. Tikadar and S. Bhattacharya. A computational model for detecting involvement of the students in a blended learning platform. *IEEE Transactions on Learning Technologies*, X(X):[under review], 2020.

[129] S. Tikadar and S. Bhattacharya. Detection of affective states of the students in a blended learning environment comprising of smartphones. *International Journal of Human–Computer Interaction*, pages 1–18, 2020.

[130] S. Tikadar, S. Bhattacharya, and V. Tamarapalli. A blended learning platform to improve teaching-learning experience. In *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*, pages 87–89. IEEE, 2018.

[131] S. Tikadar, S. Kazipeta, C. Ganji, and S. Bhattacharya. A minimalist approach for identifying affective states for mobile interaction design. In *IFIP Conference on Human-Computer Interaction*, pages 3–12. Springer, 2017.

[132] G. Urh and V. Pejović. Taskyapp: inferring task engagement via smartphone sensing. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 1548–1553, 2016.

[133] K. VanLehn, R. Freedman, P. Jordan, C. Murray, R. Osan, M. Ringenberg, C. Rosé, K. Schulze, R. Shelby, D. Treacy, A. Weinstein, and M. Wintersgill. Fading and deepening: The next steps for andes and other model-tracing tutors. In *International Conference on Intelligent Tutoring Systems*, pages 474–483. Springer, 2000.

[134] D. Västfjäll. Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae*, 5(1_suppl):173–211, 2001.

[135] R. Vicencio-Moreira, R. L. Mandryk, and C. Gutwin. Now you can compete with anyone: Balancing players of different skill levels in a first-person shooter

game. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2255–2264, 2015.

[136] L. M. Vizer, L. Zhou, and A. Sears. Automated stress detection using keystroke and linguistic features: An exploratory study. *International Journal of Human-Computer Studies*, 67(10):870–886, 2009.

[137] I. D. Wahyono, D. Saryono, M. Ashar, K. Asfani, and Sunarti. Face emotional detection using computational intelligence based ubiquitous computing. In *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*, pages 389–393. IEEE, 2019.

[138] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J. R. Movellan. The faces of engagement: Automatic recognition of student engagementfrom facial expressions. *IEEE Transactions on Affective Computing*, 5(1):86–98, 2014.

[139] A. Wigfield, J. T. Guthrie, K. C. Perencevich, A. Taboada, S. L. Klauda, A. McRae, and P. Barbosa. Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes. *Psychology in the Schools*, 45(5):432–445, 2008.

[140] B. Woolf, W. Burelson, and I. Arroyo. Emotional intelligence for computer tutors. In *Workshop on modeling and scaffolding affective experiences to impact learning at 13th international conference on artificial intelligence in education, Los Angeles, California*, 2007.

[141] B. Woolf, W. Burleson, I. Arroyo, T. Dragon, D. Cooper, and R. Picard. Affect-aware tutors: recognising and responding to student affect. *International Journal of Learning Technology*, 4(3-4):129–164, 2009.

[142] G. N. Yannakakis and A. Pavia. Emotion in games: Handbook on affective computing, 2014.

[143] M. N. A. Zabidi, M. A. Maarof, and A. Zainal. Ensemble based categorization and adaptive model for malware detection. In *2011 7th International Conference on Information Assurance and Security (IAS)*, pages 80–85. IEEE, 2011.

[144] P. Zimmermann, S. Guttormsen, B. Danuser, and P. Gomez. Affective computing—a rationale for measuring mood with mouse and keyboard. *International journal of occupational safety and ergonomics*, 9(4):539–551, 2003.

Publications

Patent out of this thesis

- **Application No:** 201831024866 (Indian Patent Office). **Title:** "A Method and a System for Controlling Learning Progress based on Identification of States of Learning Progress." **Applicant:** IIT Guwahati. **Date of application:** 04/07/2019. **Inventors:** Dr Samit Bhattacharya, and **Subrata Tikadar**. [Under Examination].

Journal papers out of this thesis

- **Subrata Tikadar**, and Samit Bhattacharya. (2020) "Detection of affective states of the students in a blended learning environment comprising of smartphones", *International Journal of Human-Computer Interaction*, Taylor & Francis. [DOI: 10.1080/10447318.2020.1861762].

- **Subrata Tikadar**, and Samit Bhattacharya. "A Computational Model for Detecting Involvement of the Students in a Blended Learning Platform", *IEEE Transactions on Learning Technologies*, IEEE. [Under Review].

- **Subrata Tikadar**, and Samit Bhattacharya. "Vedinkaksha: A 'Sensitive' Blended Learning System for Classroom Teaching", *The ACM Transactions on Intelligent Interactive System*, ACM. [Under Review].

International conference papers out of this thesis

- **Subrata Tikadar**, and Samit Bhattacharya. (2019). "A Novel Method to Build and Validate an Affective State Prediction Model from Touch-Typing." In *The 17th IFIP TC.13 International Conference on Human-Computer Interaction* (INTERACT), Springer, Cham. Paphos, Cyprus. pp. 99-119.
<u>CORE Conference Rank: A</u>

- **Subrata Tikadar**, and Samit Bhattacharya. (2019). "Predicting Students' Involvement in Blended Learning Environment" In *19th IEEE International Conference on Advanced Learning Technologies* (ICALT), IEEE. Maceio, Brazil. pp. 76-78. ***[Best Paper Award]***.
  CORE Conference Rank: B

- **Subrata Tikadar**, and Samit Bhattacharya. (2019). "How Do They Use Their Smartphones: A Study on Smartphone Usage by Indian Students" In *The 17th IFIP TC.13 International Conference on Human-Computer Interaction* (INTERACT), Springer, Cham. Paphos, Cyprus. pp. 132-151.
  CORE Conference Rank: A

- **Subrata Tikadar**, and Samit Bhattacharya. (2018). "A Blended Learning Platform to Improve Teaching-Learning Experience." In *2018 IEEE 18th International Conference on Advanced Learning Technologies* (ICALT), Springer, Cham. Paphos, Cyprus. pp. 99-119.
  CORE Conference Rank: B

- **Subrata Tikadar**, Sharath Kazipeta, Chandrakanth Ganji, and Samit Bhattacharya. (2017). "A Minimalist Approach for Identifying Affective States for Mobile Interaction Design." In *The 16th IFIP TC.13 International Conference on Human-Computer Interaction* (INTERACT), Springer, Cham. Mumbai, India. pp. 3-12.
  CORE Conference Rank: A

Vitae

Subrata Tikadar is a Research Scholar in the Department of Computer Science and Engineering of Indian Institute of Technology Guwahati. He received the MTech degree in Computer Science and Engineering from National Institute of Technical Teachers' Training and Research Kolkata, India, in 2010, and the BE degree in Computer Science and Engineering from University Institute of Technology, Burdwan University, India, in 2008. Since 2016, he has been pursuing his doctoral research in Computer Science and Engineering at Indian Institute of Technology Guwahati, India. Before joining the PhD program, Mr. Tikadar had more than five years of teaching experience at both the undergraduate and postgraduate levels. He served Bengal College of Engineering and Technology (Durgapur, WB, India) as an Assistant Professor, and Kingston Engineering College (Barasat, WB, India) as a Lecturer. He also worked at Maulana Abul Kalam Azad University of Technology which was formerly known as West Bengal University of Technology (Kolkata, WB, India) as a Visiting Faculty, and Aliah University (Kolkata, WB, India) and Central Institute of Technology Kokrajhar (BTAD, Assam, India) as a Guest Lecturer. His current research interests include Human-Computer Interaction, particularly Affective Computing and Learning Technology. He is involved in reviewing many top-tier international conferences including CHI, INTERACT, MobileHCI, and ICMI. Mr. Tikadar is a member of ACM, IEEE, and IEEE Computer Society Technical Committee on Learning Technology. His professional achievements include the recognition for outstanding review from ACM SIGCHI for a CHI paper review, the Grants (from IFIP TC13 on HCI, and the IFIP Digital Equity Committee) for Students and early stage researchers to attend INTERACT 2019 Paphos, Cyprus (September 2-6, 2019), and the Best Paper Award in the 19th IEEE International Conference on Advanced Learning Technologies (ICALT 2019, Maceio, Brazil) organized by the IEEE Computer Society and the IEEE Technical Committee on Learning Technology.

Contact Information

Email : t.subrata@iitg.ac.in,
subratatikadar@gmail.com

Web : http://iitg.ac.in/stud/t.subrata/

Address : Rangapur (Ramkrishna Pally), PO: Nilgunj Bazar
Distt. – North 24 Parganas, West Bengal, INDIA.
PIN – 700121.

Appendices

Appendix A

Consent Form Used in Controlled Experiments

A representative of the consent forms used in the controlled experiments conducted to complete the thesis work is presented in Figure A.1. It may be noted that the name of the experiment varied for each of the experiments, remaining all the other fields of the form intact. We conducted the experiments with the participants only when both the participants and experimenters agreed to all the conditions mentioned in the consent form.

Participant Consent Form

UCNET

Name of the research: Experiment for Identification of Students' Involvement in Classroom

Experiment(s) conducted at: UCCN Lab, Department of CSE, IIT Guwahati

I, the undersigned, confirm that

1. I have read and understood the information about the research and I have been given the opportunity to ask questions about the research and my participation.
2. I voluntarily agree to participate in the experiment.
3. I understand that I can withdraw at any time without giving reasons and that I will not be penalised for withdrawing nor will be questioned on why I have withdrawn.
4. The procedures regarding confidentiality have been clearly explained (e.g. use of names, pseudonyms, anonymization of data, etc.) to me.
5. The use of the data in research, publications, sharing and archiving has been explained to me.
6. I allow to record/capture video/audio/image while participating and I understand that those can be used solely for research purpose maintaining the anonymity.
7. I understand that other researchers will have access to this data only if they agree to preserve the confidentiality of the data and if they agree to the terms I have specified in this form.
8. I, along with the Researcher, agree to sign and date this informed consent form.

Participant:

S.V. KOGILAVANI	29/6/18	29/06/18
Name of Participant	Signature	Date

Researcher:

Subrata Tikadar		29/06/2018
Name of Researcher(s)	Signature	Date

Figure A.1: Consent form used in the controlled experiments

Appendix B

DFDs of a Sensitive Classroom System as per Vedinkaksha Framework

In order to design the sensitive blended learning platform based on the Vedinkaksha framework, we drew the DFDs for the system up to Level-3. In **Chapter** 7 (Subsection 7.1.1) we have presented only one of the DFDs, i.e., the 'Level-1' DFD of the system. In this appendix, we present all the DFDs along with the process decomposition details.

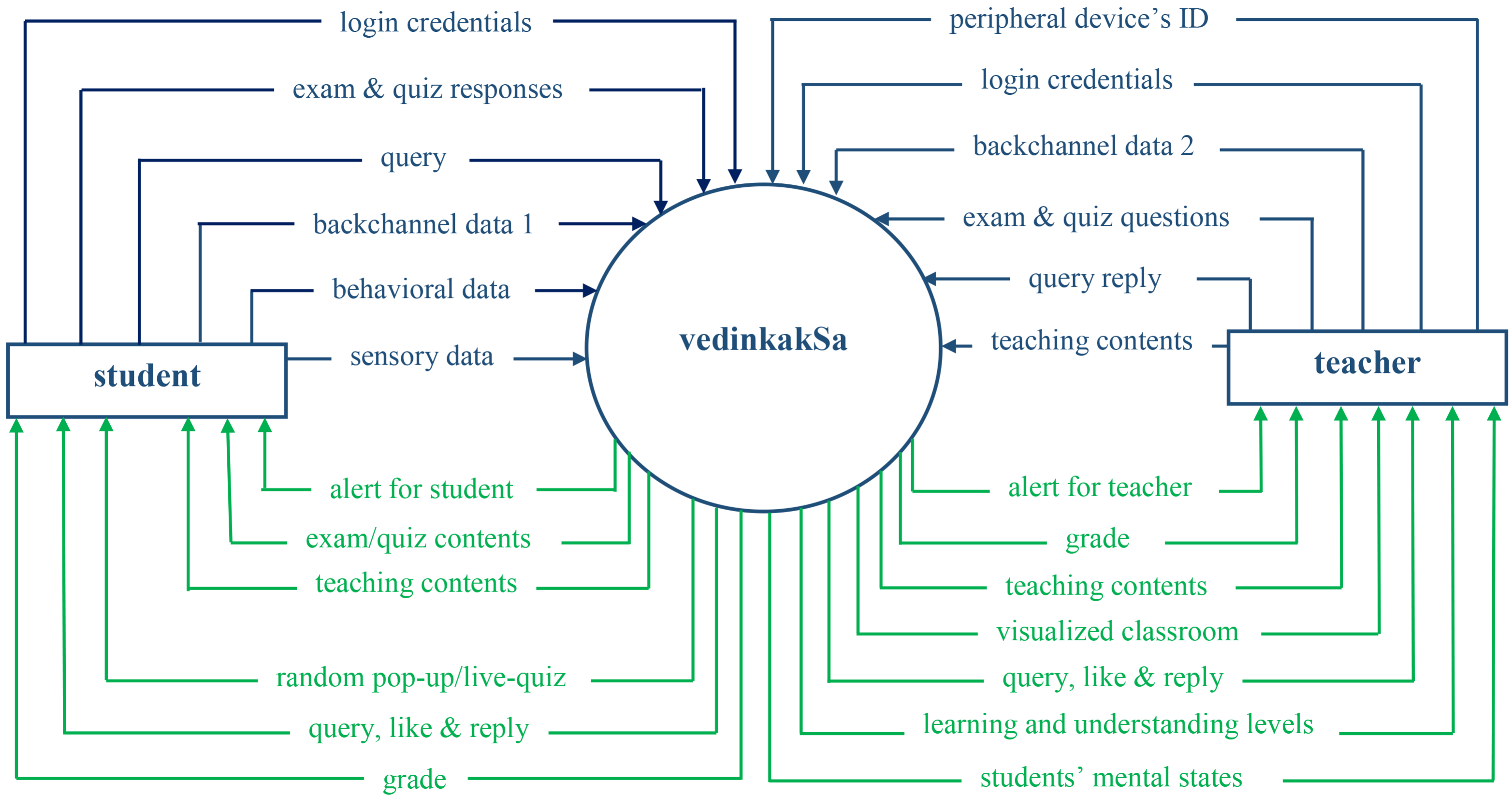


Figure B.1: Context level (0-Level) DFD of the sensitive blended learning platform

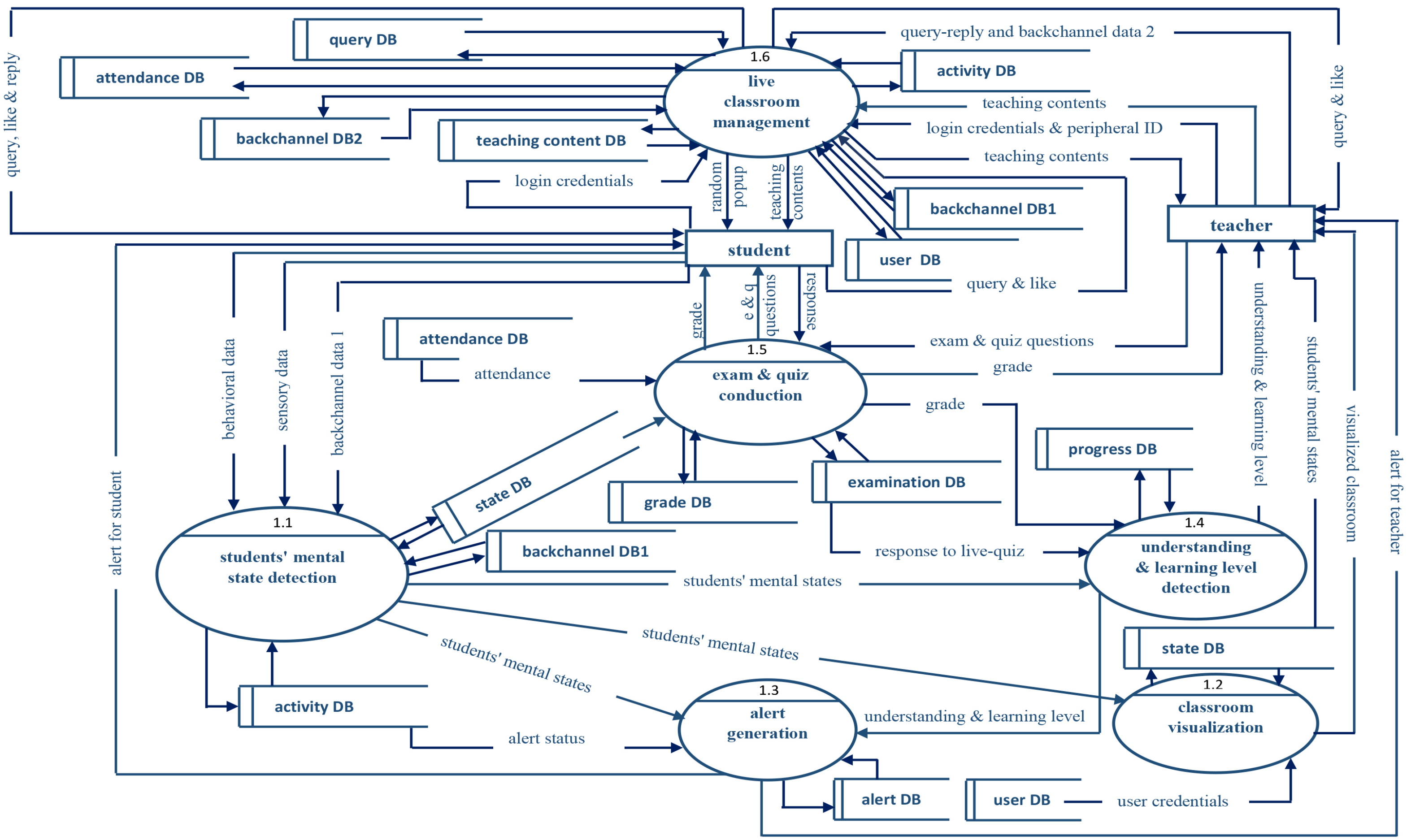


Figure B.2: Level-1 DFD of the sensitive blended learning platform

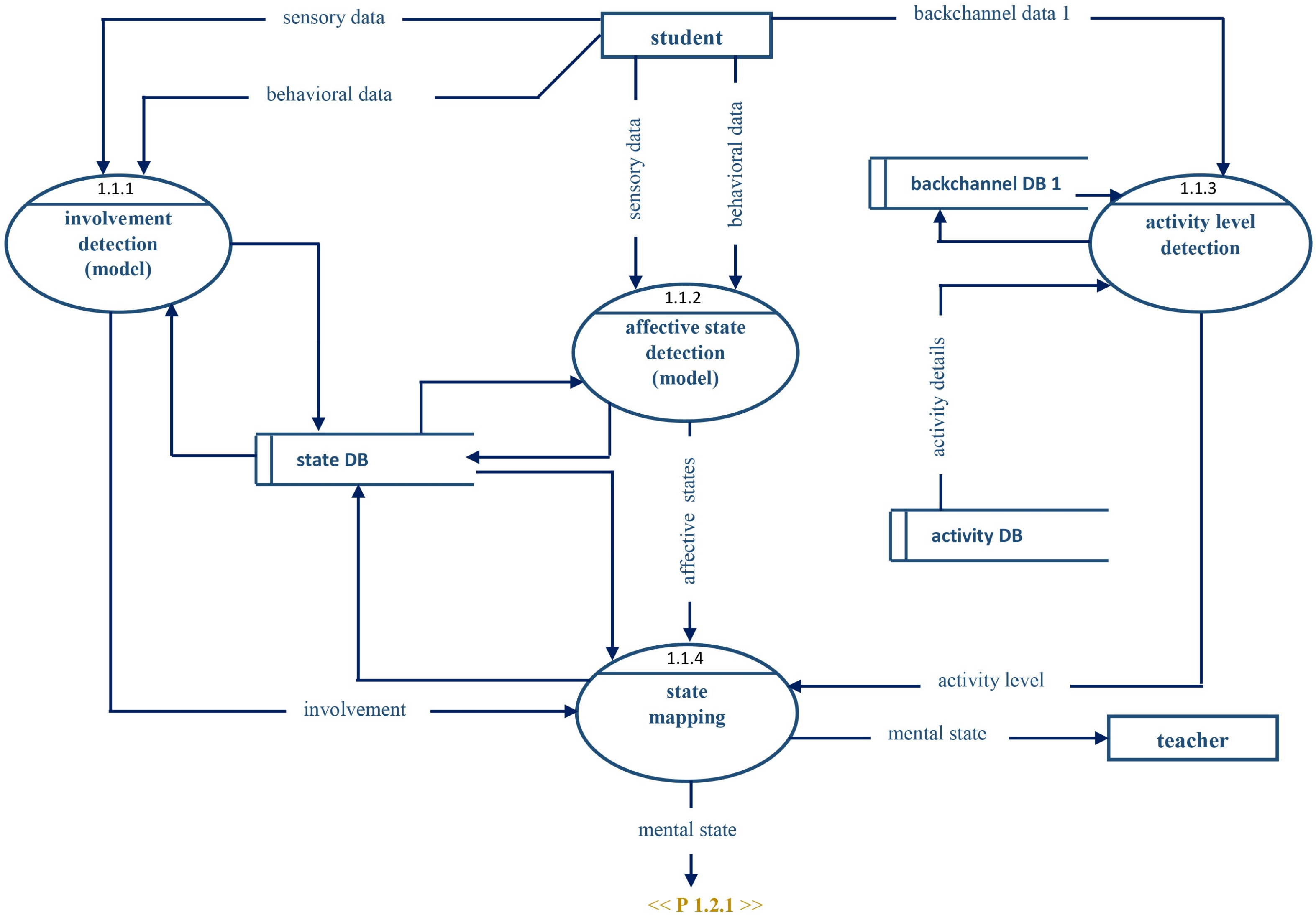


Figure B.3: Level-2 (of Process 1.1) DFD of the sensitive blended learning platform

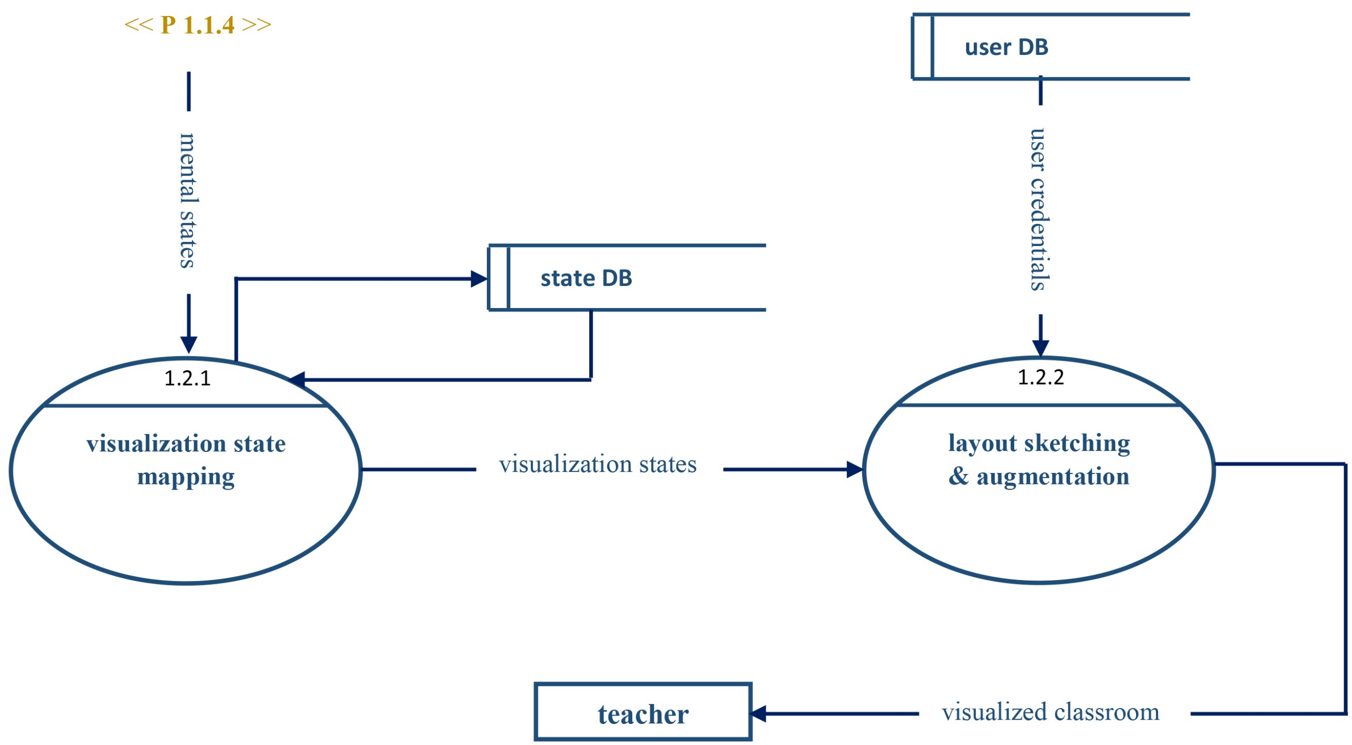


Figure B.4: Level-2 (of Process 1.2) DFD of the sensitive blended learning platform

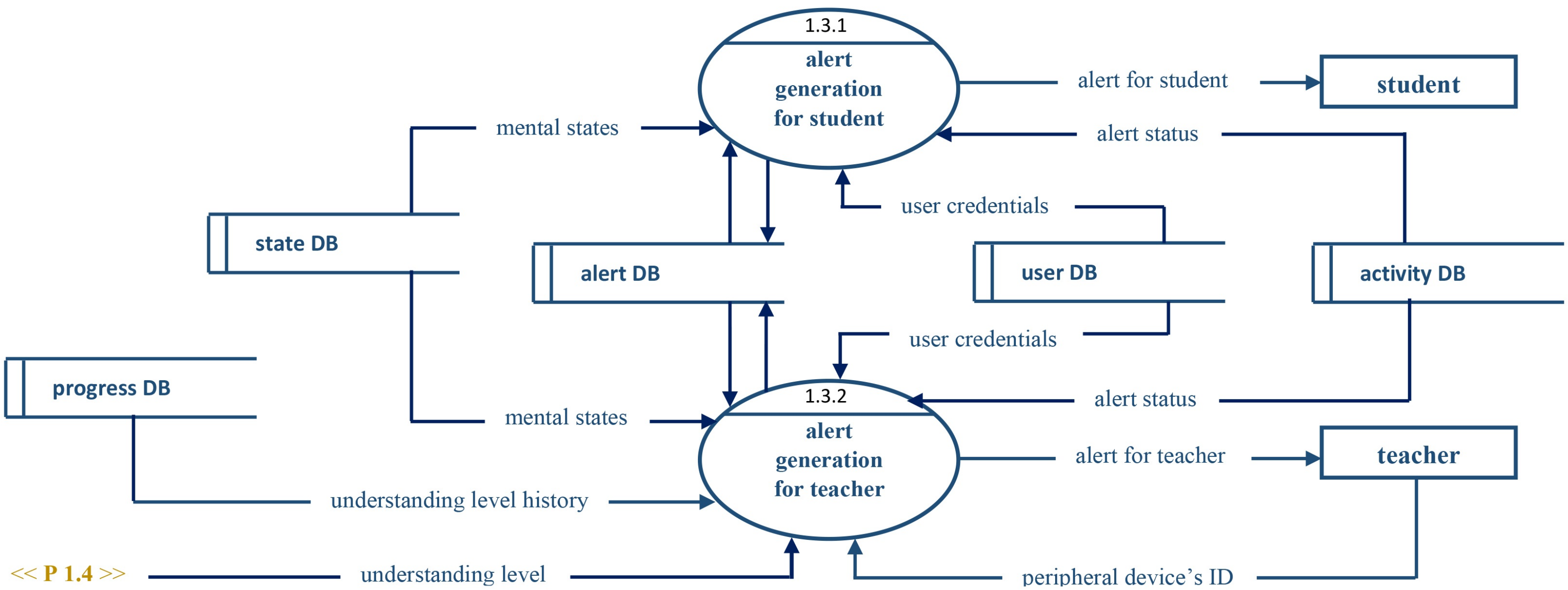


Figure B.5: Level-2 (of Process 1.3) DFD of the sensitive blended learning platform

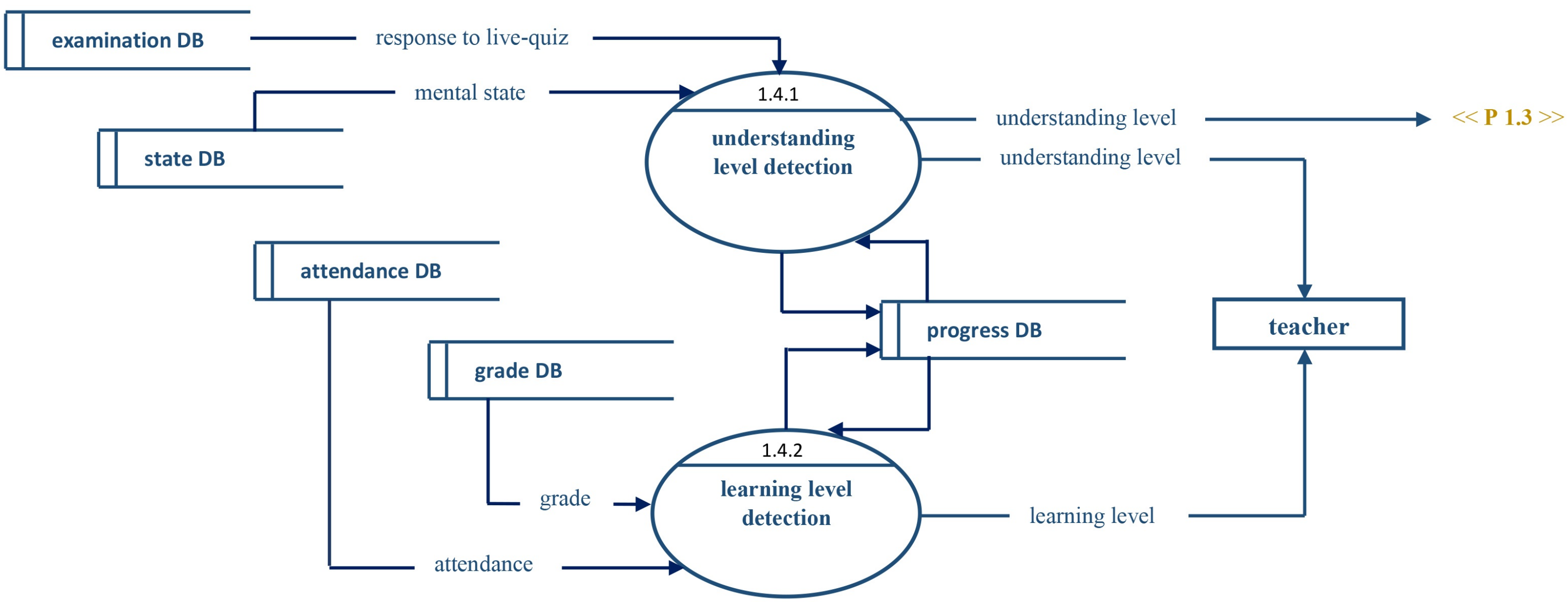


Figure B.6: Level-2 (of Process 1.4) DFD of the sensitive blended learning platform

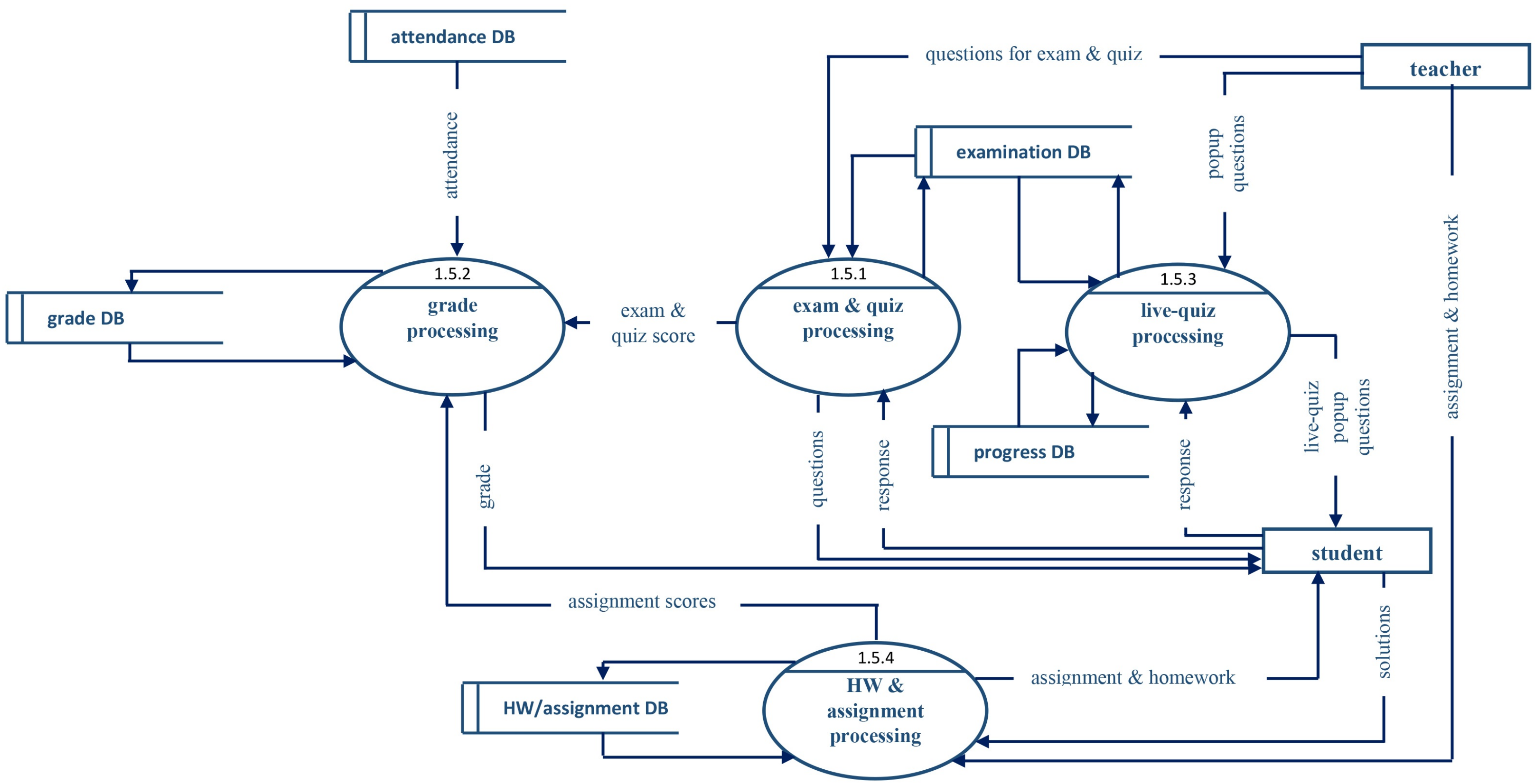


Figure B.7: Level-2 (of Process 1.5) DFD of the sensitive blended learning platform

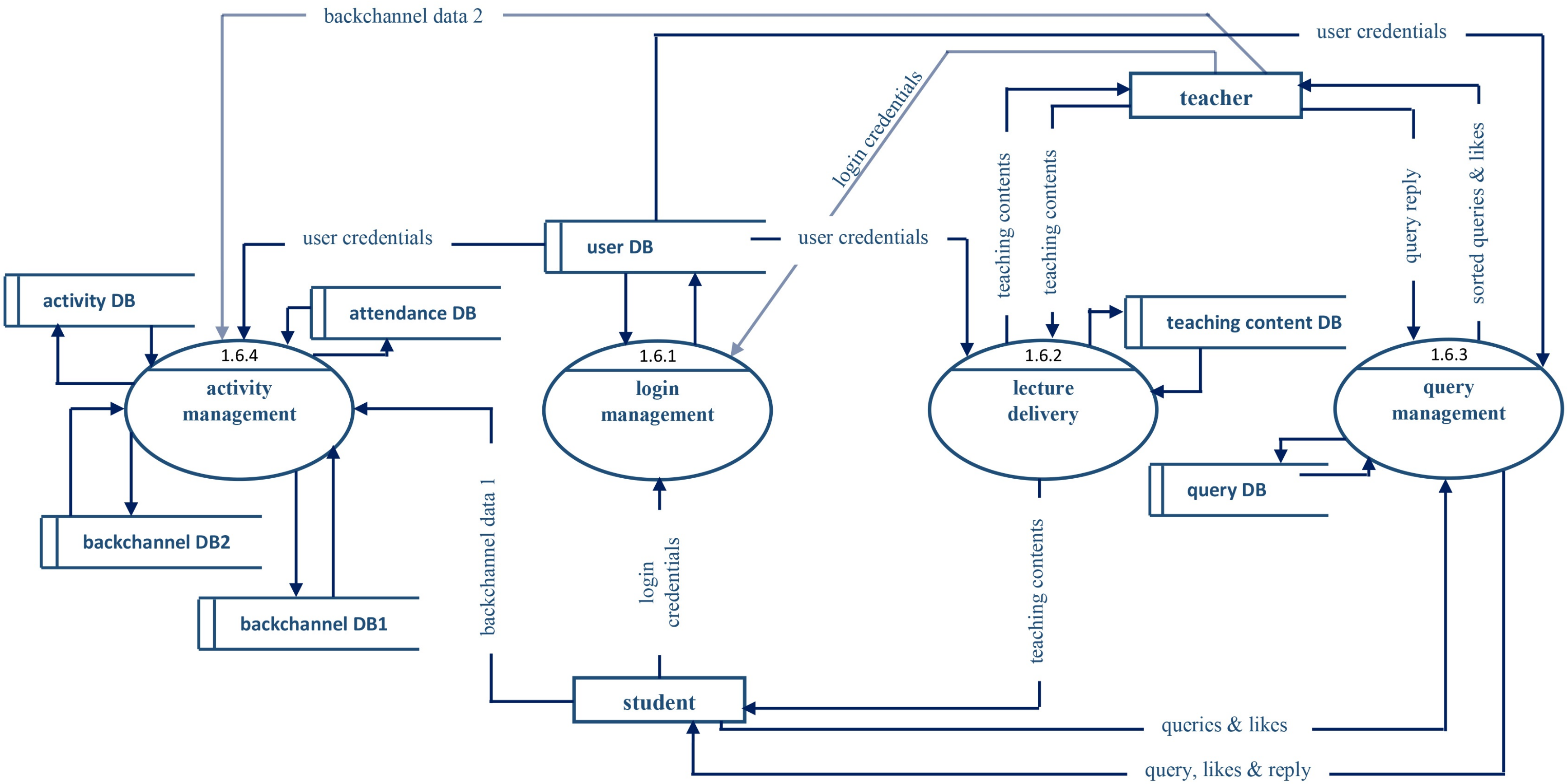


Figure B.8: Level-2 (of Process 1.6) DFD of the sensitive blended learning platform

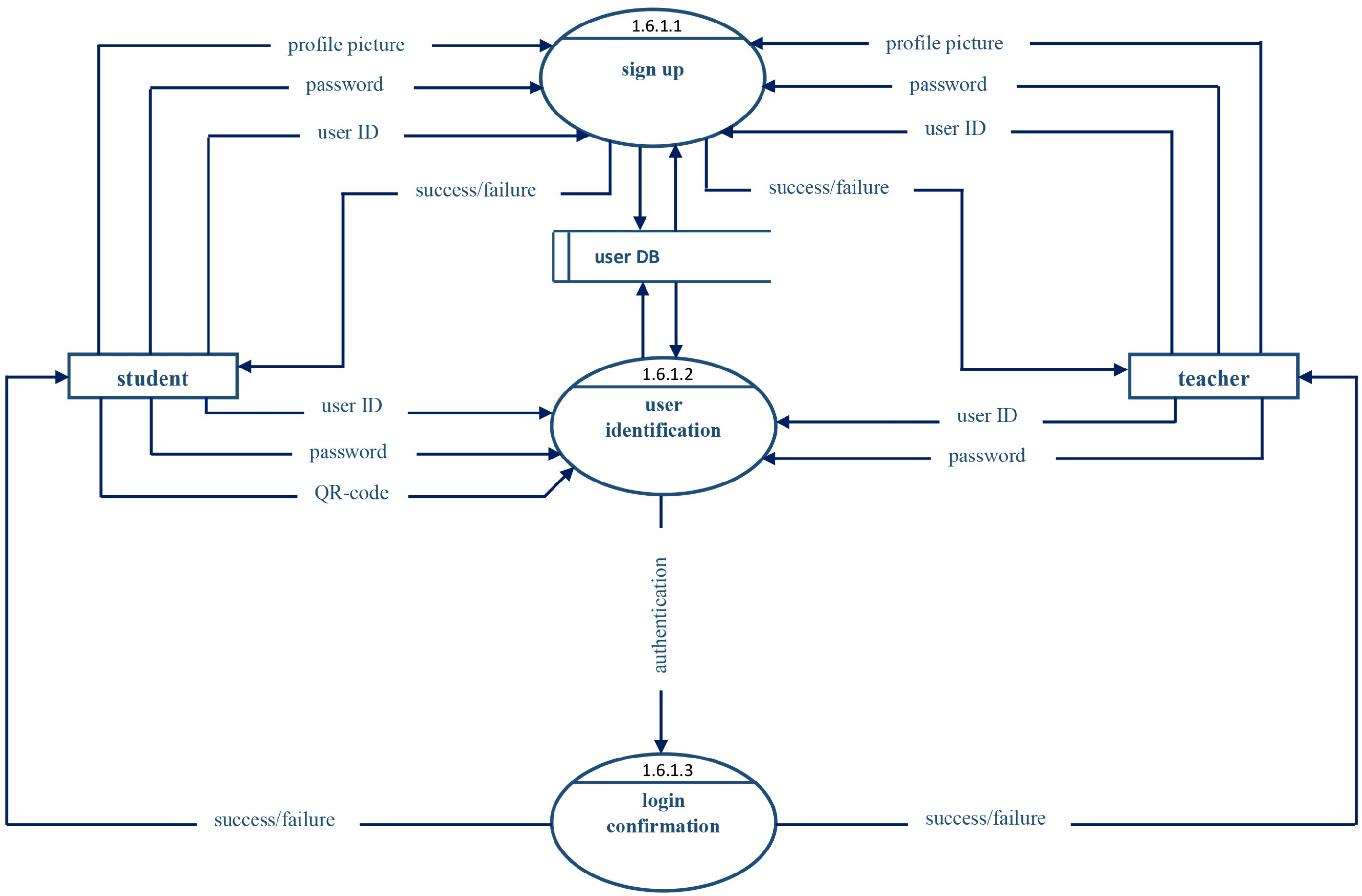


Figure B.9: Level-3 (of Process 1.6.1) DFD of the sensitive blended learning platform

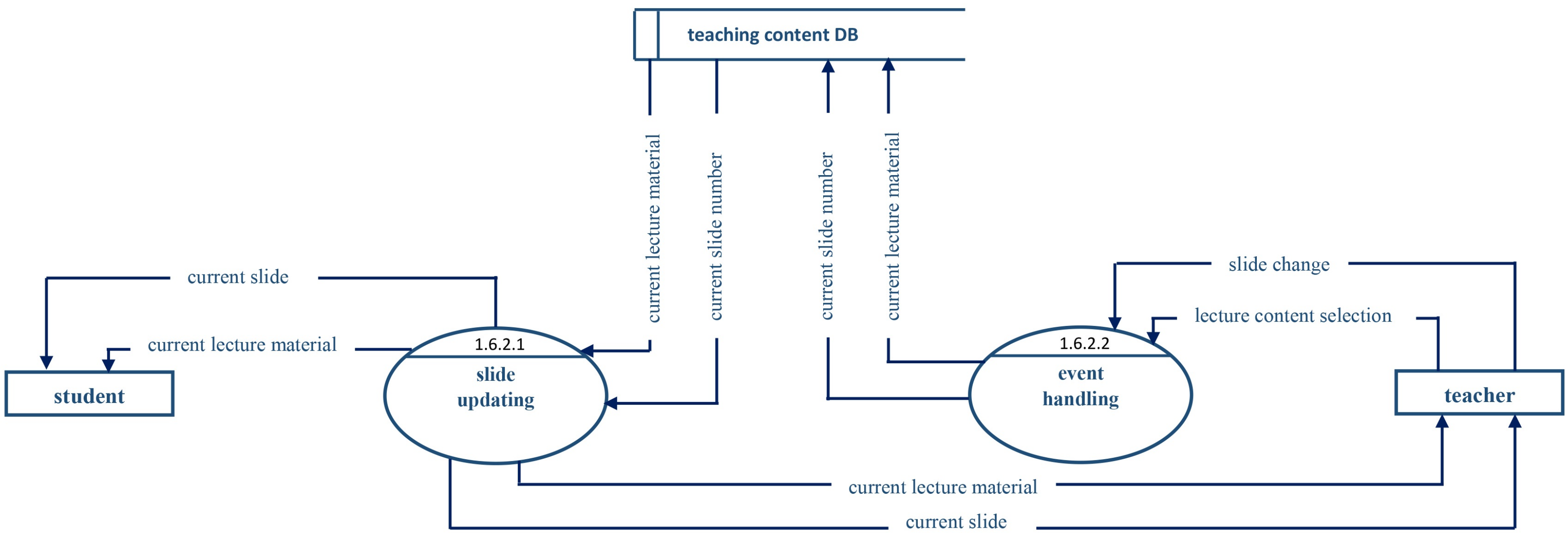


Figure B.10: Level-3 (of Process 1.6.2) DFD of the sensitive blended learning platform

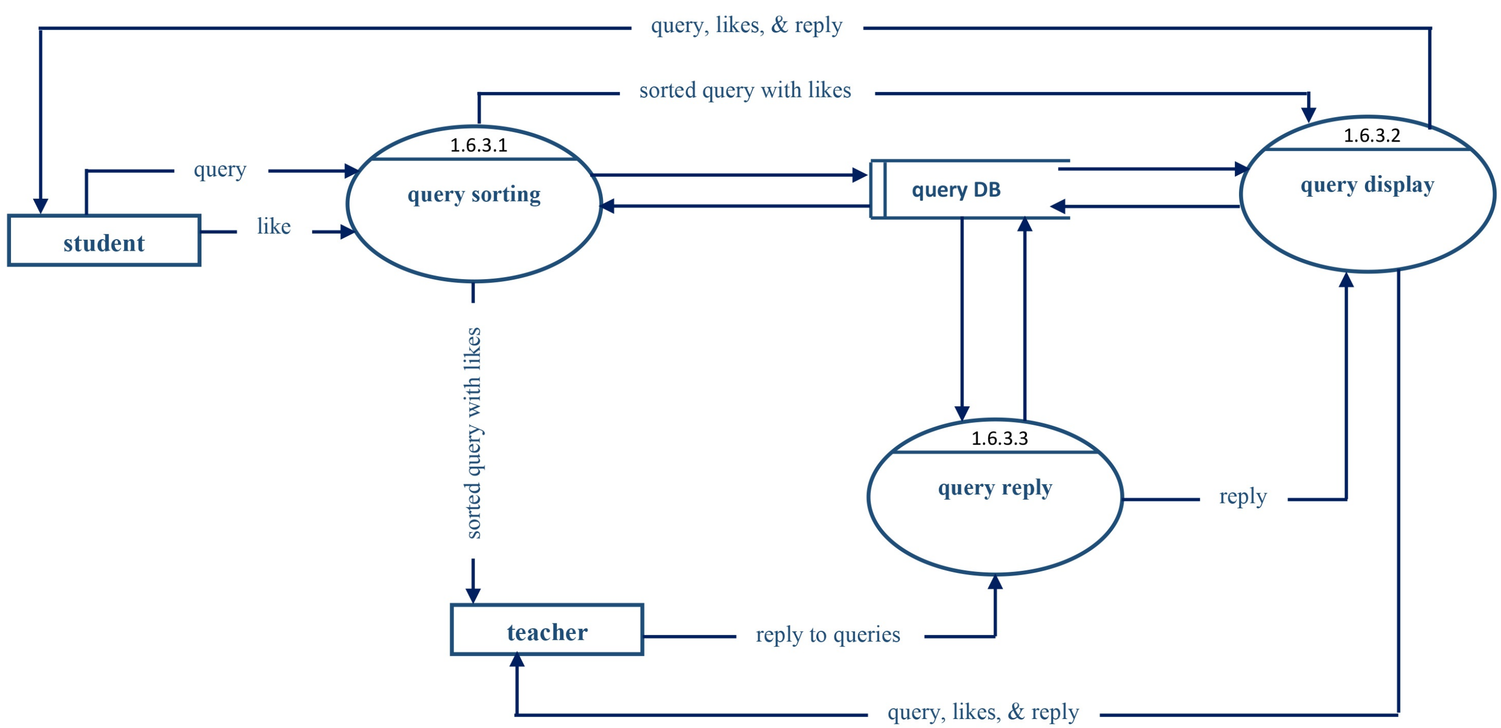


Figure B.11: Level-3 (of Process 1.6.3) DFD of the sensitive blended learning platform

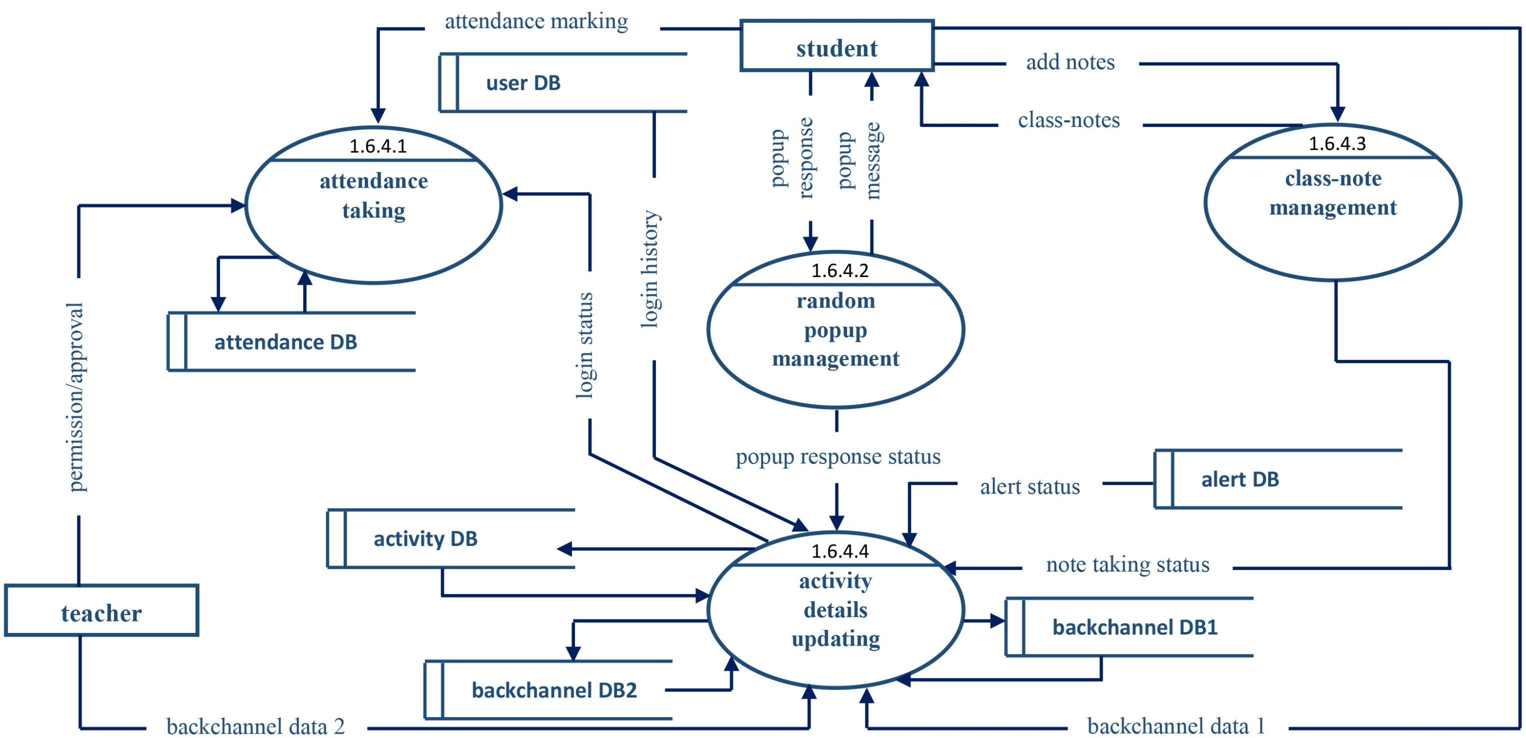


Figure B.12: Level-3 (of Process 1.6.4) DFD of the sensitive blended learning platform

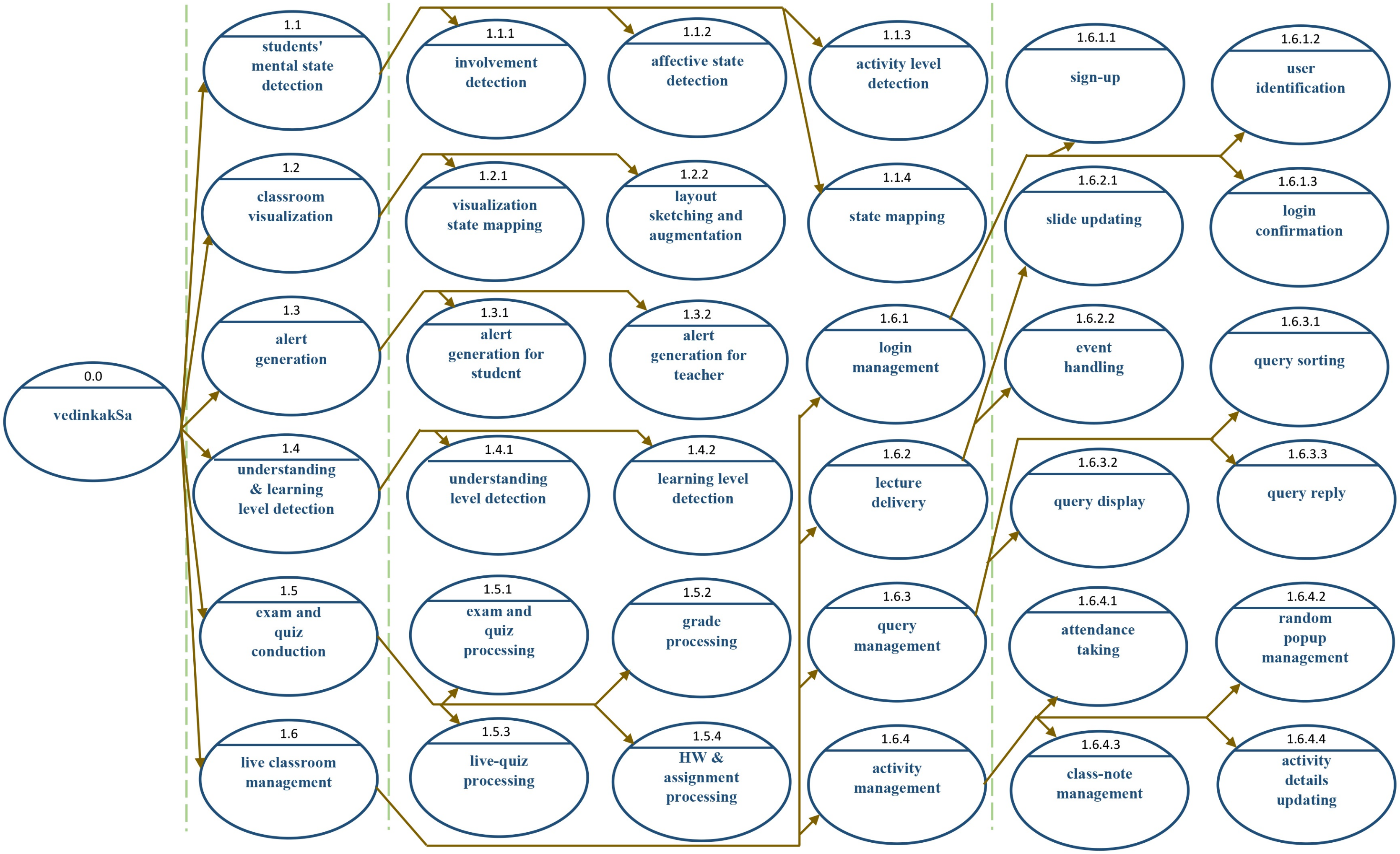


Figure B.13: Process decomposition of the sensitive blended learning system design