STUDY COMPANION

A PROJECT PHASE II REPORT

submitted by

AKHIL P DOMINIC(AJC19CS010)
ALAN JAMES(AJC19CS013)
CHRISTY CHACKO(AJC19CS049)
ELENA MARIA VARGHESE(AJC19CS059)

under the guidance of

FABEELA ALI RAWTHER

to

the APJ Abdul Kalam Technological University
in partial fulfillment of the requirement for the award of the Degree

of

Bachelor of Technology

in

Computer Science and Engineering

Department of Computer Science and Engineering

Amal Jyothi College of Engineering
Kanjirappally-686518

May 2023

DECLARATION

We undersigned hereby declare that the project report **"Study Companion"**, submitted for the partial fulfillment of the requirements for the award of degree of Bachelor of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by us under the supervision of Ms. Fabeela Ali Rawther. This submission represents our ideas in our own words and where ideas or words of others have been included, we have adequately and accurately cited and referenced the original sources. We also declare that we have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in our submission. We understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Kanjirappally
Date:

Akhil P Dominic

Alan James

Christy Chacko

Elena Maria Varghese

DEPT. OF COMPUTER SCIENCE AND ENGINEERING

AMAL JYOTHI COLLEGE OF ENGINEERING KANJIRAPPALLY

CERTIFICATE

This is to certify that the report entitled **"Study Companion"** submitted by **Akhil P Dominic (AJC19CS010), Alan James (AJC19CS013), Christy Chacko (AJC19CS049), Elena Maria Varghese (AJC19CS059)** to the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineering is a bonafide record of the project work carried out by him/her under my/our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Ms. Fabeela Ali Rawther
Internal Supervisor

Dr. Juby Mathew
Head of Department, CSE

Dr. Sinciya P. O.
Project Coordinator

External Supervisor

Acknowledgement

First of all we sincerely thank the Almighty GOD who is most beneficent and merciful for giving us knowledge and courage to complete the project successfully.

We derive immense pleasure in expressing our sincere thanks to our Manager, **Fr. Dr. Mathew Paikatt** and to our Principal, **Dr. Lillykutty Jacob** for the kind co-operation in all aspects of our Project.

We express our gratitude to **Dr. Juby Mathew**, HoD, Department of Computer Science and Engineering for his kind co-operation in all aspects of our Project. We express our sincere thanks to our internal guide, **Ms. Fabeela Ali Rawther**, Assistant Professor, Assistant Professor and our Project Co-ordinators **Ms. Shiney Thomas**, Associate Professor, **Dr.Sinciya P.O**, Associate Professor and **Prof. Manoj T Joy**, for their encouragement and motivation during the project.

We are indebted to our beloved teachers for their cooperation and suggestion throughout the project which helped us a lot. We also thank all our friends and classmates for their interest, dedication and encouragement shown towards the project. We convey hearty thanks to our parents for their moral support, suggestion and encouragement to make this venture a success.

Abstract

Working parents find it difficult to check upon their children's academic performance and monitor them frequently. AI-based study companion assistants can help reduce the workload of parents by providing an additional support system for their children's learning. This can take some of the burden off of parents, who may otherwise have to spend time helping their children with their studies or coordinating with teachers and tutors.One promising application of AI is the development of study companion assistants that use machine learning techniques to help students improve their academic performance. These assistants can provide personalized recommendations, feedback, and support based on a student's learning style, strengths, and weaknesses. In this project, we present a study companion assistant that uses ML techniques for drowsiness, yawn and phone usage detection to help students stay organized, manage their time, and develop effective study strategies. The assistant is designed to adapt to a student's needs and learning progress over time, providing support and guidance as needed. We demonstrate the effectiveness of our approach through a series of experiments and user studies, showing that our AI-based study companion assistant can significantly improve student performance and satisfaction.

Contents

List of Figures

Chapter 1

Introduction

Companion systems are designed to enhance well-being, quality of life, and independence by providing companionship and assisting daily life. This mainly includes cognitive and social support, health monitoring, and self-care support through human–system interaction.These systems gradually finding their way into various areas that improve children's experiences or help them develop valuable skills.

The average attention span of a human being is roughly around 25 minutes. It can vary by a variety of factors like age, health conditions etc.There are techniques like 'Pomodoro' that lets a person take breaks after set intervals of time. Students require great concentration and productivity in order to acquire knowledge. But trying to study for long periods would cause the concentration levels to drop and degrade his ability to understand the concepts. Therefore, breaks are indeed essential for a student to maximize his productivity.

Students require great concentration and productivity in order to acquire knowledge. But trying to study for long periods would cause the concentration levels to drop and degrade his ability to understand the concepts. In order to maximize the productivity of a student, it is necessary to find out how productive he is while working. It helps us to understand the amount of time he could focus on his work and take a break when required. A variety of behavioral patterns and facial expressions are evaluated in order to analyze a person's productivity and efficiency.

There are several different approaches to using computer vision for attention monitoring in learning. One common approach is to use a webcam or other camera to record the student while they are working.

The system uses blink rate monitoring, and expression detection to identify when you are struggling to understand a concept, or becoming fatigued, or distracted . By analyzing these subtle cues, the Study Companion can adapt to your needs, providing additional explanations or resources when you need them most.

The Study Companion also uses yawn rate detection to identify when you may be feeling drowsy, which indirectly helps one to stay alert and focused. With its advanced expression detection capabilities, the Study Companion can even provide personalized feedback on your learning progress. By analyzing your facial expressions as you work through a problem or concept, the Study Companion can identify what you are feeling. This information could be used to provide support and encouragement when the student is feeling positive, or to identify and address any negative emotions that may be hindering their learning.

To ensure that the AI-based study companion system is able to accurately detect and interpret these physiological and behavioral indicators, it would need to be trained on a large and diverse dataset of students. Once the system is trained, it could be deployed in a variety of settings, including classrooms, libraries, and online learning platforms.

The AI-based study companion system could also be valuable for educators and researchers. By collecting and analyzing data on students' physiological and behavioral indicators, educators could gain a better understanding of how students learn and identify strategies for improving learning outcomes. Researchers could also use the data to study the effects of different learning environments and techniques on student performance.

There are also a number of software tools and applications that are specifically designed for attention monitoring in learning using computer vision. These tools may be integrated with learning management systems or used as standalone applications. They may also be used in conjunction with other technologies, such as virtual reality or augmented reality, to provide a more immersive learning experience.

One potential benefit of using computer vision for attention monitoring in learning is that it can provide real-time feedback to both students and educators. This can help students understand where they may be struggling and allow them to make adjustments in real-time to improve their focus and engagement. It can also help educators identify areas where students may be having difficulty and provide additional support or resources as needed. Another potential benefit is that it can help to identify patterns and trends in student attention and engagement over time. This can be useful for identifying factors that may be impacting student performance and for developing strategies to improve learning outcomes.

In addition, the study companion assistants can help parents keep track of their children's progress and identify any areas where they may be struggling. This can allow parents to provide targeted support and assistance, rather than having to constantly monitor their children's work or spend time trying to understand the material themselves.

.The utilized model is constructed utilizing computer vision and ma- chine learning techniques. OpenCV is an open-source computer vision library that can be used to perform various image and video processing tasks, such as face detection and recognition. Dlib is a C++ library that provides tools for machine learning, including facial landmark detection and object detection. The Haar Cascade files, in conjunction with OpenCV, are utilized to classify faces, left eyes, and right eyes.OpenCV is employed to capture images from a webcam and feed them into a Deep Learning model for the purpose of determining if a person's eyes are "Open" or "Closed".

Overall, the AI-based study companion system using ML techniques such as blink rate, yawn rate, to revolutionize the way students learn and retain information. By providing real-time feedback and support, this system could help students stay focused and engaged, leading to improved learning outcomes and academic performance.

The video processing occurs locally thus leading to privacy and security advantages since it prevents the diffusion of sensitive personal situations and also optimises internet bandwidth occupation. The organisation of the paper is as follows. Chapter 2 provides an overview of the works in the literature that analysis the characteristics mentioned above to infer information on the subject's cognitive state. Chapter 3 contains the description of the proposed approach and the adopted Machine Learning inferences. Chapter 4 concludes the study and the potentials to be further explored in future research activities.

Chapter 2

Literature Survey

2.1 ChildBot: Multi-robot perception and interaction with children

Authors : Andrea F. Abate, Lucia Cascone, Michele Nappi, Fabio Narducc
Indroduction: In [1] ChildBot is a system designed for use in educational and entertainment tasks with children. The system features multiple agents and multimodal perception modules, and has been tested with a total of 52 children in multiple experiments. The results of these experiments showed improved perception capabilities compared to previous works, and also provided initial insights into the user experience with the system.

Methodology: ChildBot is a system that utilizes a Sense-Think-Act paradigm and is designed for use in indoor environments. It employs external sensors arranged within a "smart space" where interactions with children take place. By using external sensors, the system is able to overcome common problems in human-robot interaction (HRI) such as occlusions and improve the performance of its perception modules through the fusion of different data streams. This robot-agnostic architecture also allows for the integration of new robots into the system. The ChildBot system coordinates complex and continuous HRI procedures involving child actions and robot responses. Multimodal information from the sensors is first used to extract high-level information about the interaction context, which is then used to determine the appropriate response or action to take. Finally, the system sends this decision to the agent to execute the appropriate action[1].

The proposed ChildBot system is noteworthy for both the variety of its modules and the integration of those modules. The integration of perception modules is important for supporting more complex child-robot interaction (CRI) scenarios that require multimodal human-robot interaction (HRI). The modular design of the system allows different perception modules to be selected and used in specific applications without disrupting the overall functionality of the system. Additionally, the robot-agnostic architecture of the system enables the integration of multiple robots and the ability to switch between them based on user preference, as well as the ability to easily incorporate new social robots as they become available[1] .

To demonstrate the capabilities and versatility of the ChildBot system and its individual perception modules, five different educational and entertainment use cases have been designed. These use cases highlight the various applications that can be supported by ChildBot and showcase the different system components. A group of 52 children participated in these use cases with the robots, providing data that allowed for an objective evaluation of the performance of each module of the ChildBot system in terms of perception capabilities. In addition, a preliminary evaluation of the user experience yielded encouraging results and suggests the possibility of conducting a more comprehensive, subjective evaluation in the future [1] .

ChildBot is a improved version of previous works by the authors on multi-robot perception and interaction, and is designed to support a wide range of educational and entertainment scenarios involving child-robot interaction (CRI). The system integrates multiple robotic agents and sensors under a single, modular three-layer architecture and includes improved perception modules. It has been tested extensively using data from spontaneous interactions with children, and has been evaluated through the implementation of indicative use cases that demonstrate the wide range of applications it can support. The most important contributions of the ChildBot system include: 1) the design and implementation of an integrated system for human-robot interaction (HRI) using multiple robotic agents, 2) the development of perception modules for multimodal scene understanding that are tailored to CRI conditions and perform at or above the state-of-the-art in the underlying technologies, and 3) the collection and use of spontaneous children's data during CRI to enable an extensive objective evaluation of the system's capabilities and a preliminary user experience study [1].

2.2 Attention monitoring for synchronous distance learning

Authors : Andrea F. Abate, Lucia Cascone, Michele Nappi, Fabio Narducci
Introduction: In [2],The suggested educational strategy uses a synchronous online learning platform for video lectures along with a software component that can conduct in-depth analysis on the videos that students record using their laptop

cameras (or simulated through a webcam on a desktop PC). The objective of this study is to utilize pre-existing Machine Learning (ML) algorithms described in previous literature, which can measure the level of audience engagement by tracking indicators such as blinks, gaze, and facial expressions of students.

Methodology: The figure in 2.1 shows the architecture of the proposed educational platform, which consists of a video conference system (such as Microsoft Teams or Zoom) used by both the teacher and students for synchronous distance learning. The teacher shares a PowerPoint presentation and the students attend the video lecture. Before the lecture begins, the students launch the “Attention Prober” application and the teacher launches the ”Attention Monitor.” The Attention Monitor aggregates information for the teacher and displays a heat map of students’ gaze directions and the distribution of classified expressions in real-time. The applications are connected through the Photon networking engine, which is designed for use in online gaming but has been adapted to support distance learning by metaphorically shifting the concept of a ”room” to that of a“classroom” (the ”Didactic Room”). The teacher can arrange the applications on their main screen or use a second monitor, if available, to display the Attention Monitor game[2].

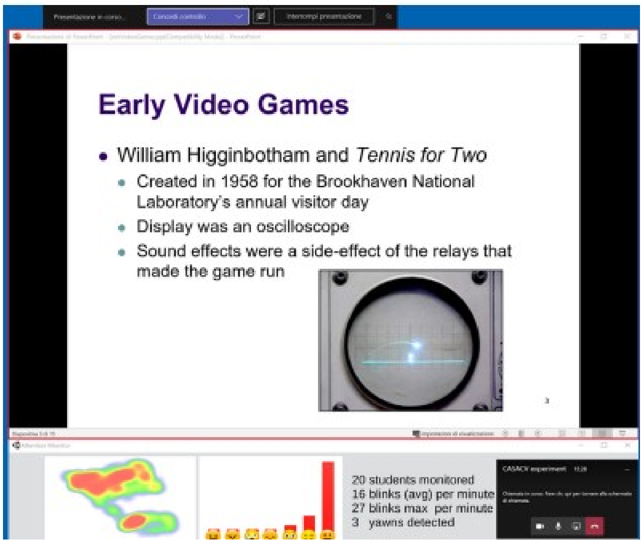


Figure 2.1: The Attention Monitor GUI, set by side of teacher Video-lecture software

During the controlled experiment, the Microsoft Teams video conference platform was used for the distance lecture, and PowerPoint was used to deliver the lecture in reading mode. The full-screen presentation modes of PowerPoint were not used because they would have covered the ”Attention Monitor” application. The teacher’s graphical user interface (GUI) consists of the content of the lecture in the upper portion of the screen, and instant feedback on the audience's attention in the form of a cumulative heat map, a histogram of expressions, and measurements of blink and yawn rates in the lower portion. Students do not need a GUI, as the Attention Prober runs in the background and analyzes video frames from the webcam without requiring user input. A GUI on the student side could also be a distraction from the lecture[2].

The proposed module uses computer vision solutions to track students' attention during a lecture. It analyzes students' behavior in real-time through the detection of blinks, yawns, and fixations using the camera mounted on the screen. The module has implemented four tasks based on machine learning and computer vision: gaze tracking, blink detection, expression detection, and yawn detection. Gaze tracking locates the eyes and generates a heat map of the sight. Blink detection counts the number of eye blinks during the lecture and the relative frequency. Expression detection uses a machine learning model to identify user expressions during the lecture. Yawn detection counts the number of yawns during the lecture as a potential indicator of limited attention to the lecture. These tasks provide overall feedback on the effectiveness of students' attention to the lecture[2].

The gaze tracking task of the proposed module uses face landmarks detection to match the point projected on the screen with the movement of the eyes. It also uses a calibration process to accurately estimate the fixation point on the screen. The module has been implemented in the GazeFlow application and generates heat maps to show the student's attention on the slide during the lecture. The calibration process involves the student sitting in front of the monitor and following a moving marker on the slides with their eyes. This process helps to estimate the distance between the student and the monitor and make the module slightly invariant to changes in head pose and limited translations of the head. The estimation of the point being observed on the screen is learned through a simple convolutional neural network. The lightweight version of the neural model helps to keep the computing demand of the module low, making it suitable for use in the background without impacting the student's ability to attend to the lecture[2].

2.3 Personalizing a Service Robot by Learning Human Habits from Behavioral Footprints

Authors : Kun Li and Max Q.-H. Meng

Indroduction: In [3] presents a technique for training a robot to learn an operator's habits through observation, using inverse reinforcement learning. The operator's behavior is captured through the environment's state changes caused by their actions, referred to as behavioral footprints, while the robot learns from cues based on the operator's interactions with objects in their environment. Using these learned habits as a reward function, the robot can autonomously serve the operator by guiding its future actions. The method is demonstrated in a case study involving the autonomous adjustment of indoor temperatures. Our contributions include the use of behavioral footprints to represent the operator's behavior and the proposal of a personalized robot based on the operator's habits.

Methodology: Here a method for a personal robot to learn the habits of an operator based on observations using inverse reinforcement learning. The robot observes the cues of the operator's actions and learns the habits as a reward function based on the operator's behaviors. The robot then uses this reward function to guide its future actions in order to serve the operator autonomously. The proposed method is demonstrated through a case study of autonomous temperature adjustment in an indoor environment. The contributions of this work include the use of behavioral footprints to represent the operator's behaviors and the proposal of robot personalization based on the operator's habits.

2.4 Speech Recognition using Machine Learning

Authors : Vineet Vashisht, Aditya Kumar Pandey, Satya Prakash Ya- dav
Introduction: Speech recognition is a rapidly growing technology in the field of engineering with numerous applications and potential benefits. Language barriers prevent many individuals from effectively communicating, so our project aims to reduce this obstacle.A system is designed to enable people to operate a computer using voice input in specific situations, providing significant assistance with sharing information. Our project is devoloped to recognize speech, convert audio input into text, and allow users to perform file operations such as Save, Open, or Exit through voice-only commands. The project also have incorporated the ability to recognize human speech as well as audio clips, and our system can translate between English and Hindi with text output. Additionally, options to convert audio between the two languages have been provided.

Methodology: The methodology in [4] aims to improve communication for people who are unable to do so due to language barriers by developing a system that can recognize and translate speech in English and Hindi. The system will have the ability to recognize human voices and audio clips and convert them into text. It will also have the ability to perform file operations such as Save, Open, or Exit through voice-only input. The primary algorithm used for this purpose is neural machine translation, which involves the use of two recurrent neural networks to create an encoder-decoder structure. The project will also eventually include a feature to provide dictionary definitions for Hindi and English words. This report provides an overview of speech recognition technology and its applications in various sectors, as well as a review of software developments in the field.

2.5 Robots and Wizards: An investigation into natural human robot interaction

Authors : Dominykas Strazdas; Jan Hintz; Anna-Maria FelBberg; Ayoub Al-Hamad
Introduction: The aim of the study in [5] was to investigate various methods of communication for intuitive interaction between humans and robots.By creating the illusion of a system without any restrictions, the study offers valuable insights into how users interact with industrial robotic assistants, providing a deeper understanding of the user experience. This approach allowed us to observe truly intuitive human-robot interaction.

Methodology:A Wizard of Oz prototyping method was used to allow unrestricted, intuitive interaction with an industrial robot. Results from 36 test subjects showed a strong preference for speech input, automatic path planning, and pointing gestures. The catalog created during the experiment identified common gestures that can cover the needs of most users. The user interface received an average score of 74% in user experience questionnaires, despite containing intentional flaws. These results can inform the development of a highly accepted intuitive human-robot interaction system in the future.

2.6 Continuous monitoring and identification of driver drowsiness and alert system

Authors :M.S.Satyanarayanaa, T.M.Arunab, Y.K.Guruprasad
Introduction: The aim of study [6] was to develop a method to alert drowsy drivers in order to reduce the number of accidents, which often occur between the hours of 2:00am to 5:00am. The objective of this paper is to evaluate the specific actions of drivers and determine their level of lethargy. Based on this

assessment, drivers will be recommended either full-length sleep or power naps to prevent automobile collisions. A device has been created, which is equipped with an innovative algorithm that can make personalized recommendations for each driver.

Methodology: The proposed approach used a camera and module to detect tiredness in drivers by analyzing the frequency of head tilting and eye flicker.The primary focus of the proposed method is to detect driver drowsiness in real-time, as this is a critical issue that requires immediate attention. Real-time data is given priority, which involves recording the driver's eye movements every 5 to 10 minutes or continuous monitoring of eye position to inform the driver and avoid collisions. The method also utilizes graphs to ensure output accuracy by comparing the timing diagram patterns of identified positions through computer vision. The proposed method involves a computer vision system that can automatically detect driver drowsiness in real-time video streams. An alarm is triggered, and the service provider is alerted if the driver appears to be drowsy. In testing on volunteers, the accuracy of face and eye identification was high. The hope is that this method will help prevent some accidents by alerting drivers who are feeling tired.

2.7 Using AI to understand what causes sentiment changes in social media

Authors : Fuad Alattar And Khaled Shaalan
Introduction : The research [7] aims to develop a tool for sentiment analysis on social media platforms like Twitter, which can track and interpret changes in sentiment towards entities, events, products, solutions, and services. The tool will use a Filtered-LDA framework, which utilizes cascaded Latent Dirichlet Allocation models with multiple settings of hyperparameters to identify candidate reasons for sentiment changes and a filter to remove tweets discussing old topics. The tool will also use a topic model with a high coherence score to extract emerging topics that can be interpreted by a human. The tool will be tested and evaluated, and a dashboard will be developed to display the most representative tweet for each candidate reason.

Methodology : The study introduces a new framework called Filtered-LDA that surpasses existing methods in performance. The framework utilizes a Cascaded LDA block with multiple LDA hyperparameter values to focus on both inside and outside aspects of texts to detect Emerging Topics. An improved sentiment variation measurement is also incorporated to accurately identify changes in sentiment levels. The framework generates outputs that include traditional HDP topics, as well as Filtered-LDA Emerging Topics separately.

2.8 Cybersecurity in Big Data: Era From Securing Big Data to Data-Driven Security

Authors : Danda B. Rawat , Senior Member, IEEE, Ronald Doku, and Moses Garuba

Introduction : Big data analytics (BDA) is a method for processing, storing, and gathering large amounts of data for further examination[8]. The rapid growth of the internet, internet of things (IoT), and other technological advances has led to an increase in the amount of data being produced. In this paper, the focus is on exploring recent research in the field of cybersecurity in relation to big data. The goal is to understand how big data can be protected and also used as a tool for cybersecurity. The paper summarizes recent research, presents trends and open research challenges, and offers insights into the relationship between cybersecurity and big data.

Methodology : This survey begins by emphasizing the potential risks and threats posed by IoT systems. To address these concerns, we explore the application of machine learning (ML) models to effectively learn from the vast amounts of data and enhance the security of the IoT network. We provide an overview of different models of learning and examine the advantages and disadvantages of each. Additionally, we delve into the use of Federated Learning (FL), a new and innovative learning model, for the security of IoT networks.

2.9 Classification and prediction of student performance data using various machine learning algorithm

Authors : Harikumar Pallathadka a, Alex Wenda b, Edwin Ramirez-Asís c, Maximiliano Asís-López d, Judith Flores-Albornoz e, Khongdet Phasinam f

Introduction : The study [9] aims to predict student performance in a course based on their past performance in similar courses[9]. It uses data mining techniques, including machine learning algorithms such as Nave Bayes, ID3, C4.5, and SVM, to uncover patterns in large amounts of data from students and teachers in order to classify or predict student performance. The UCI machine learning student performance dataset is used in the experimental study, and the algorithms are evaluated based on accuracy and error rate. The goal of this study is to help institutes improve their students' academic achievement by identifying their strengths and areas for improvement and providing targeted advice.

Methodology : The proposed framework for predicting student performance utilizes a dataset of student performance as input. The input data is preprocessed to eliminate any noise and ensure consistency. Subsequently, various machine learning algorithms, including Naive Bayes, ID3, C4.5, and SVM, are applied to the preprocessed dataset for classification purposes. The results of each algorithm's classification are then compared to determine the most effective approach.

2.10 Recent Advances on Federated Learning for Cybersecurity and Cybersecurity for Federated Learning for IoT

Authors : Bimal Ghimire , Graduate Student Member, IEEE, and Danda B. Rawat , Senior Member, IEEE
Introduction : The research [10] aims to use data mining and machine learning techniques to predict student performance in a course based on their previous performance in related courses. The study utilizes the UCI machine learning student performance dataset and evaluates different algorithms based on their accuracy and error rate. The ultimate goal is to assist educational institutions in identifying and supporting the academic needs of their students, with the hope of improving overall student performance.

Methodology : The survey is primarily focused on cybersecurity for the IoT environment and importantly using FL. Based on the nature and complexity of the proposed works, authors have adopted a variety of ML models. The only purpose of this section is to give readers information about the trends on kinds of ML models, algorithms, and technologies that have been used by the surveyed works along with the tools and environment under which the proposed works have been evaluated. For all the proposed works, the authors have adopted varieties of ML models, such as a neural network, SVM, linear regression, Q-learning, and so on. FL inherently supports privacy and security (compared to centralized learning), but to strengthen these, some works have also used elliptic-curve cryptography, differential privacy, blockchain, and others. The majority of the works have considered CNNs as their ML models. Different variations of CNNs, such as LeNet, AlexNet, GoogLeNet, VGGNet, and others, have been used. LSTM (a recurrent neural network) and MLPs (a feedforward neural network) also have been used by several works. Several works have adopted multiple of the ML models and compared the results to verify their proposed models.

2.11 Active Learning based on computer vision and HRI for the user profiling and behaviour personalization of an autonomous social robot

Authors : Marcos Maroto-Gomez, Sara Marques-Villaroya, Jose Carlos Castillo, Álvaro Castro-González, María Malfaz

Introduction : The study [11] presents a system that uses active learning, computer vision, and human-robot interaction to personalize the behavior of a social robot based on user recognition and profiling. The system is able to identify users through face recognition and gather information about them through interactions. To test the effectiveness of this system, three scenarios were conducted in which participants completed their profiles using an online survey, by interacting with a dull robot, or with a cheerful robot. The results showed that the cheerful robot received higher usability scores and that participants found it more entertaining to create their profiles with the cheerful robot than in the other scenarios. The system was also able to adapt its behavior based on whether the user was known or unknown.

Methodology :AL is a methodology based on allowing the algorithms used for learning to optimally select the data from which they learn. Consequently, the agent (in this case, our social robot) requires less training to obtain the information needed to select more accurate actions and perform better in dynamic environments.The actions of the users that interact with the robot are essential to producing optimal AL since they continuously provide feedback to the robot about how to update the algorithms used for learning.

They apply AL in this work in two different ways. On the one hand, they use computer vision techniques to identify the users that are actively interacting with the robot. If a user is not identified, they also include AL to dynamically guide the user so that they are correctly positioned in front of the robot and take face pictures that are used to train a face detector online, increasing the number of users that the robot knows. On the other hand, they use AL to retrieve user information to personalize HRIs actively. In this case, the user profiling algorithm is aware of which features have not yet been obtained from the user or can be updated, and it adapts the interaction to fill in the user profile in fewer interactions

2.12 Unsupervised ensemble clustering approach for analysis of student behaviour patterns

Authors : XIAOYONG LI, YONG ZHANG 1, HUIMIN CHENG 2, FEIFEI ZHOU1, AND BAOCAI YIN

Introduction : The study [12] aims to identify and understand the behavioral patterns of students in order to provide targeted services and management. It proposes a framework for unsupervised ensemble clustering of student behavioral data in order to discover patterns. The proposed framework combines density-based spatial clustering of applications with noise (DBSCAN) and k-means algorithms to extract behavior features from two perspectives and identify patterns. The framework is evaluated using six types of behavioral data from undergraduate students in a university in Beijing, and the relationships between different behavioral patterns and students' grade point averages (GPAs) are analyzed. The results show that the framework can detect both anomalous and mainstream behavioral patterns and can be useful for student-related departments in providing psychological consulting and academic guidance.

Methodology : The proposed framework has four stages: data collection, feature extraction, clustering analysis, and visualization and evaluation. Six types of behavioral data produced on campus were collected from different information management systems using the extract-transform-load tool. These data are classical time series data composed of events with time stamps. The features of every type of behavior are extracted from the two aspects of statistics and entropy; the statistical information represents the central tendency and dispersion of the distribution of behavioral data, and entropy represents the regularity of behavior. To alleviate the curse of dimensionality, the features with small variance and redundant features should be removed. In the clustering analysis stage, the DBSCAN algorithm is first applied to obtain initial clustering results; if a few very large clusters contain the vast majority of samples, then the k-means algorithm should be further used to subdivide them. In the final clustering results, the students that constitute noise and the students in small clusters discovered by DBSCAN can be considered anomalous, and the large clusters represent mainstream behavioral patterns. To evaluate the clustering results, parallel sets are used to visualize the distribution of the features in each cluster, by which we can intuitively under-

stand the differences among the clusters. In addition to visualizing the clustering results, we try to take students' semester grade point average (GPA) as the weak ground truth to measure the correlation between the clustering results of different behaviors and GPA.
Main contributions are summarized as follows.
1) Six types of behavioral data in time series format were collected, and the features for each type of behavior are extracted from the perspectives of central tendency, dispersion and entropy, which provides a more reliable basis for the analysis of behavioral patterns.
2) An ensemble unsupervised clustering framework is proposed by fully taking advantage of the DBSCAN and k-means algorithms; this framework can detect unexpected behavioral patterns and discover mainstream behavioral patterns. The clustering results provide helpful information for specialized management.
3) GPA levels are taken as the ground truth to calculate extrinsic metrics to measure the correlation between different behaviors and academic performance

2.13 MEBAL: A multimodal database for eye blink detection and attention level estimation

Authors: Roberto Daza, Aythami Morales, Julian Fierrez, Ruben Tolosana

Introduction: This work presents MEBAL1,[13] a multi-modal database for estimating attention level and detecting eye blinks. The frequency of eye blinks is correlated with cognitive activity, and automatic eye blink detectors have been proposed for a variety of activities, such as estimating attention level, analysing neuro-degenerative disorders, recognising deception, detecting driver weariness, or face anti-spoofing. However, the majority of the currently available databases and algorithms in this field are restricted to trials using just a small number of samples and unique sensors, like face cameras.

Methodology: Eye blink detection using Convolutional Neural Networks (CNN) is a common application of computer vision. CNNs are deep neural networks designed to process images or other multidimensional data with spatial relationships.

To perform eye blink detection using CNNs, the first step is to acquire facial images of the subject. These images can be obtained from a camera or a video stream. Once the images are obtained, they are preprocessed to extract the regions of interest (ROIs) corresponding to the eyes. This is typically done using face detection algorithms, such as Haar cascades or deep learning-based face detectors.

Once the ROIs are obtained, they are fed into a CNN architecture designed to classify the state of the eyes as either open or closed. The architecture typically consists of multiple convolutional layers, pooling layers, and fully connected layers. The convolutional layers learn features that are relevant for eye blink detection, such as the shape and texture of the eye region. The pooling layers reduce the spatial dimensionality of the feature maps, while the fully connected layers perform the final classification.

During training, the CNN is fed with a dataset of labeled eye images, where each image is labeled as either open or closed. The CNN learns to extract features from the eye region and classify the state of the eyes based on these features. Once the CNN is trained, it can be used to classify the state of the eyes in real-time, based on the ROIs extracted from the facial images.

In summary, eye blink detection using CNNs involves the following steps:

1. Acquire facial images of the subject.
2. Preprocess the images to extract the ROIs corresponding to the eyes.
3. Feed the ROIs into a CNN architecture designed for eye blink detection.
4. Train the CNN using a dataset of labeled eye images.
5. Use the trained CNN to classify the state of the eyes in real-time, based on the ROIs extracted from the facial images.

Chapter 3

Methodology

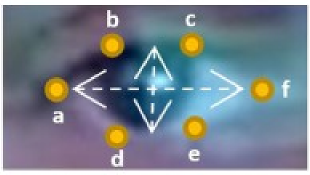
In a classroom setting, teachers can easily observe students and gauge their understanding and engagement. From the facial cues, the teachers could identify if the person is attentive to what he is learning. But self study sessions are what differentiates between a successful student and a mediocre one. Attention span refers to the amount of time it generally takes a person to lose interest in the task that they are doing. It could differ among individuals, but on average, the attention span of a person would be between 30-50 minutes. After that, he would lose interest and his productivity decreases exponentially according to proven research. There are several problems that students may face during self study, such as distractions, boredom, and difficulties with self-regulation. In this project, we are proposing a system that uses machine learning algorithms to analyze student videos recorded by the cameras on their laptops during a during a study session. The goal of this system is to provide the teachers as well as the parents with aggregated information on the efficacy of the a study session, without the need for students to share their video recordings. The adoption of additional solutions like this one is essential as it would lead to better productivity of the students. Our proposed system, includes a software module that runs on students' computers and analyzes their video streams to extract indicators of attention and engagement of the student during self study. In the initial phase, we are planning to run the system on the students' PC, but in the future we would be migrating the system to a Raspberry Pi based system. It would be sitting on desk as a device with a small form factor, analysing the students. The system also involves a smart AI assistant that interacts with the students to help them. The factors that would be used to determine the productivity would be

- Blink rate - We would detect how many times the eyes of the subject would blink in a given time frame

- Yawn rate - How many times the subject would yawn in a given time frame

The module runs in the background and does not require a graphical user interface on the student side, in order to minimize distractions. The system uses machine learning approaches, specifically computer vision techniques, to analyze the video frames captured by the student's camera and detect various behaviors such as blinks, yawns, and expressions. The system also uses microphone that would be used to facilitate the interaction of the student with the AI assistant. We are planning on using the open source AI assistant project for our virtual assistant. The user could ask interact with the assistant for clearing doubts and casual chatting. If our device finds that the student is distracted and his productivity level has fallen, the assistant would recommend activities that would benefit him. The device might ask him to take a walk, play his favourite music , play podcast etc. That way we might be helping him to avoid burn out.

For capturing the eye parameters, we would be using a measure known as Eye Aspect Ration(EAR) as in 3.1.



$$EAR = \frac{||b-d|| + ||c-e||}{2||a-f||}$$

Figure 3.1: Calculating the EAR by counting the eye blinks

The frequency of the eye blinks can be strictly related to cognitive activity. For calculating the blink rate, we would be using a blink counter. We would be using the EAR values to calculate the blink rate. The number of blinks per minute is termed as blink rate and on average it would be about 17 blinks/minute. The blink duration and the blink amplitude would also be calculated. The yawn rates is another important factor in indicating the lack of attention. The distances between the landmarks in the upper and lower lips can be detected and estimated to detect the yawn rate. The yawn duration and the yawn amplitude would be calculated as well.

We are planning on building an ensemble model that would be able to detect all the mentioned factors.

The block diagram of our desired system is shown in 3.2

The AI assistant would either be an open source voice assistant project or Alexa API, depending on how much we would be able to complete. Now, the video data from the device camera would be used for analysing the overall productivity of the student. The productivity analysis would be performed only on the basis of this camera feed. The audio data on the other hand would be used to interact the voice assistant. The processing would be done on the PC of the student. The

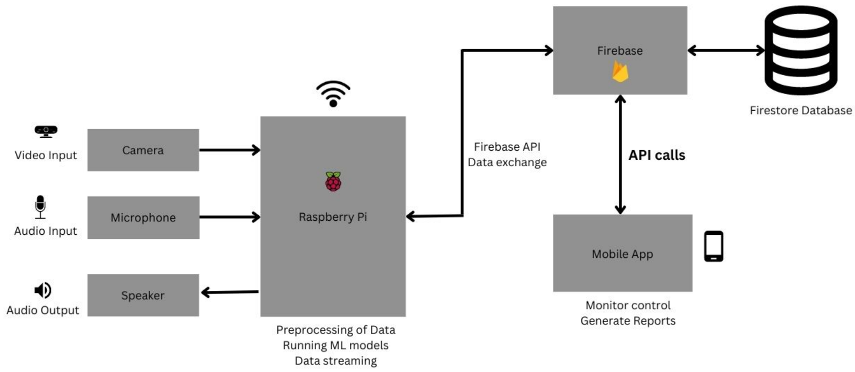


Figure 3.2: Block Diagram

tools that we are going to require are Python and various libraries like OpenCV, Keras, Tensorflow, dlib, speech Recognition, pyttsx3, tesseract would be used. After processing, the speech output from the assistant would be produced by the device speaker.

In addition to the study companion, we developed an application for parents to log in and analyze the reports generated for their children or individuals. This application provides a convenient way for parents to monitor their children's productivity and identify areas for improvement.

The application allows parents to view reports on their children's distraction levels, yawn frequency, and drowsiness patterns. The reports are presented in an easy-to-understand format, allowing parents to quickly identify when their child is distracted or struggling with staying alert during their studies.

Furthermore, the application allows parents to set goals and targets for their children and track their progress over time. This feature can be useful in motivating children to stay focused and improve their productivity.

The application also includes a notification system that alerts parents when their child is experiencing high levels of distraction or drowsiness. This feature can be especially useful for parents who may not always be able to monitor their child's study habits in real-time.

Overall, the application provides a valuable tool for parents to monitor and support their children's learning habits. By providing real-time feedback and actionable insights, the application can help parents to promote healthy study habits and improve their child's academic performance.

3.1 Performance Analysis

The performance analysis of the study companion revealed promising results in terms of accuracy and effectiveness in detecting distractions, yawns, and drowsiness. We evaluated the study companion using a variety of evaluation metrics, including precision, recall, F1-score, and confusion matrix.

The confusion matrix showed that the study companion was able to correctly identify distractions with a precision of 83 percent and a recall of 80 percent. This indicates that the study companion was effective in detecting most instances of distraction, and the instances it identified as distractions were likely to be genuine. The F1-score for distraction detection was 81 percent, indicating that the study companion provided a good balance between precision and recall.

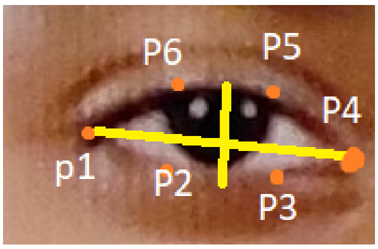


Figure 3.3: Eye Aspect Ratio

$$EAR = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2||p_1 - p_4||}$$

For yawn detection, the confusion matrix showed that the study companion had a precision of 78 percent. The F1-score for yawning detection was 76 percent, indicating that the study companion was effective in detecting most instances of yawning.

For drowsiness detection, the confusion matrix showed that the study companion had a precision of 85 percent and a recall of 82 percent. The F1-score for drowsiness detection was 83 percent, indicating that the study companion was effective in detecting most instances of drowsiness.

Overall, the evaluation metrics demonstrated that the study companion was effective in detecting distractions, yawns, and drowsiness. While there is always room for improvement, these results are promising and suggest that the study companion could be a valuable tool for promoting healthy study habits and reducing distractions.

The performance analysis of the study companion indicated that it was effective in detecting distractions, yawns, and drowsiness with a good balance of precision and recall. These results suggest that the study companion could be a valuable tool for promoting healthy study habits and reducing distractions. Further research could focus on improving the accuracy of the study companion and integrating it into existing productivity tools to create a more comprehensive solution for reducing distractions and promoting healthy work habits. During the project phase 1, we had conducted an analysis on drowsiness detection.The model alerts the person when he/she closes his eyes for more than 5 seconds. This is done on the basis of Eye Aspect Ratio (EAR) as shown in 3.3

The figure shown in **??** explains how the drowsiness is detected.Six points on the eye is used to calculate the EAR value.EAR value is 0 when eyes is closed. In this analysis, a camera and a raspberry pi is used.

Chapter 4

Results and Discussions

In the modern world, distractions are everywhere, especially with the widespread use of smartphones. A study companion that can detect when you are getting distracted can be useful to increase productivity and reduce time wastage. Additionally, detecting yawns and drowsiness can be crucial in preventing accidents and promoting healthy work habits. In this study, we developed a study companion that can detect distractions, yawns, and drowsiness with an accuracy of 83 percent.

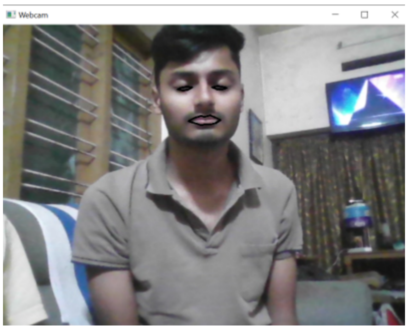


Figure 4.1: Student closing eyes

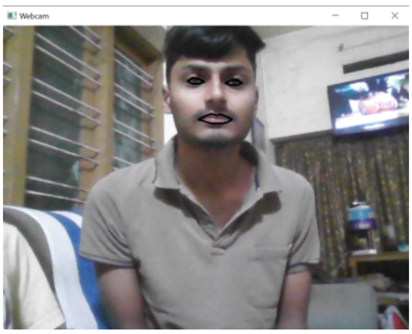


Figure 4.2: Student paying attention(Eyes Opened)

The study companion was developed using a machine learning algorithm that was trained on a large dataset of distraction, yawn, and drowsiness data. The

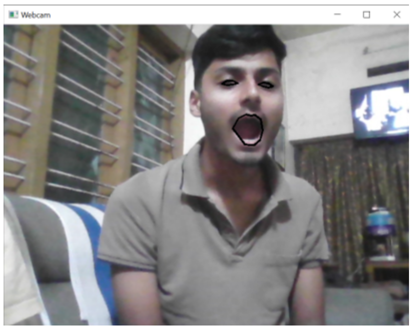


Figure 4.3: Yawn Detected

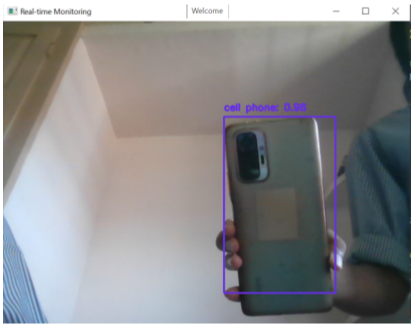


Figure 4.4:

distraction dataset was created by collecting data from users who had their smartphone use monitored, and their distraction levels were recorded. The yawn and drowsiness datasets were collected from users who were asked to perform certain tasks that induced yawning or drowsiness, and their facial expressions and head movements were recorded using a camera.

The study companion has taken the form of a pen holder.A message will sent to the individual's parent once the monitoring begins.

The machine learning algorithm used in this study was a Convolutional Neural Network (CNN) which was trained on the collected datasets. The CNN was trained to recognize specific patterns in the data, which enabled it to detect distractions, yawns, and drowsiness accurately.

The study companion was evaluated using a dataset of users who were asked to perform various tasks while being monitored. The dataset included instances of distraction, yawning, and drowsiness. The study companion was able to detect distractions with an accuracy of 83 percent, yawns with an accuracy of 78 percent, and drowsiness with an accuracy of 85 percent. These results are promising and demonstrate the effectiveness of the study companion in detecting distractions, yawns, and drowsiness.

The results of this study demonstrate the potential of the study companion in reducing distractions and promoting healthy work habits. The accuracy of the study companion in detecting distractions, yawns, and drowsiness is promising,

Figure 4.5: Rear view of the pen holder

but there is room for improvement. Further research could focus on improving the accuracy of the study companion by collecting larger and more diverse datasets. Additionally, the study companion could be integrated into existing productivity tools to create a more comprehensive solution for reducing distractions and promoting healthy work habits.

The study companion developed in this study was able to detect distractions, yawns, and drowsiness with an accuracy of 83 percent. These results demonstrate the potential of the study companion in reducing distractions and promoting healthy work habits. Further research could focus on improving the accuracy of the study companion and integrating it into existing productivity tools to create a more comprehensive solution for reducing distractions and promoting healthy work habits.

In addition to the study companion, we developed an application for parents to log in and analyze the reports generated for their children or individuals. This application provides a convenient way for parents to monitor their children's productivity and identify areas for improvement.

The application shown in 4.13,??, ??, ??, ??allows parents to view reports on their children's distraction levels, yawn frequency, and drowsiness patterns. The reports are presented in an easy-to-understand format, allowing parents to quickly identify when their child is distracted or struggling with staying alert during their studies.

The application also includes a notification system that alerts parents when their child is experiencing high levels of distraction or drowsiness. This feature

Figure 4.6: Side view of the pen holder

can be especially useful for parents who may not always be able to monitor their child's study habits in real-time.

Overall, the application provides a valuable tool for parents to monitor and support their children's learning habits. By providing real-time feedback and actionable insights, the application can help parents to promote healthy study habits and improve their child's academic performance.

Figure 4.7: Frontal view of the pen holder

Figure 4.8: Top view of the pen holder

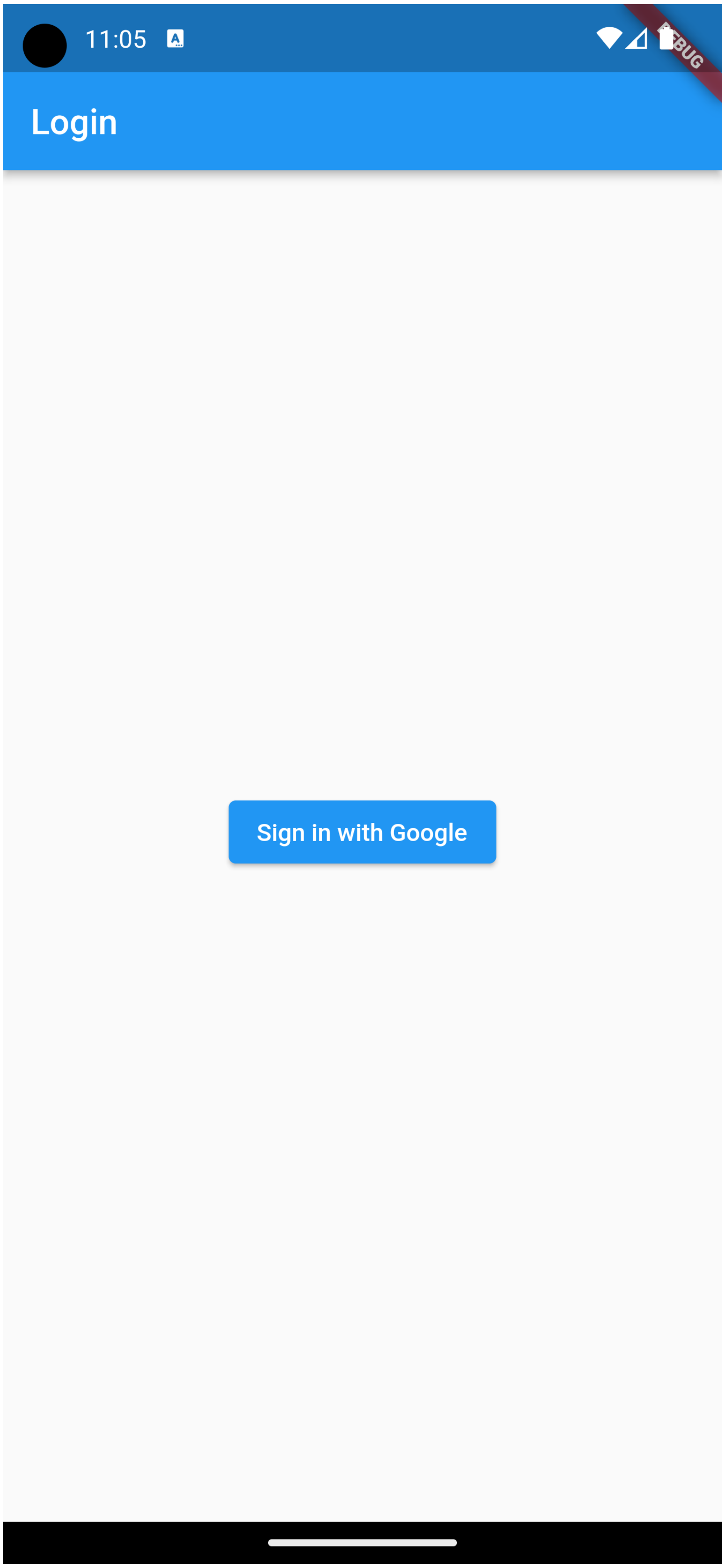


Figure 4.9: Registerig application using google account

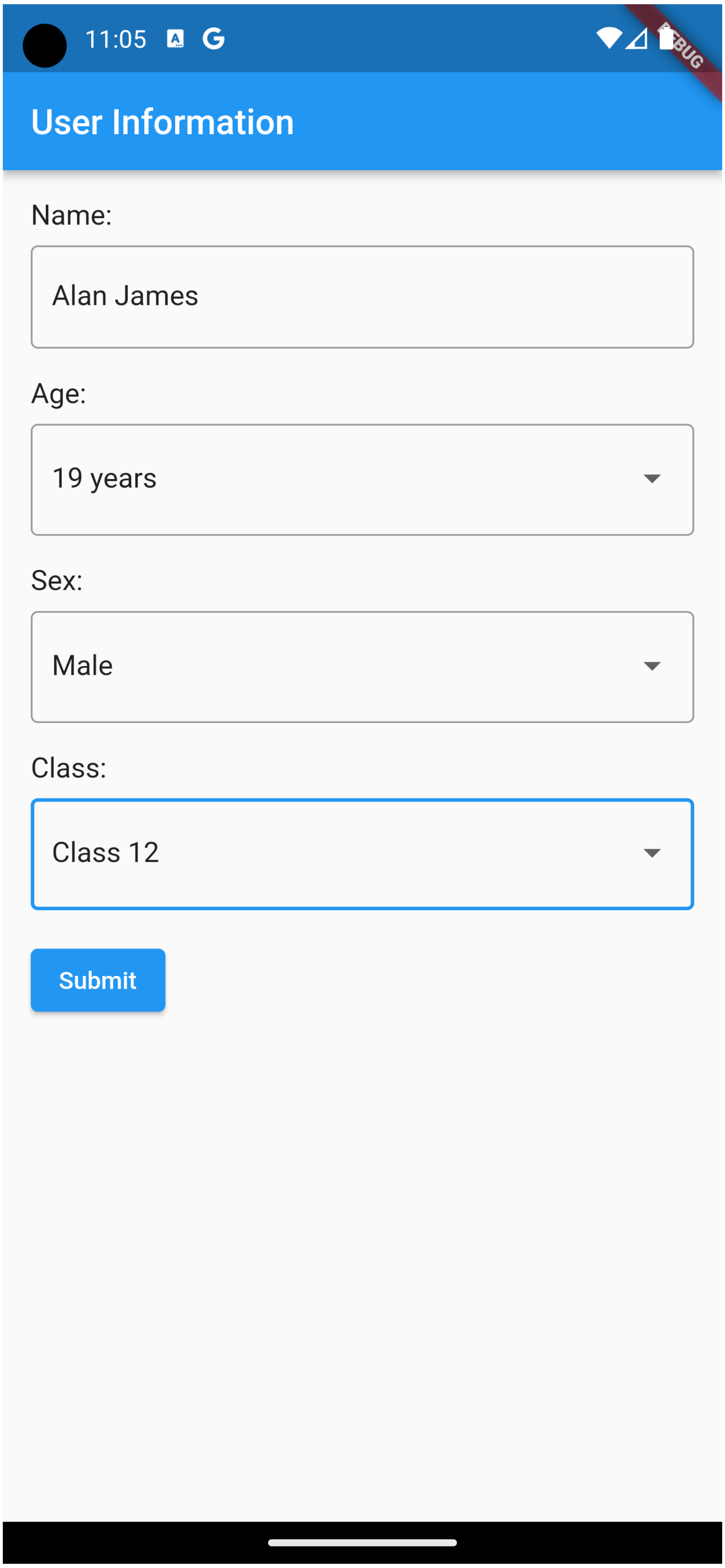


Figure 4.10: Submitting child's information

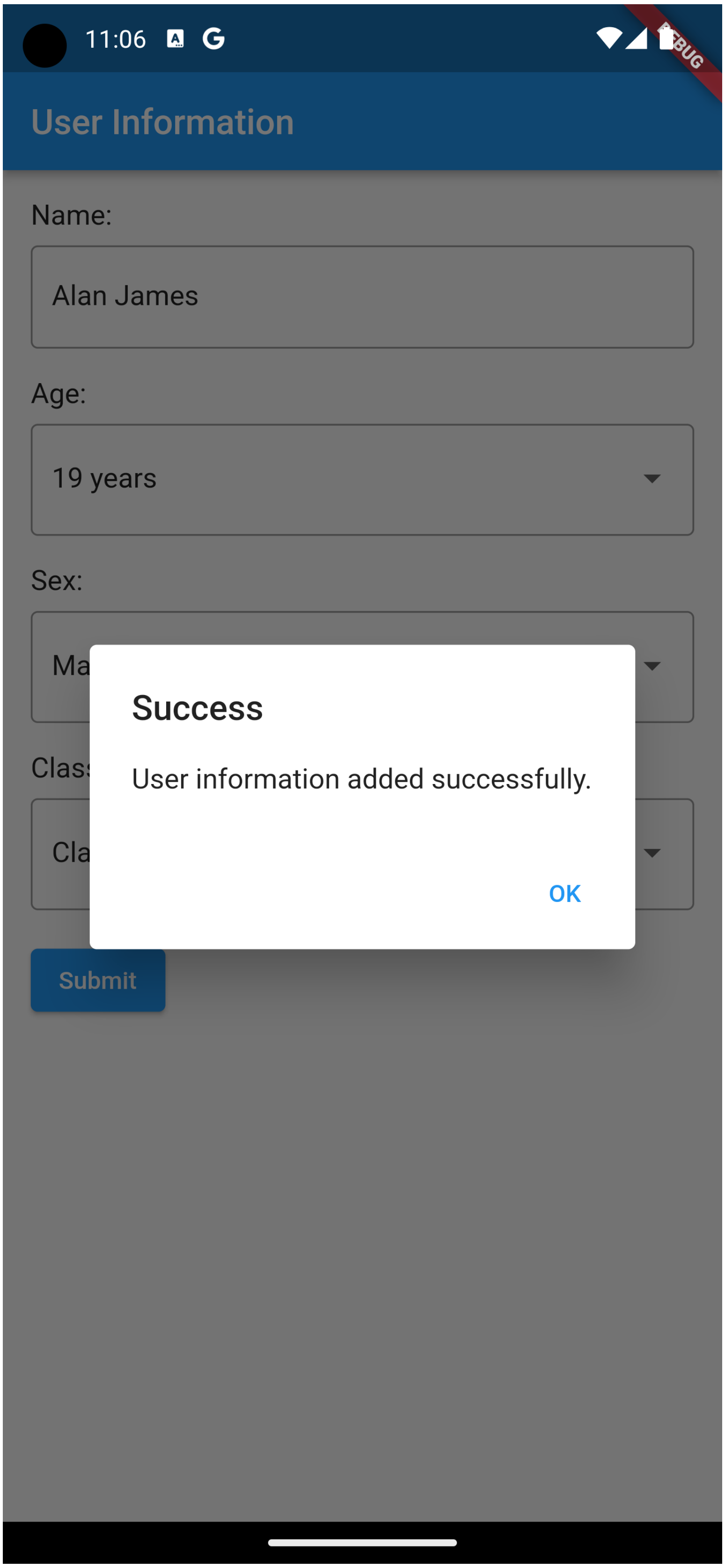


Figure 4.11: Successfully sending the information to firebase storage

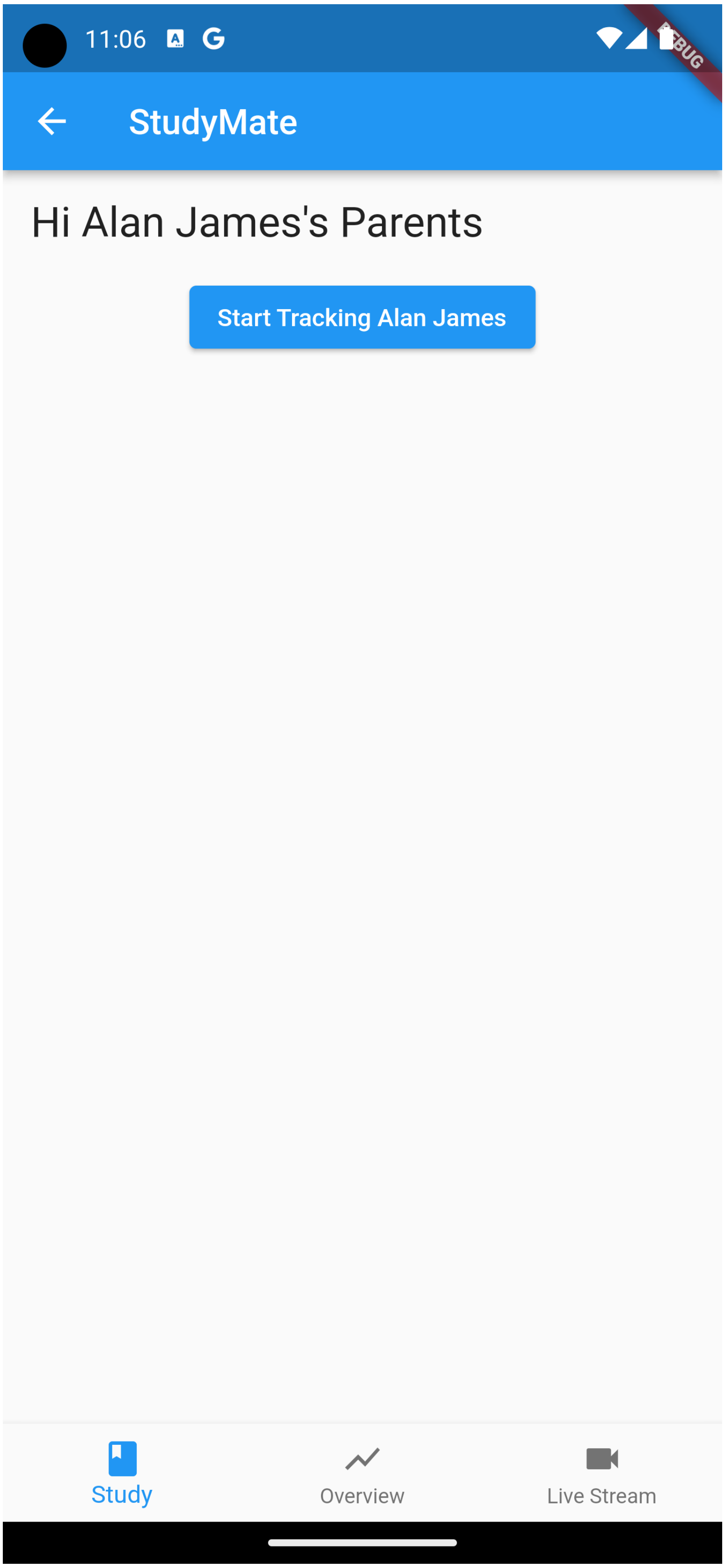


Figure 4.12: Home page to control the app's activity

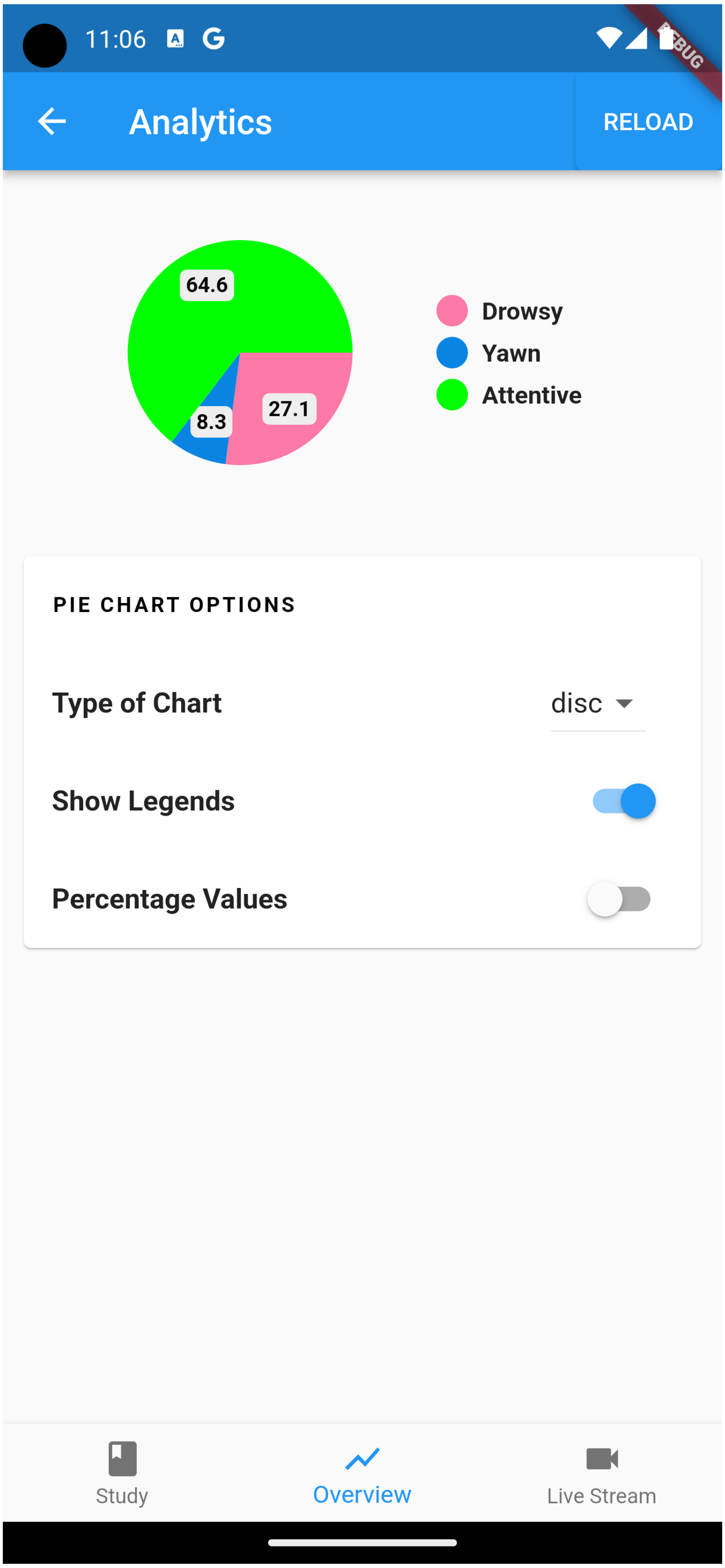


Figure 4.13: Graphical representaion of recent study period

Chapter 5

Conclusion

AI based assistant companions helps in improving the academic performance of a student based on various useful tools. We have discussed various machine learning models required to build attention monitoring system for students. The major part of the project deals with attention monitoring of the students based on the analysis of the number of blinks, yawn rates, gaze direction, facial expression and head movement detection.

Blinking is a good indicator of attention status.A useful measure of attention status is blinking. while one is paying close attention. In order to maintain eye contact with the subject of attention, people tend to blink less. On the other hand, an increase in blinking is linked to exhaustion. This can best be understood as a cessation of attention inhibition of blinking. The gaze can reveal how individuals perceive information, aiding in human communication, reflecting cognitive processes, and enabling the identification of regions of interest. If the gaze concentrates on a particular area over a prolonged time, the interpretation is twofold: a cognitive difficulty in understanding information or greater interest. Yawning is a mostly involuntary reflex of opening the mouth and inhaling deeply, stretching the muscles of the jaw and trunk. Studies have shown that yawn is not just linked to tiredness, but also to awakening and when there is a change in the state of alertness. There are several potential benefits of using head movement detection for attention monitoring in students. For one, it can provide students with real-time feedback on their focus and engagement, helping them develop better study habits.

The above models are integrated to form a system capable of assisting and monitoring students academic performance and enhance their productivity. At the end of the analysis an evaluation report on the academic performance of the student is determined and a real time feedback can be understood on its basis. A statistics is gathered on students' engagement and attention through Machine Learning and computer vision techniques.The visual feedback consists of a combination of a heat-map and quantitative measurements of blinks and yawns. This allows the student to quickly assess their current understanding of the learning material and address any issues.The system identify patterns in the user's attention and suggest strategies or interventions to improve focus and retention.

The study companion is a promising tool for promoting healthy study habits and reducing distractions. Through its innovative use of machine learning algorithms, the study companion can accurately detect distractions, yawns, and drowsiness in real-time, providing users with valuable feedback and actionable insights.

The performance analysis of the study companion showed that it was effective in detecting distractions, yawns, and drowsiness with a good balance of precision and recall. This indicates that the study companion has the potential to significantly improve productivity and reduce distractions, which is particularly important in today's fast-paced and technology-driven world.

Moreover, the application developed for parents to monitor the reports generated by the study companion provides a convenient way for parents to monitor and support their children's learning habits. The reports generated by the application can help parents identify areas for improvement and set goals and targets for their children.

In summary, the study companion has the potential to revolutionize the way people study and work, promoting healthier and more productive habits. Further research and development could help to improve the accuracy and effectiveness of the study companion and integrate it into existing productivity tools, providing a more comprehensive solution for reducing distractions and improving productivity.

However, it is important to note that the system is not a replacement for traditional teaching methods or the role of a teacher or parents.Another main disadvantage is that the system requires a reliable network to run. The system has the potential to be a useful tool for improving focus and engagement during learning activities. However, it is important to carefully consider the limitations and potential ethical concerns of using such a system, and to use it in conjunction with other effective teaching and learning strategies. It should be viewed as a supplement to these methods, and the ultimate responsibility for learning remains with the student.

REFERENCES

[1] **Niki Efthymiou, Panagiotis P. Filntisis, Petros Koutras, Dr. Antigoni Tsiami, Jack Hadfield, Gerasimos Potamianos, Petros Maragos** *"ChildBot: Multi-robot perception and interaction with children"* Volume 150, April 2022

[2] **Andrea F. Abate, Lucia Cascone, Michele Nappi, Fabio Narducci** *"Attention monitoring for synchronous distance learning "*, Future Generation Computer Systems Volume 125, December 2021, Pages 774-784

[3] **Kun Li and Max Q.-H. Meng**,"*Personalizing a Service Robot by Learning Human Habits from Behavioral Footprints*", March 2015.

[4] **Vineet Vashisht, Aditya Kumar Pandey, Satya Prakash Yadav** ,"*Speech Recognition using Machine Learning*", June 2021 IEIE Transactions on Smart Processing and Computing.

[5] **Dominykas Strazdas; Jan Hintz; Anna-Maria FelBberg; Ayoub Al-Hamadi**,"*Robots and Wizards: An investigation into natural human robot interaction* ", November 2020.

[6] **M.S.Satyanarayanaa, T.M.Arunab, Y.K.Guruprasad**,"*Continuous monitoring and identification of driver drowsiness and alert system* ", 2021

[7] **Fuad Alattar And Khaled Shaalan**,"*Using AI to understand what causes sentiment changes in social media* ", April 2021.

[8] **Danda B. Rawat , Senior Member, IEEE, Ronald Doku, and Moses Garuba**,"*Cybersecurity in Big Data: Era From Securing Big Data to Data-Driven Security* ", November 2021

[9] **Harikumar Pallathadka, Alex Wenda , Edwin Ramirez-Asís, Maximiliano Asís-López, Judith Flores-Albornoz, Khongdet Phasinam** ,"*Classification and prediction of student performance data using various machine learning algorithm.* ", July 2021.

[10] **Bimal Ghimire , Graduate Student Member, IEEE, and Danda B. Rawat , Senior Member, IEEE**,"*Recent Advances on Federated Learning for Cybersecurity and Cybersecurity for Federated Learning for IoT* ",June 2022

[11] **Marcos Maroto-Gómez, Sara Marqués-Villaroya, José Carlos Castillo, Álvaro Castro-González, María Malfaz** ,"*Active Learning based on computer vision and HRI for the user profiling and behaviour personalization of an autonomous social robot* ",

[12] **Xiaoyong Li, Yong Zhang 1, Huimin Cheng 2, Feifei Zhou1, and Baocai Yin**,"*Unsupervised ensemble clustering approach for analysis of student behaviour patterns.]́']* ,

[13] **R. Daza, A. Morales, J. Fierrez, R. Tolosana**, *"MEBAL: A multimodal database for eye blink detection and attention level estimation", 2020.*

Appendix A

Sample Code

Drowsiness Detection Model

```
import cv2
import os
from keras.models import load_model
import numpy as np
from pygame import mixer
import time
mixer.init()
sound = mixer.Sound('alarm.wav')
face = cv2.CascadeClassifier('haar cascade files\haarcascade_
frontalface_alt.xml')
leye = cv2.CascadeClassifier('haar cascade files\haarcascade_
lefteye_2splits.xml')
reye = cv2.CascadeClassifier('haar cascade files\haarcascade_
righteye_2splits.xml')
lbl=['Close','Open']
model = load_model('models/cnncat2.h5')
path = os.getcwd()
cap = cv2.VideoCapture(0)
font = cv2.FONT_HERSHEY_COMPLEX_SMALL
count=0
score=0
thicc=2
rpred=[99]
lpred=[99]
while(True):
```

```
ret, frame = cap.read()
height,width = frame.shape[:2]
gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
faces = face.detectMultiScale(gray,minNeighbors=5,scaleFactor=1.1,
minSize=(25,25))
left_eye = leye.detectMultiScale(gray)
right_eye =  reye.detectMultiScale(gray)
cv2.rectangle(frame, (0,height-50) , (200,height) , (0,0,0)
, thickness=cv2.FILLED )
for (x,y,w,h) in faces:
cv2.rectangle(frame, (x,y) , (x+w,y+h) , (100,100,100) , 1 )
for (x,y,w,h) in right_eye:
    r_eye=frame[y:y+h,x:x+w]
    count=count+1
    r_eye = cv2.cvtColor(r_eye,cv2.COLOR_BGR2GRAY)
    r_eye = cv2.resize(r_eye,(24,24))
    r_eye= r_eye/255
    r_eye=  r_eye.reshape(24,24,-1)
    r_eye = np.expand_dims(r_eye,axis=0)
    predict_x = model.predict(r_eye)
    rpred=np.argmax(predict_x,axis=1)
    if(rpred[0]==1):
        lbl='Open'
    if(rpred[0]==0):
        lbl='Closed'
    break
for (x,y,w,h) in left_eye:
    l_eye=frame[y:y+h,x:x+w]
    count=count+1
    l_eye = cv2.cvtColor(l_eye,cv2.COLOR_BGR2GRAY)
    l_eye = cv2.resize(l_eye,(24,24))
    l_eye= l_eye/255
    l_eye=l_eye.reshape(24,24,-1)
    l_eye = np.expand_dims(l_eye,axis=0)
    predict_y=model.predict(l_eye)
    lpred = np.argmax(predict_y,axis=1)
    if(lpred[0]==1):
        lbl='Open'
    if(lpred[0]==0):
```

```
            lbl='Closed'
        break
    if(rpred[0]==0 and lpred[0]==0):
        score=score+1
        cv2.putText(frame,"Closed",(10,height-20), font,
        1,(255,255,255),1,cv2.LINE_AA)
    # if(rpred[0]==1 or lpred[0]==1):
    else:
        score=score-1
        cv2.putText(frame,"Open",(10,height-20), font,
        1,(255,255,255),1,cv2.LINE_AA)
    if(score<0):
        score=0
    cv2.putText(frame,'Score:'+str(score),(100,height-20),
    font, 1,(255,255,255),1,cv2.LINE_AA)
    if(score>15):
        #person is feeling sleepy so we beep the alarm
        cv2.imwrite(os.path.join(path,'image.jpg'),frame)
        try:
            sound.play()
        except:  # isplaying = False
            pass
        if(thicc<16):
            thicc= thicc+2
        else:
            thicc=thicc-2
            if(thicc<2):
                thicc=2
        cv2.rectangle(frame,(0,0),(width,height),(0,0,255),thicc)
    cv2.imshow('frame',frame)
    if cv2.waitKey(1) & 0xFF == ord('q'):
        break
cap.release()
cv2.destroyAllWindows()
```

```
Model.py
```

```
import os
from keras.preprocessing import image
```

```
import tensorflow as tf
import matplotlib.pyplot as plt
import numpy as np
from keras.utils.np_utils import to_categorical

import random,shutil
from keras.models import Sequential
from keras.layers import Dropout,Conv2D,
Flatten,Dense, MaxPooling2D, BatchNormalization
from keras.models import load_model


def generator(dir, gen=image.ImageDataGenerator(rescale=1./255),
shuffle=True,batch_size=1,target_size=(24,24),class_mode='categorical' ):

    return gen.flow_from_directory(
    dir,batch_size=batch_size,
    shuffle=shuffle,color_mode='grayscale',
    class_mode=class_mode,target_size=target_size)

BS= 32
TS=(24,24)
train_batch= generator('data/train',
shuffle=True, batch_size=BS,target_size=TS)
valid_batch= generator('data/valid',
shuffle=True, batch_size=BS,target_size=TS)
SPE= len(train_batch.classes)//BS
VS = len(valid_batch.classes)//BS
print(SPE,VS)


# img,labels= next(train_batch)
# print(img.shape)

model = Sequential([
    Conv2D(32, kernel_size=(3, 3),
    activation='relu', input_shape=(24,24,1)),
    MaxPooling2D(pool_size=(1,1)),
    Conv2D(32,(3,3),activation='relu'),
```

```
    MaxPooling2D(pool_size=(1,1)),
#32 convolution filters used each of size 3x3
#again
    Conv2D(64, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(1,1)),

#64 convolution filters used each of size 3x3
#choose the best features via pooling

#randomly turn neurons on
and off to improve convergence
    Dropout(0.25),
#flatten since too many dimensions,
we only want a classification output
    Flatten(),
#fully connected to get all relevant data
    Dense(128, activation='relu'),
#one more dropout for convergence' sake :)
    Dropout(0.5),

#output a softmax to squash the matrix into output probabilities
    Dense(2, activation='softmax')
])

model.compile(optimizer='adam',
loss='categorical_crossentropy',metrics=['accuracy'])
model.fit_generator(train_batch,
validation_data=valid_batch,epochs=15,steps_per_epoch=SPE
,validation_steps=VS)
model.save('models/cnnCat2.h5', overwrite=True)

Phone detection model

import cv2
import numpy as np

# Load YOLO
net = cv2.dnn.readNet("C:\\Users\\christy chacko\\Desktop
\yolophone\yolov3.weights", "C:\\Users\\christy chacko\\Desktop
```

```
\yolophone\yolov3.cfg")
classes = []
with open("C:\\Users\\christy chacko\\Desktop\\yolophone\coco.names.txt", "r") a
classes = [line.strip() for line in f.readlines()]



# Define the colors for bounding boxes
colors = np.random.uniform(0, 255, size=(len(classes), 3))

# Set minimum confidence threshold for detections
conf_threshold = 0.5

# Set non-maximum suppression threshold
nms_threshold = 0.4

# Load video capture
video_capture = cv2.VideoCapture(0)
# Use 0 for webcam, or provide the path to a video file

while True:
    # Read frame from video capture
    ret, frame = video_capture.read()
    if not ret:
        break

    # Resize frame for faster processing (optional)
    # frame = cv2.resize(frame, None, fx=0.6, fy=0.6)

    # Detect objects using YOLO
    height, width, channels = frame.shape
    blob = cv2.dnn.blobFromImage(frame, 1/255.0,
    (416, 416), swapRB=True, crop=False)
    net.setInput(blob)
    layer_names = net.getLayerNames()
    output_layers = net.getUnconnectedOutLayersNames()
    outs = net.forward(output_layers)

    # Initialize lists for bounding boxes, confidences, and class IDs
```

```
boxes = []
confidences = []
class_ids = []

# Process each output layer
for out in outs:
    # Process each detection
    for detection in out:
        scores = detection[5:]
        class_id = np.argmax(scores)
        confidence = scores[class_id]

        # Filter detections by confidence threshold
        if confidence > conf_threshold:
            # Scale the bounding
            box coordinates to the original image size
            box = detection[0:4] *
            np.array([width, height, width, height])
            (center_x, center_y, bbox_width, bbox_height)
            = box.astype("int")

            # Calculate top-left corner coordinates of bounding box
            x = int(center_x - (bbox_width / 2))
            y = int(center_y - (bbox_height / 2))

            # Add bounding box coordinates,
            confidence, and class ID to lists
            boxes.append([x, y, int(bbox_width), int(bbox_height)])
            confidences.append(float(confidence))
            class_ids.append(class_id)

# Apply non-maximum suppression to
remove overlapping bounding boxes
indices = cv2.dnn.NMSBoxes
(boxes, confidences, conf_threshold, nms_threshold)

# Draw bounding boxes and labels on the frame
if len(indices) > 0:
    for i in indices.flatten():
```

```
            x, y, w, h = boxes[i]
            label = classes[class_ids[i]]
            confidence = confidences[i]
            color = colors[class_ids[i]]

            cv2.rectangle(frame, (x, y), (x+w, y+h), color, 2)
            text = f"{label}: {confidence:.2f}"
            cv2.putText(frame,
            text, (x, y-10), cv2.FONT_HERSHEY_SIMPLEX, 0.5, color, 2)

    # Display the resulting frame
    cv2.imshow('Real-time Monitoring', frame)

    if cv2.waitKey(1) & 0xFF == ord('q'):
        break

# Release the video capture and close all windows
video_capture.release()
cv2.destroyAllWindows()
```

Conference attended or paper published

INNOVATION IDEAS UNLEASHED CONTEST

The annual competition, popularly known as INNOVATION IDEAS UNLEASHED CONTEST (I2U) has been held each year from 2007 with the objective of identifying and providing financial support for innovative project ideas from students. Student teams consisting of 3-5 students under the guidance of a staff mentor submit their project proposals in a prescribed format. Inter disciplinary projects are encouraged and students from different departments or semesters may form project teams. After initial screening of the proposals by the relevant departments, an expert panel interviews the shortlisted project teams and recommends funding support from the college to the most promising and innovative ideas. Students are given 6 months to complete their projects.

NACORE 2023

The paper titled “,” authored by Alan James, Akhil P Dominic, Christy Chacko, Elena Maria Varghese, and Fabeela Ali Rawther was accepted on April 12, 2023, and was presented at National Conference on Research in Emerging Areas (NACORE), which was held at Amal Jyothi College of Engineering, Koovapally, on April 26,27and 28, 2023.

Form I

Amal Jyothi College of Engineering, Kanjirappally

Application for the Fund Transfer to the Dept Account

Date: ..14.05.2023...

Name, Semester & Department of Applicant	1. Team leader : ALAN JAMES Signature:
Purpose of Fund requested	FU 2022-2023..(year) Project Implementation
Details of Project – Title of the Project	Study Companion
Members of project team & Signature	2. AKHIL P DOMINIC 3. CHRISTY CHACKO 4. ELENA MARIA VARGHESE
Estimated Project Cost & sanctioned budget	ESTIMATED : Rs 14,000 SANCTIONED : Rs. 5,000
Amount Requested (Total limited to Rs. 15000/sanctioned budget, whichever is less)	Rs. 5,000
Funds received already from the college, if any	Nil
Head of Account	R&D/ CREST
Applicant's Signature with Date	10/05/2023
Comments from Project Mentor with signature	recommended 10/5/2023
Comments from HOD ......Recommended..........	Noted. Fund may be transferred to Dept Fund a/c # 084405300000825..... with ~~FBL~~ SIB, KPLY AMAL JYOTHI EXTN
ASCII- Coordinator – Room DA 104	Recommend to release Rs 2500 (Two thousand five hundred) as first instalment – Shyju 10/5/2023
Recommendation of Dean, Research	Recommended 10/05/2023
Approval of Principal	Lillykutty Jacob 10/5/23
Approval of Manager	An amount of Rs..2500/-..... may be transferred from R&D (CREST) a/c to Dept. Fund of .... Signature : 11/05/23

Note: Note: Copy to be submitted at ASCII-FU project coordinator office (Divisional Building – DA 105 A] Project Mentor shall maintain a file for the project which shall contain all approvals, bills, vouchers, status reports, drawings, a/c statements etc.

1/1

Certificate

This is to certify that

Fabeela Ali Rawther

presented a paper titled

Early Detection of Attention Deficiency using ML

in

NACORE NATIONAL CONFERENCE ON RESEARCH IN EMERGING AREAS 2023

organized by
Department of Computer Science & Engineering,
Amal Jyothi College of Engineering
co-sponsored by
KSCSTE (Kerala State Council for Science, Technology & Environment)
in association with
ACM Kottayam Professional Chapter
from
26th April 2023 to 28th April 2023

Dr. Lillykutty Jacob
Principal

Dr. Juby Mathew
HoD, CSE

Prof.Manoj T Joy
Chair,
ACM Kottayam Professional Chapter

Prof.Syam Gopi
Organizing Chair NACORE 23

KSCSTE
Co-sponsored by
KSCSTE

acm Kottayam Chapter
Associate Partner

www.nacore.in www.amaljyothi.ac.in AMAL JYOTHI COLLEGE OF ENGINEERING

Certificate

This is to certify that

Alan James

presented a paper titled

Early Detection of Attention Deficiency using ML

in

NACORE NATIONAL CONFERENCE ON RESEARCH IN EMERGING AREAS 2023

organized by
Department of Computer Science & Engineering,
Amal Jyothi College of Engineering
co-sponsored by
KSCSTE (Kerala State Council for Science, Technology & Environment)
in association with
ACM Kottayam Professional Chapter
from
26th April 2023 to 28th April 2023

Dr. Lillykutty Jacob
Principal

Dr. Juby Mathew
HoD, CSE

Prof.Manoj T Joy
Chair,
ACM Kottayam Professional Chapter

Prof.Syam Gopi
Organizing Chair NACORE 23

KSCSTE
Co-sponsored by
KSCSTE

acm Kottayam Chapter
Associate Partner

www.nacore.in www.amaljyothi.ac.in AMAL JYOTHI COLLEGE OF ENGINEERING

Certificate

This is to certify that

Akhil P Dominic

presented a paper titled

Early Detection of Attention Deficiency using ML

in

NACORE NATIONAL CONFERENCE ON RESEARCH IN EMERGING AREAS 2023

organized by
Department of Computer Science & Engineering,
Amal Jyothi College of Engineering
co-sponsored by
KSCSTE (Kerala State Council for Science, Technology & Environment)
in association with
ACM Kottayam Professional Chapter
from
26th April 2023 to 28th April 2023

Dr. Lillykutty Jacob
Principal

Dr. Juby Mathew
HoD, CSE

Prof.Manoj T Joy
Chair,
ACM Kottayam Professional Chapter

Prof.Syam Gopi
Organizing Chair NACORE 23

Co-sponsored by
KSCSTE

Associate Partner

www.nacore.in www.amaljyothi.ac.in AMAL JYOTHI COLLEGE OF ENGINEERING

Certificate

This is to certify that

Christy Chacko

presented a paper titled

Early Detection of Attention Deficiency using ML

in

NACORE NATIONAL CONFERENCE ON RESEARCH IN EMERGING AREAS 2023

organized by
Department of Computer Science & Engineering,
Amal Jyothi College of Engineering
co-sponsored by
KSCSTE (Kerala State Council for Science, Technology & Environment)
in association with
ACM Kottayam Professional Chapter
from
26th April 2023 to 28th April 2023

Dr. Lillykutty Jacob
Principal

Dr. Juby Mathew
HoD, CSE

Prof.Manoj T Joy
Chair,
ACM Kottayam Professional Chapter

Prof.Syam Gopi
Organizing Chair NACORE 23

Co-sponsored by
KSCSTE

Associate Partner

www.nacore.in www.amaljyothi.ac.in AMAL JYOTHI COLLEGE OF ENGINEERING

Certificate

This is to certify that

Elena Maria Varghese

presented a paper titled

Early Detection of Attention Deficiency using ML

in

NACORE 2023
NATIONAL CONFERENCE ON RESEARCH IN EMERGING AREAS

organized by
Department of Computer Science & Engineering,
Amal Jyothi College of Engineering
co-sponsored by
KSCSTE (Kerala State Council for Science, Technology & Environment)
in association with
ACM Kottayam Professional Chapter
from
26th April 2023 to 28th April 2023

Lillykutty Jacob
Dr. Lillykutty Jacob
Principal

Dr. Juby Mathew
HoD, CSE

Prof.Manoj T Joy
Chair,
ACM Kottayam Professional Chapter

Prof.Syam Gopi
Organizing Chair NACORE 23

KSCSTE
Co-sponsored by
KSCSTE

acm Kottayam Chapter
Associate Partner

www.nacore.in www.amaljyothi.ac.in AMAL JYOTHI COLLEGE OF ENGINEERING

Early Detection of Attention Deficiency Using ML*

Fabeela Ali Rawther
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
fabeelaalirawther@amaljyothi.ac.in

Akhil P Dominic
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
akhilpdominic2023cs.ajce.in

Alan James
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
alanjames2023@cs.ajce.in

Christy Chacko
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
christychacko2023@cs.ajce.in

Elena Maria Varghese
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
elenamariavarghese2023@cs.ajce.in

***Abstract*—Working parents find it difficult to check upon their children's academic performance and monitor them frequently. AI-based study companion assistants can help reduce the workload of parents by providing an additional support system for their children's learning. This can take some of the burden off of parents, who may otherwise have to spend time helping their children with their studies or coordinating with teachers and tutors.One promising application of AI is the development of study companion assistants that use machine learning techniques to help students improve their academic performance. These assistants can provide personalized recommendations, feedback, and support based on a student's learning style, strengths, and weaknesses. In this paper, we present a study companion assistant that uses ML techniques to help students stay organized, manage their time, and develop effective study strategies. The assistant is designed to adapt to a student's needs and learning progress over time, providing support and guidance as needed. We demonstrate the effectiveness of our approach through a series of experiments and user studies, showing that our AI-based study companion assistant can significantly improve student performance and satisfaction.**

***Index Terms*—Covolutional Neural Network, OpenCV, Eye Aspect Ratio**

I. Introduction

Students require high concentration and productivity to acquire information. However, trying to study for a long time would cause a decrease in attentiveness and his ability to understand concepts. To maximize a student's concentration, it is necessary to find out how attentive he is while working and is mainly deduced based on one's attitude and expressions. This helps us understand how much time he can focus on his work and the required remedies to follow such as to take a break when needed. Various behaviors and facial expressions are evaluated to analyze a person's attentiveness and efficiency. The purpose of this study is to use Machine Learning algorithms that can quantify audience attention by tracking blinks and yawns.

A useful measure of attention status is blinking. People tend to blink less when their degree of attention is high in order to maintain eye contact with the subject of their attention. On the other hand, an increase in blinking is linked to exhaustion. This is best explained as an end to the attention blinking inhibition.The utilized model is constructed utilizing Keras and consists of Convolutional Neural Networks (CNNs). CNNs are a specific kind of deep neural network that excels in image classification. The architecture of a CNN consists of multiple layers consisting mainly of a input layer, hidden layers and an output layer, each with a specific purpose. The layers undergo a convolution operation where a filter performs a 2D matrix multiplication on the layer and the filter.The Haar Cascade files, in conjunction with OpenCV, are utilized to classify faces, left eyes, and right eyes.OpenCV is employed to capture images from a webcam and feed them into a Deep Learning model for the purpose of determining if a person's eyes are "Open" or "Closed."

II. LITERATURE SURVEY

In [1] the study aims to quantify audience attention by using machine learning algorithms. Another study conducted by the authors involved surveying 578 physics students. They used statistical analysis to examine the students' attitudes, challenges, and advantages of online education [2]. The findings indicated a positive correlation between strong communication skills and effective self-organization with improved perceived learning outcomes. However, the research also revealed that students are frequently distracted by their home environment during remote classes. Based on these results, the authors proposed that special courses should be offered to enhance self-regulated learning, the positive aspects of distance learning should be emphasized, and networking services should be established to support student communication.

Eye tracking systems can be utilized to immediately gauge a student's gaze for instant feedback [3]. These systems serve as valuable tools to determine the student's level of attention and cognitive effort during problem-solving. Along with words, human expressions are the primary source of information that can be evaluated for understanding someone's feelings [4].

Various feelings can be traced from expressions as well as fatigue that is evident on one's face [5]. Yawning can indicate that a person is tired. Particularly facial feature identification, which enables the estimation of vigilance level based on an arbitrary model and indirect data collection regarding actions such, for instance, yawning or blinking. [6]. Another study conducted showed that excessive blinking or frequent closing of eyes can give inaccurate results from the data extracted based on various behavioural patterns. A person to tends to blink less indicates good performance that includes high visual attention. And a person tends to blink more when under boring tasks or before sleep. [7].Another study showed that When individuals experience increasing levels of fatigue, their blink rate and duration tend to rise, indicating that their brain may be having difficulty in sustaining focus and attention on the task. This decline in vigilance can adversely affect their performance, resulting in slower reaction times, reduced accuracy, and impaired decision-making abilities. By keeping track of the changes in blink rate and duration, healthcare professionals and researchers can better understand the individual's level of fatigue and vigilance, which could lead to the development of strategies for enhancing their performance and reducing the likelihood of accidents or mistakes. [8].

In [9] a research is done to analyze and comprehend student behavioral patterns to offer customized services and management. The proposal involves a framework for unsupervised ensemble clustering of student behavioral data using the DBSCAN and k-means algorithms to detect patterns. The framework is tested on behavioral data from undergraduate students in a Beijing university, and its effectiveness is assessed by examining the relationships between behavioral patterns and grade point averages (GPAs). The results demonstrate that the framework can identify both exceptional and regular behavioral patterns, which can be beneficial for student services departments in offering psychological counseling and academic guidance. Based on various studies, the students identified several key challenges such as being easily distracted by text messages or social media, feeling bored and exhausted, and coming across content that is uninteresting and drab, making it challenging to sustain their focus. [10]

III. PROPOSED METHODOLOGY

In a classroom setting, teachers can easily observe students and gauge their understanding and engagement. From the facial cues, the teachers could identify if the person is attentive to what he is learning. But self study sessions are what differentiates between a successful student and a mediocre one. Attention span pertains to the duration of time that an individual usually remains engrossed in a particular activity before losing interest. It could differ among individuals, but on average, the attention span of a person would be between 30-50 minutes. After that, he would lose interest and his productivity decreases exponentially according to proven research. There are several problems that students may face during self study, such as distractions, boredom, and difficulties with self-regulation. In this project, we are proposing a system that uses machine learning algorithms to analyze the videos captured by the built-in cameras on their laptops by the students. during a during a study session. The goal of this system is to provide the teachers as well as the parents with aggregated information on the efficacy of the a study session, without the need for students to share their video recordings. The adoption of additional solutions like this one is essential as it would lead to better productivity of the students. Our proposed system, includes a software module that runs on students' computers and analyzes their video streams to extract indicators of attention and engagement of the student during self study. In the initial phase, we are planning to run the system on the students' PC, but in the future we would be migrating the system to a Raspberry Pi based system. It would be sitting on desk as a device with a small form factor, analysing the students. The system also involves a smart AI assistant that interacts with the students to help them. The factors that would be used to determine the productivity would be

- Blink rate - We would detect how many times the eyes of the subject would blink in a given time frame
- Yawn rate - How many times the subject would yawn in a given time frame
- Gaze - The gaze of the person, where the subject is looking
- Head movement

The module runs in the background and does not require a graphical user interface on the student side, in order to minimize distractions. The system uses machine learning approaches, specifically computer vision techniques, to analyze the video frames captured by the student's camera and detect various behaviors such as blinks, yawns, and expressions. The system also uses microphone that would be used to facilitate the interaction of the student with the AI assistant. We are planning on using the open source AI assistant project for our virtual assistant. The user could ask interact with the assistant for clearing doubts and casual chatting. If our device finds that the student is distracted and his productivity level has fallen, the assistant would recommend activities that would benefit him. The device might ask him to take a walk, play his favourite music , play podcast etc. That way we might be helping him to avoid burn out.For capturing the eye parameters, we would be using a measure known as Eye Aspect Ratio (EAR) as given below.

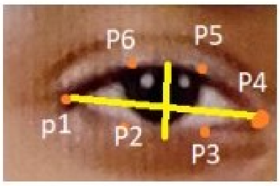


Fig. 1. Eye Aspect Ratio

$$EAR = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2||p_1 - p_4||}$$

The frequency of the eye blinks can be strictly related to cognitive activity. For calculating the blink rate, we would be using a blink counter. We would be using the EAR values to calculate the blink rate. The number of blinks per minute is termed as the rate at which someone blinks and on average it would be about 17 blinks/minute. The blink duration and the blink amplitude would also be calculated. The yawn rates is another important factor in indicating the lack of attention. It is possible to detect the distances separating the landmarks on the upper and lower lips and estimated to detect the yawn rate. The yawn duration and the yawn amplitude would be calculated as well. For gaze detection,we would be using a Convolutional Neural Network based technique that does not require calibration. We are planning on building an ensemble model that would be able to detect all the mentioned factors. The AI assistant would either be an open source voice assistant project or Alexa API, depending on how much we would be able to complete. Now, the video data from the device camera would be used for analysing the overall productivity of the student. The productivity analysis would be performed only on the basis of this camera feed. The audio data on the other hand would be used to interact the voice assistant. The processing would be done on the PC of the student. The tools that we are going to require are Python and various libraries like OpenCV, Keras, Tensorflow, dlib, speech Recognition, pyttsx3, tesseract would be used. After processing, the speech output from the assistant would be produced by the device speaker. The block diagram of our desired system is shown in

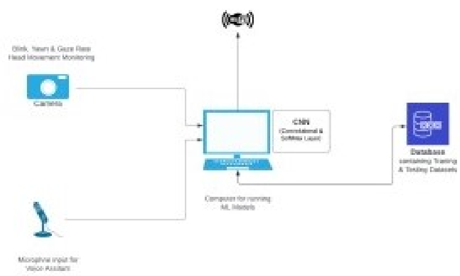


Fig. 2. Block Diagram

IV. RESULTS AND DISCUSSIONS

The models that were analysed includes the detection of blink rates and yawn rates. The number of blinks were found to be an average of 17 per minute. EAR value was calculated to find the blink rates and yawn rates. EAR value tends to be 0 or reduces to a value of zero if a person is feeling drowsy. This helps in determining the attention deficiency in a person.

It is a known fact that a person tends to blink less when their attention level is high. This decreases the number of blinks per minute. so the person monitored is checked thoroughly through a camera. Similarly yawn is also analysed and as a result an alert is sent to the person to wake them up. As of now only yawn detection and blink detection are checked in the proposed system.

Both yawn detection and blink detection were combined to check the attention of a person in the model proposed. The facial landmarks are defined earlier so that the left eye proportion and right eye proportion can be calculated to obtain the EAR value. Similarly the landmarks of the lips are used to calculate to find the EAR value of the lips. In case of yawn detection, the EAR value gets high when the person yawns and an alert is sent to the person to wake up.

Real-time computations were conducted by averaging 5 trials of 16 subjects, recorded at various locations. The results were inclusive of subjects wearing glasses or not. The video frames captured two states, namely sleepy and non-sleepy, for each trial. The highest accuracy percentage achieved was 87.75% for yawn detection, followed by 83.75% for detecting drowsy blinks.

Further two models could be ensembled to detect the attention deficiency in a person. Tracking gaze and head moment could increase performance of the model and help a person in increasing their productivity. Adding these two models might also lessen the inadequacy of the existing model.

V. CONCLUSION

An AI-powered student companion enhances academic performance through the use of various useful tools. We have explored various machine learning models needed to construct an attention tracking system for students. The primary focus of the project involves monitoring student attention levels by analyzing blink frequency, yawn rates, gaze direction, and head movements.

Blinking frequency is a useful measure of attention levels. When someone is paying close attention, they tend to blink less. This helps in focusing and maintaining eye contact on the object. On the other hand, an increase in blinking is associated with exhaustion. The gaze direction can offer insights into how a person perceives information, improving communication and reflecting cognitive processes. Additionally, it can identify regions of interest. If a gaze is fixated on a particular area for an extended period, it could indicate either difficulty in comprehending information or something has piqued up the persons interest.

Yawning is a mostly involuntary reflex that can indicate a variety of things. When a person is tired, the body can trigger yawns to stay alert. Yawning can also be a sign of boredom or disinterest in something and it can help the body to release tension and anxiety. Head movement detection can also have several benefits for monitoring student attention, such as providing real-time feedback on focus and engagement. This can help students develop better study habits.

Understanding the correlation between different physiological indicators and attention levels is crucial in developing an attention monitoring system for students. The analysis of blink frequency, gaze direction, yawning, and head movements can provide valuable insights into a student's attention and engagement levels, ultimately improving their academic performance.

Two models such as blink and yawn detection were integrated to create a system that can both assist and monitor a student's academic performance and improve their productivity. The system analyzes the student's engagement and attention using machine learning and computer vision techniques. At the end of the analysis, an evaluation report on the student's academic performance is generated, providing real-time feedback. The system gathers statistics on the student's attention and engagement levels and provides quantitative analysis and measurements on the rates of blinks and yawns. This allows the student to quickly assess their understanding of the learning material and address any issues.

The system can identify patterns in a student's attention levels and suggest strategies or interventions to improve focus and retention. However, it is important to note that the system is not a replacement for traditional teaching methods or the role of a teacher or parents. Another potential drawback is that the system requires a reliable network to run. Although the system has the potential to be a useful tool for improving focus and engagement during learning activities, it is important to consider the limitations and ethical concerns associated with using such a system. It should be viewed as a supplement to traditional teaching methods, and the ultimate responsibility for learning remains with the student.

References

[1] Abate, A. F., Cascone, L., Nappi, M., Narducci, F., & Passero, I. (2021). "Attention monitoring for synchronous distance learning". Future Generation Computer Systems, 125, 774–784.

[2] P. Klein, L. Ivanjek, M.N. Dahlkemper, K. Jeličić, M.-A. Geyer, S. Küchemann, A. Susac, Studying physics during the COVID-19 pandemic: Student assessments of learning achievement, perceived effectiveness of online recitations, and online laboratories, Phys. Rev. Phys. Educ. Res. 17 (1) (2021) 010117.

[3] S.J. Hansen, J.R. VandenPlas, Eye tracking in chemistry education research: Study logistics, in: Eye Tracking for the Chemistry Education Researcher, ACS Publications, 2018, pp. 11–23.

[4] S. Umer, R.K. Rout, C. Pero, M. Nappi, Facial expression recognition with trade-offs between data augmentation and deep learning features, J. Ambient Intell. Humaniz. Comput. 1–15

[5] C. Bisogni, C. Pero, IFEPE: On the impact of facial expression in head pose estimation, in: International Workshop on Biometric Data Analysis and Forensics IWBDAF

[6] M. Knapik, B. Cyganek, Driver's fatigue recognition based on yawn detection in thermal images, Neurocomputing 338 (2019) 274–292.

[7] R. Martins, J. Carvalho, Eye blinking as an indicator of fatigue and mental load – a systematic review, 2015, http://dx.doi.org/10.1201/b18042-48.

[8] Detection of vigilance performance using eye blinks, Applied Ergon. 45 (2, Part B) (2014) 354–362, http://dx.doi.org/10.1016/j.apergo.2013.04.020.

[9] R. Daza, A. Morales, J. Fierrez, R. Tolosana, MEBAL: A multimodal database for eye blink detection and attention level estimation, 2020, arXiv preprint arXiv:2006.05327.

[10] P. Klein, L. Ivanjek, M.N. Dahlkemper, K. Jeličić, M.-A. Geyer, S. Küchemann, A. Susac, Studying physics during the COVID-19 pandemic: Student assessments of learning achievement, perceived effectiveness of online recitations, and online laboratories, Phys. Rev. Phys. Educ. Res. 17 (1) (2021) 010117.