Machine Learning and Eye-tracking Framework to Detect Engagement in Online Learning

MSc Research Project in Data Anlaytics
Research Project Report

Yogalakshmi Chandrasekar
Student ID: x20221665@student.ncirl.ie

School of Computing
National College of Ireland

Supervisor: Dr.Paul Stynes , Dr.Anu Sahni, Dr.Pramod Pathak

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Yogalakshmi Chandrasekar
Student ID:	x20221665@student.ncirl.ie
Programme:	Research Project Report
Year:	2022
Module:	MSc Research Project in Data Anlaytics
Supervisor:	Dr.Paul Stynes , Dr.Anu Sahni, Dr.Pramod Pathak
Submission Due Date:	15/08/2022
Project Title:	Machine Learning and Eye-tracking Framework to Detect Engagement in Online Learning
Word Count:	5673
Page Count:	17

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Yogalakshmi Chandrasekar
Date:	18th September 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Machine Learning and Eye-tracking Framework to Detect Engagement in Online Learning

Yogalakshmi Chandrasekar
x20221665@student.ncirl.ie

Abstract

This research proposes a machine learning framework that helps to identify the online learning engagement of students. The biggest challenge is to detect the cognitive processes of the students, whether they are engaged in an online learning ecology or in a passive phase that is: not involved in the lecture. This framework combines machine learning model with eye-tracker device. A lab video on android app development is shown to students from all the academic background for 7 minutes. Wherein, the participant's eye movements are tracked using an eye tracking device. The experiment begins by filling the participants fill out the personality questionnaire to access their cognitive processes, learning ability , and alertness level. Followed by that participants answered multiple choice questionnaire based on the tutorial video. The video recorded in the eye tracking devices is processed to prepare it ready for analysing it in the machine learning models. In order to detect correlation between the cognitive presence of the students and engagement. The data collected from the eye-tracking device and the questionnaire is used in the research to get the result. The models are evaluated based on accuracy, balanced accuracy, ROC AUC (receiver operating characteristic), F1 score. The model KNN K-Nearest Neighbour has been consistent across all the 3 experiments and provided accuracy ranges from 67 to 79 percentage. The algorithm Logistic Regression has outperformed all the other model in experiment 3 with the highest accuracy of 89 percentage, precision: 1, Recall: 0.75, F1-score: 0.86 and 91 percentage weighted average.

Area - Machine learning, online learning, cognitive presence, Eye tracking, engagement detection.

1 Introduction

The recent adaptation to the online learning has opened a wide arena for research and studies on online learning. There are several studies conducted using eye tracking devices in the past couple of years to detect engagement in the e-learning. This study focuses on identifying engagement among the participants while watching a tutorial video on android development. Eye-tracking can be an effective tool to enhance the outcome of e-learning models. It can lively capture the students' activities to predict the attention level, empathy, engagement, whether they are stuck in a concept, bored, and even the order of reading the contents. Eye tracking is a process of capturing the eye movement to comprehend where the participants are focusing and the duration of it. They are recorded by calculating fixation, saccade , fixation duration, blink frequency, pupil diameter changes, and dilation.

Fixation is the duration amount of time required taken by fovea to to process an image by fovea.

Saccade is the time interval measured to change focus from an image to another .

The recorded data from an eye-tracker could help the researchers to identify students' interest level, and attention level by capturing the gaze position, blink rate, fixation counts and time. Also Calvi et al. (2008), other emotional entities such as stress levels, tiredness, problem solving can be detected. Pupil dilation Lim JZ (2020) is observed in majority of the experiments if a participant is emotionally toned or watching interesting visuals.

In the recent years several research have been done to detect the emotions of students such as engagement, fatigue, and so on. However, to the best of my knowledge no study has been conducted yet to detect engagement and cognitive and cognitive processes due to the current alertness using the KSS scale (sleepiness scale) among students using eye-tracking device and make predictions using machine learning . This is the novelty factor and motivation in my research paper to detect engagement using eye tracking device combined with machine learning framework.

Identifying the online learning outcome is highly crucial to access the students' performance. The positive performance of a student is based on their state of mind during the lectures. Two years into the remote learning ecology it is still a challenge for students to get accustomed with the online learning. The types of lectures really make a difference in learning outcome. However, it is a topic still not researched yet. It is important for the educators to understand the level of student involvement in theory classes vs the tutorials/lab sessions. Multiple studies have been conducted to identify the involvement of students in online learning using facial recognition, posture of the head, body language, and general posture. In this study an Eye-tracker device is used to record the point gaze of the students that will aid the teachers to understand the involvement of the pupils. This study also helps the data analysts to learn the optimum machine and deep learning model to be applied on the small dataset as it is an emerging need.

This research has a broad scope of benefiting both the teaching and data analysis community. The experiment is conducted on 35 participants from the computational background and 5 from non-computing background. Firstly, it will benefit the teaching community by making them understand what causes students to distract. According to the study Bhardwaj et al. (2021) it is essential to understand when and why students disengage or engage in a lecture and If it is linked exclusively digital learning. Secondly, this research would also aid the data analyst community by providing a real time dataset on SMI eye-tracking device. Not only the dataset but also the questionnaire, and the tutorial video that has been used for the research if the future researchers want to make further development in their studies. Finally, this study also focuses on employing optimum machine and deep learning model that are suited for the small dataset wherein the records are in the range of 20 to 100.

This research explores whether a student is in engaged or disengaged state during digital learning. Most importantly it tries to investigate how accurate the framework can detect the students' engagement level using the data that is derived from the eye-tracking device. Also there have been multiple studies on detecting student engagement but the combination of personality questionnaire and eye tracking data is a novel approach.

To address the research question, the following specific sets of research objectives were derived:

•Conducting eye-tracking experiments on participants to generate video data and a dataset.

•Design machine learning framework to detect the engagement.

•Implement a machine learning framework employing models that are optimum to classify the engagement for a small dataset. To detect the engagement using machine learning algorithms such as SVM, logistic regression, decision tree, and KNN are employed.

•The major contribution of this research is its inter disciplinary nature. Identifying the engagement involves emotional engineering using eye tracker and followed by that making prediction using machine learning models.

The paper is divided in to following sections. The consequent section comprises the detailed discussion of related study in the domain. It also contains the key findings, observation, the tools and techniques that have been used, and result of those papers. Majorly the machine learning approaches and how to identify engagement. section 3 describes the project methodology. Section 4 contains the detailed discussion on the design specification of the eye tracking device, participants, video details and the questionnaire that is going to be applied in this research. Section 5 discusses the implementation of machine learning model applied on the dataset generated from the experiment. Section 6 outlines the result and evaluation of the model. Section 7 is a conclusion of this report which summarises the key findings, future work and the accomplishment.

2 Related Work

2.1 Detection of Student Engagement

Detecting engagement in the student has always been a area of interest in the research scholars related to education section, and emotion analytical scholars. Engagement detection in students has a well defined structureDewan et al. (2018) and divided in to three main divisions. Manual, automatic, and semi-automatic. This is based on the dependencies and the methodologies. In this study semi-automatic method is adapted. The need is more because recently Zhang et al. (2017) it was found that the drop rate from schools is growing exponentially high.

In online learning environment Dewan et al. (2018) for an effective outcome engagement is highly crucial. It is a motivational factors for both parties involved: Teacher and the students. The new generation E-learners definitely need the additional functionally that could track the live engagement status. This would help them immensely by providing timely interventions. Students Al-Hendawi (2012) need personalized teaching model, pedagogical style that they are comfortable with learning. Detecting the online learning and fostering it has been more crucial now than ever.

Like how the ways to detect the engagement is structured in different ways, the engagement process itself in learning is classified as Bosch (2016) affective, behavioral, and cognitive. The cognitive engagement is what aimed to be detected in this study. Cognitive engagement Anderson et al. (2004) represents the collective behavioral pattern of the students during the lectures such as being thoughtful, attentive, understanding the complex concepts and getting back to the teacher with doubts or answering to the questions. Researches emphasize that measuring engagement just with the observational data is not sufficient. In other words taking decision on the eye tracking data alone is not justified. There must be a data that is unique for the individual being detected for the engagement. In this study two kinds of internalized data is employed : personality questionnaire which covers a broad spectrum of the participants' behaviour, sleepiness, and alertness and then the MCQ to access how well they understood the concepts.

2.2 Student Engagement Detection using Eye-Tracker device

Eye tracking experiments are highly beneficial and time effective Lim JZ (2020) in the field of emotion detection. Emotion engineering is a upsurging interest for several researcher and there are multiple studies carried out to identify the engagement but using the eye tracker devices and machine learning there are really a few. The engagement detection is useful not only in the learning outcome perspective, it can also help a lecturer to identify if a student is consistently disengaged. It can be a effective tool for mental health supervision, security, and safe driving.

Eye tracking tool could be a profound tool in some business such as website companies. It acts as a ultimate tool Bulygin and Kashevnik (2022) in identifying the order of consumption of content in the website by the users, if a particular content captures more attention, and the user experience.

Studies were able to detect engagement using eye-tracker with good frequency Bhardwaj et al. (2021) achieved 93.6 percentage accuracy, Precision:94.48, and Recall: 87 percentage which is really good result for a mission like this. The study Wang et al. (2020) that used web camera images and videos to detect engagement in the learners in online classes used CNN based deep learning model used 24911 images. It was able to achieve the accuracy of 35 percentage only. This is because of the limited number of features. It is taken into account and data collected as much as possible to predict the engagement level of the students precisely.

The study Binh et al. (2019) that proposed a new algorithms such as CNN, CNN with VGG16, CNN with VGG19 , and CNN with VGG16 and VGG19 was able to achieve the accuracy of 68.8, 74.9, 76 and 80.8 percentage. The study was aimed to create an intelligent tutoring system but this study concludes that the abstract details of the student posture or the gaze itself is not adequate to identify the engagement as it is highly subjective. The information such as domain knowledge, student-trait, human-machine interface, teacher student familiarity and comfort level would increase the accuracy as well as the logic of prediction. Another study Liu and Wang (2006), Matusz et al. (2020) that set to detect engagement among 92 participants was able to achieve Average SU score of 74.68 percentage. It used eye gaze and gaze gesture as a prominent variables.

2.3 Median Split studies to classify the participants

In order to categorize the participants in to two categories median spilt is considered DeCoster et al. (2011) in this research project. There are multiple ways to divide the groups in to category when the values are continuous value. It is the most convenient and easiest way to interpret and come to a conclusion in a study. However, the other methods such as proportional split is also analysed. But it is suitable for a group which has more number of population as well as it is the strategy to be used to match the original population's proportion. Mean split is not suitable for this study. Median split is found optimum as distribution is normal. The extreme group analysis was not considered as there are no extreme low groups in the participants.

Although there are a few mentions about median split approach being a major cause for producing type 1 error and also loss of information. However, the study done by Iacobucci et al. (2015b) disproves it and states multiple reason for using median split approach to dichotomies a continuous variable.The median splits are completely a rationale approach when the independent variables are not correlated Iacobucci et al. (2015a).

3 Research Methodology

In this section the complete research methodology followed and its rationale will be explained in detail. Therefore, this section comprises the details of experimental setup, procedure, data extraction, pre-processing, data cleaning, storage, feature extraction, and finally,the machine learning algorithms chosen.

There are totally 26 participants got involved in this research project. Among the 26 participants 20 males and 6 females had a normal healthy eye vision. They are aged between 19 and 52 years. The participants recruited for the experiment are not expected to have any prior knowledge on app development;however, a few of them were from the computational background and most of them are from the educational background such as arts, finance, and healthcare.

3.1 Experimental Setup

The primary equipment that is employed in this research is the SMI eye tracking device which is like a glasses that has three cameras installed in total. The camera on the outside capture the visuals that is seen by the participants once the experiment is started in the ETG application. The camera inside the rim of the eye tracking device is used to capture the point gaze, number of blinks, saccade, and the fixation. Hence, the the data captured is able to generate multiple features which will be discussed later in this3.4 section. For the experiment to run every time the mobile application called 'ETG viewer' must be used. This is installed in the mobile phone that comes along with the eye tracking device. The screen recorder application named 'Mobizen' was used for each experiment to screen capture the entire experiment. The logged experimental data are processed in the laptop that is again exclusive to the SMI eye tracking device that is explained in detail under 3.3 data extraction section.

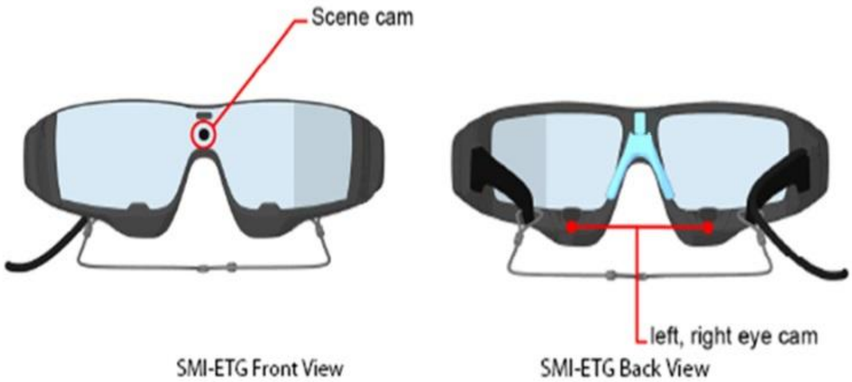

Figure 1: SMI eye tracker[1]

3.2 Experimental Procedure

The most challenging part of this research project is conducting experiment on the 26 participants who signed up to dedicate their time for the benefit of this research. The participants are presented with the consent form [2] before the experiment begins. The consent form comprises exhaustive information such as purpose of the study, procedure, duration, contact details of the researcher,and the incentives. Once the participants consent to conduct the experiment they are asked to fill out the personality questionnaire which is recorded to measure their basic behaviour and also their current alertness level.

After that the participants are made to sit in a quiet place and asked to wear the eye tracking device. The eye tracker device was calibrated, it was a three point calibration essentially the top right, top left and the bottom centre of the screen. Once the calibration is done successfully and the experiment is started using the ETG mobile application that is connected to the eye tracking device the participants are requested not to move their head as it might impact the experiment and the point gaze capturing process.

After the successful completion of watching the 7 minutes practical tutorial video on android app development using android studio.The participants are requested to answer an MCQ[3] on Android app development lab. prepared exclusively for this research study. The maximum points a participant could score is 15. This exercise is done for classifying the participants in to two categories: engaged and disengaged. The strategy employed was median split and this will be given more clarification in the 5 implementation section.

3.3 Data Extraction

The data extraction part of this research project is extensive and involves a lot of steps need to be carried out carefully for all the 26 experiments carried out. The BeGaze guidelines[4] provided from the manual is closely studied and applied for data extraction. Each experiment is stored inside a folder in a memory card mounted on the SMI eye tracking mobile phone. The folder is named by the mobile application automatically by

[2]consent form and personality questionnaire: `https://forms.gle/MY5c9KPfuuz7MbRr6`

[3]MCQ on Android app development lab:`https://forms.gle/TFCMDV8MawU2b9Hx9`

[4]BeGaze Manual: `https://psychologie.unibas.ch/fileadmin/user_upload/psychologie/Forschung/N-Lab/SMI_BeGaze_Manual.pdf`

the application. It uses the information provided at the beginning of the experiment. Essentially, the folder name is as same as the participants' name. This intuitive feature makes it easy to understand and differentiate between the experiments.

The entire recorded eye tracking data folders are transferred from the memory card to the computer and with the use of BeGaze software the folder is located and loaded inside the tool and once the video is processed the metrics are exported by default in a 'txt' format. BeGaze software export trial metrics in multiple forms. It provides event statistics where it generates the details of the trial for each frame and there is a trial summary statistics option which provides the summary statistics for a trial (each experiment) in a single line in a few cases up to 4 rows.

3.4 Data Pre-processing, Feature Extraction and storage

The data gathered in text format from the BeGaze tool had lots of features which can be removed even before any statistical analysis. The txt file is loaded and the following features are removed first: 'Left Mouse Click Count', 'Left Mouse Click Frequency [count/s]', 'Right Mouse Click Count','Right Mouse Click Frequency [count/s]', 'Scanpath Length [px]'. The reason for this is experiment did not involve any activity where the participants have to use the mouse and scanpath summary is irrelevant for this study therefore they are dropped. The MCQ answers are exported in an excel file. The median value of the score is calculated. The median value is 12. Followed by, the participants are categorized into engaged or disengaged. So the participants who scored the points above or equal to the median is categorized as engaged and given the value '1' and rest of them are given the value '0' to indicate that they belong to disengaged category and all the records are stored under the variable name 'Median splits'.

A variable called 'Perclos' is calculated from the feature 'blink count total' and 'total time interval'. After that the variables are tested for high correlation using pearson statistics. The list of variables that had high correlations were removed.

The data from the personality questionnaire and the eye-tracking data was merged for the third experiment. All the data-sets are again stored in onedrive as well as in the github repository.

4 Design Specification

The machine learning integrated eye tracking framework combines the SMI eye tracker metrics and the machine learning algorithm to detect the engagement of students in an online learning ecology. Engagement detection is also a part of emotion engineering. The eye tracking element of the framework consist of SMI eye-tracking device, SMI mobile, screen recorder, BeGaze software tool. The machine learning element of the framework uses the binary classification algorithms which is detailed in section 4.3.

4.1 Creating a dataset

In order to detect the engagement in online learning among students a dataset has to be created. Therefore a tutorial on creating an android application using android studio was recorded exclusively for this research. The video was shown to 26 participants and eye tracking information was recorded. Using the data extracted a small dataset was created which would be helpful for the data analytical community to conduct future research.

4.2 Features and description

This section details the variables that are included in the study to detect the online learning engagement of students. Due to the page constraints only the list of the variables used for the study is shown in the figure below.

1	Trial
2	Stimulus
3	Export Start Trial Time [ms]
4	Export End Trial Time [ms]
5	Participant
6	Color
7	Visual Intake Count
8	Visual Intake Frequency [count/s]
9	Visual Intake Duration Total [ms]
10	Visual Intake Duration Average [ms]
11	Visual Intake Duration Maximum [ms]
12	Visual Intake Duration Minimum [ms]
13	Visual Intake Dispersion Total [px]
14	Visual Intake Dispersion Average [px]
15	Visual Intake Dispersion Maximum [px]
16	Visual Intake Dispersion Minimum [px]
17	Saccade Count
18	Saccade Frequency [count/s]
19	Saccade Duration Total [ms]
20	Saccade Duration Average [ms]
21	Saccade Duration Maximum [ms]

The detailed description of the each metrics could be found in the BeGaze software manual [5] .

21	Saccade Duration Maximum [ms]
22	Saccade Duration Minimum [ms]
23	Saccade Amplitude Total [°]
24	Saccade Amplitude Average [°]
25	Saccade Amplitude Maximum [°]
26	Saccade Amplitude Minimum [°]
27	Saccade Velocity Total [°/s]
28	Saccade Velocity Average [°/s]
29	Saccade Velocity Maximum [°/s]
30	Saccade Velocity Minimum [°/s]
31	Saccade Latency Average [ms]
32	Blink Count
33	Blink Frequency [count/s]
34	Blink Duration Total [ms]
35	Blink Duration Average [ms]
36	Blink Duration Maximum [ms]
37	Blink Duration M

Figure 2: Variables list

[5]BeGaze Manual (page no 253-284): https://psychologie.unibas.ch/fileadmin/user_upload/psychologie/Forschung/N-Lab/SMI_BeGaze_Manual.pdf

4.3 Eye-tracker components

The eye tracker unit of the study has sub elements such as recruiting the suitable participants, circulating the consent, eye tracker, screen recorder, mobile application to operate the experiments, and finally, the BeGaze software tool. The tutorial video is shown to the participants while wearing the eye tracker device. The device has three cameras to capture the eye movement and stimulus of the user. This recorded stimulus data is then fed into the BeGaze software tool to export the metrics.

4.4 Machine Learning Algorithms

Based on the studies and from the state of the art machine learning algorithms to detect the student online engagement were chosen. The datasets are fed in to each model to have a comparative study of the result and to understand the optimum model for the small dataset. The list of classification machine learning algorithms chosen are: Support vector machine, K-nearest neighbor, Decision tree, logistic regression and Adaboost.

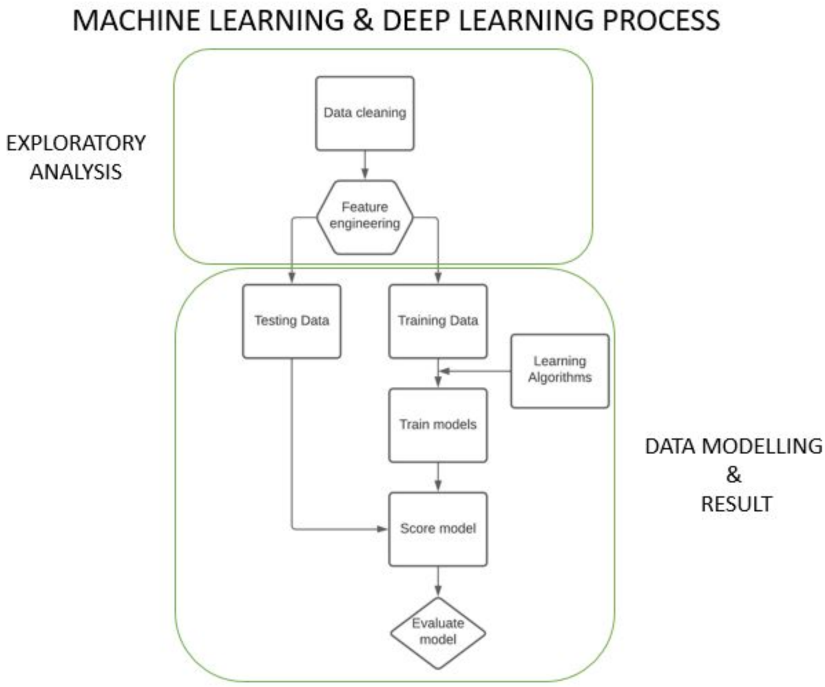

Figure 3: workflow diagram

5 Implementation

The Machine Learning and Eye-tracking Framework to detect the online engagement among the participants is implemented with the help of Python programming language. Jupyter notebook is used as an integrated development environment (IDE) version 6.9.8. Standard libraries for machine learning project implementation offered by Python namely Pandas, Seaborn, Numpy, Scikit learn are used. The data-sets were extracted using PD.read-csv.

There are totally 3 different data-sets that are merged inorder to implement this binary classification model. The data exported from the BeGaze software, two data-sets from the personality questionnaire and the MCQ. The data-sets are concatenated using unique identification number assigned for the experiments. Additionally, a calculated variable called 'PERCLOS' and its values are stored in the dataset.

The most challenging part of this research project were to make the data exported from the BeGaze software collectively to one data file. It is stored in a 'txt' format and the metrics are tab separated by default. In some of the files extracted there was accidental creation of extra folder during the experiment that created trouble while merging the data vertically. The number of columns should be the same while merging so those taken care of and also in some of the records there were accidental extra tab that was counted as a feature by Python and they were checked manually and removed. Final treatment that is done to the metrics exported from the eye-tracking data was to typecast them as numeric as all of them are assigned the type string.

Another challenge in this research study is that masking the personal details of the participants. As it is unethical to have the personal information identify-able in any places of the study. The personal information that are collected in this study includes name, gender, age, email id. These are crucial to collect in order to make a statistical finding based on the demographics and also to identify which experiment belongs to which participants. The metrics exported from the eye tracking data contains name of the participants, The MCQ data contains email id, and the personality questionnaire has age, email id and gender information. To merge all the three collected data into one, firstly, each participant is assigned a unique serial number. secondly, all the three data-sets are carefully re-arranged with respect to the serial number. Thirdly, the number of records are checked and also the number assigned to the participants is same across all the three data-sets. Finally, masking of personal information is done by dropping the columns that contains names and email IDs.

The age and gender columns are retained to make informed prediction based on the age, gender. However, the gender had values such as 'Man' and 'Woman' they are encoded to '1' and '0' to make them possible to be included in the machine learning model. Consequently, the dataframe is split in to 8:2 proportion for training and testing, respectively. The splitting of data was performed using scikit learn library. Finally, the machine learning models such as SVM (Support Vector MAchine), logistic regression, KNN, Decision tree, and Adaboost are implemented on the training dataset employing scikit learn library in Python.

6 Evaluation

This section of the report discusses the details of all the experiments conducted to detect the online learning engagement among the learners using an eye tracking device designed by SMI. First three sections discusses the result on experiment level and the final section discusses on a project level perspectives. There are 3 experiments carried out in this study with three different combination and utilization of the data collected to detect the engagement. In order to explore and find the optimum model for the dataset and the features, multiple machine learning models were applied on the dataset. Therefore, the optimum model will be identified by comparing the validation parameters such as accuracy, precision, recall, and F1-score. Consequent subsections discusses each experiment, result, comparison of performance among the different models and the result is captured in a single table to make the comparison among the algorithms.

6.1 Experiment 1: Implementation of models using Metrics collected from the SMI eye tracker device.

This experiment has included all the metrics extracted from the BeGaze software that showed no to less correlation between each other. The variables included in the experiment 1 to detect the online engagement includes 'Visual Intake Count', 'Saccade Count', 'Saccade Frequency ', 'Saccade Duration Total ','Blink Count', 'Blink Frequency '. Furthermore, all the variable related to saccade, blink is added in this experiment. The data is split in to train and test in 8:2 ratio in a random shuffle.

Machine Learning Models	Accuracy	Precision	Recall	F1-score	Macro Average	Weighted Average
SVM	0.44	0.4	0.5	0.44	0.45	0.46
KNN	0.78	0.75	0.8	0.75	0.78	0.78
Logistic Regression	0.44	0.33	0.25	0.29	0.42	0.43
Decision Tree	0.78	1	0.5	0.67	0.86	0.84
Adaboost	0.67	0.6	0.75	0.67	0.68	0.68

Figure 4: Results Experiment-1

The table illustrates the results obtained by each machine learning model. In order to not limit the evaluation to just accuracy, confusion matrix was generated to obtain results such as precision, recall, F1-score additional to accuracy. **It is observed that decision tree has out performed other 4 machine learning models with the accuracy of 78 percentage** for the dataset that contains eye tracker metrics.The next experiments aims to investigate how accurate a machine learning model could detect engagement using calculated variable from the eye tracker data.

6.2 Experiment 2: Implementation of models using Metrics collected from the SMI eye tracker device and the calculated variable

This experiment has included all the metrics extracted from the BeGaze software that showed no to less correlation between each other and the calculated 'perclose' from the metrics blink count and total time interval .The goal is to investigate how accurate a model can detect the engagement.The aim is to find out if the model accuracy can be improved by including the calculated variable 'PERCLOS'

Machine Learning Models	Accuracy	Precision	Recall	F1-score	Macro Average	Weighted Average
SVM	0.44	0.4	0.5	0.44	0.45	0.46
KNN	0.78	0.75	0.75	.8 0	0.78	0.78
Logistic Regression	0.44	0.5	0.6	0.55	0.42	0.44
Decision Tree	0.33	0.33	0.5	0.4	0.33	0.33
Adaboost	0.33	0.38	0.75	0.5	0.38	0.33

Figure 5: Results Experiment-2

The table provides the result obtained by different machine learning model in the second experiment. **The highest accuracy achieved in this experiment is 0.78 by the model KNN.** Followed by that SVM and Logistic Regression has performed with 0.44 accuracy. It is evident that including the calculated variable does not have any impact in the performance of the model. In the following experiment the data collected from the questionnaire.

6.3 Experiment 3: Implementation of models using Metrics collected from the SMI eye tracker device,'PERCLOS', and Personality Questionnaire

This experiment has included all the metrics extracted from the BeGaze software that showed no to less correlation between each other and the data collected from the personality questionnaire. The aim of this experiment is again to explore if the records from the personality questionnaire such as behaviour, attitude and current alertness level has any impact (either positive or negative) in the outcome of the result.

Machine Learning Models	Accuracy	Precision	Recall	F1-score	Macro Average	Weighted Average
SVM	0.67	0.6	0.75	0.67	0.68	0.68
KNN	0.67	0.67	0.8	0.73	0.67	0.67
Logistic Regression	0.89	1	0.75	0.86	0.92	0.91
Decision Tree	0.78	1	0.5	0.67	0.86	0.84
Adaboost	0.67	0.6	0.75	0.67	0.68	0.68

Figure 6: Results Experiment-3

The table shows the result obtained from the experiment. **The model Logistic regression has predicted the engagement of the learners with the highest accuracy of 89 percentage.** The consequent section discusses in detail the performance of the model, key findings of this research and its significance.

6.4 Discussion

This section details the overall performance observation of the experiments in a project level. The project began by co-ordinating with participants to involve in the experiment. After the eye tracker experiment is done the video is processed using the BeGaze software to obtain the matrics. Followed by that data cleaning was taken place using python programming language in Jupyter Notebook as IDE. The correlation among the variable is checked prior to applying the model and the features that showed high correlations were removed systematically. Below chart provides the collective information of the accuracy in all the machine learning algorithms employed across all the experiments. Also the results obtained with the help of confusion matrix is shown below for the optimum models from each of the experiments. Confusion matrix Chicco Davide Tötsch (2021) is very useful in the data analytical project as picking the best fit model based on just the accuracy can be misleadingMangalathu et al. (2020).

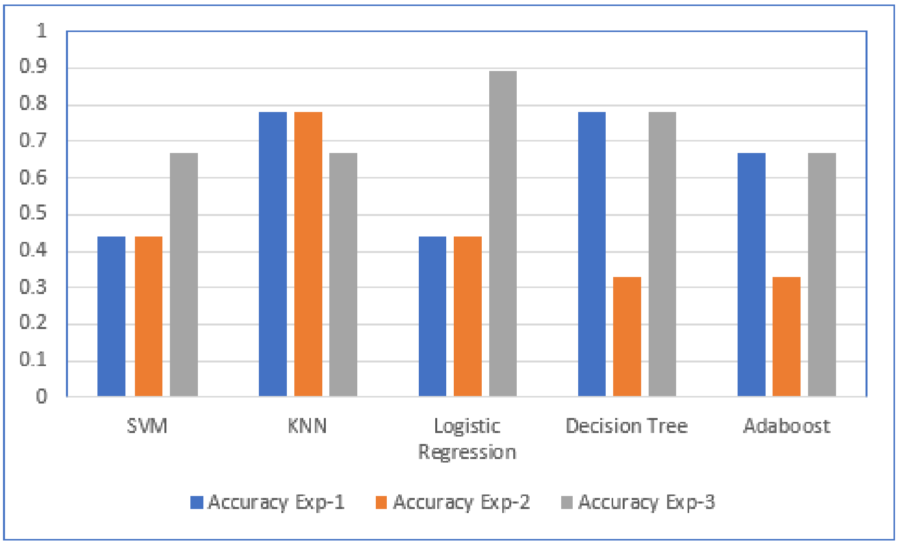

Figure 7: Overall Accuracy comparison of Machine Learning Models

It is observed from the chart that experiment 3 has achieved maximum accuracy in detecting the online engagement in student using logistic regression algorithm. The accuracy obtained was 89 percentage. The third experiment included the data collected from the eye tracker device, calculated variable and the data recorded from the personality questionnaire. Experiment 2 was conducted to find out if the model performance improve when the calculated variable 'PERCLOS' is included in the variable as it is generated by the blink count. As blink behaviour is an indication if a person is involved in the learning. However, the calculated variable negatively impacted the accuracy. Because the researchesBulygin and Kashevnik (2022) , Banfi et al. (2019) that used 'perclos' to detect student engagement actually found the variable improved the model performance. In a research study Patel (2021) to find out zoom fatigue the variable 'PERCLOS' was used and it showed the same characteristics in that research. It is surprising to see the blinking behaviour is negatively impacting in the online learning engagement detection.

```
Classification Result for KNN for First experiment :
[[3 1]
 [1 4]]
              precision    recall  f1-score   support

         0.0       0.75      0.75      0.75         4
         1.0       0.80      0.80      0.80         5

    accuracy                           0.78         9
   macro avg       0.78      0.78      0.78         9
weighted avg       0.78      0.78      0.78         9
```

Figure 8: Confusion Matrix for the Best Fit Machine Learning Model in Experiment-1

```
Classification Report for KNN for Experiment 2 :
[[3 1]
 [1 4]]
              precision    recall  f1-score   support

         0.0       0.75      0.75      0.75         4
         1.0       0.80      0.80      0.80         5

    accuracy                           0.78         9
   macro avg       0.78      0.78      0.78         9
weighted avg       0.78      0.78      0.78         9
```

Figure 9: Confusion Matrix for the Best Fit Machine Learning Model in Experiment-2

It is also evident from the figure 8 that KNN model was consistent among all the experiments in this research and provided accuracy above 67 percentage. To be precise 67 percentage in experiment 3 and 79 percentage in both experiment 1 and 2. SVM and Logistic Regression performed very poor in the first 2 experiments. Decision tree and Adaboost were least performing with the accuracy of 33 percentage in second experiment.

```
Classification Report for Logistic Regression for experiment 3 :
              precision    recall  f1-score   support

         0.0       1.00      0.75      0.86         4
         1.0       0.83      1.00      0.91         5

    accuracy                           0.89         9
   macro avg       0.92      0.88      0.88         9
weighted avg       0.91      0.89      0.89         9
```

Figure 10: Confusion Matrix for the Best Fit Machine Learning Model in Experiment-3

It is observed from the chart that experiment 3 has acheived maximum accuracy in detecting the online engagement in student using logistic regression algorithm. The accuracy obtained was 89 percentage. The third experiment included the data collected from the eye tracker device, calculated variable and the data recorded from the personality questionnaire. Experiment 2 was conducted to find out if the model performance improve when the calculated variable 'PERCLOS' is included in the variable as it is generated by the blink count. As blink behaviour is an indication if a person is involved in the learning. However, the calculated variable negatively impacted the accuracy. In a research study to find out zoom fatigue the variable 'PERCLOS' was used and it showed the same characteristics in that research. It is surprising to see the blinking behaviour is negatively impacting in the online learning engagement detection.

Therefore it is evident based on the performance of the model that KNN would be the optimum model for the research involving small data-sets ranges from 30 to 100 instances. This is one of the research objectives. But as far as the dataset in this 'Machine learning based Eye Tracking Framework to detect student Engagement' is concerned the optimum model is Logistic Regression and the dataset to be included is the combined records of eye tracker data and the personality questionnaire. The next section discusses the conclusion, future work and the challenges faced in this study.

7 Conclusion and Future Work

This research project successfully proposes a **Machine Learning and Eye-tracking Framework** to Identify Engagement in Online Learning. The novelty in the research is that it combines the eyetracker data, personality questionnaire and the calculated variable 'PERCLOS' to detect the engagement using a tutorial video. This data analytical sector of this research is done utilizing the video data logged from the eye tracking experiment by which the metrics are exported employing the BeGaze software tool. Additionally, the data from personality questionnaire and the scores from the MCQ played a major role in the research outcome. The classification of the participants based on the score they secured in the MCQ questionnaire with the help of median split strategy.

This research project is able to present that the machine learning models such as SVM, KNN, Logistic Regression, Decision Tree and Adaboost provide highest accuracy when the data collected from eye tracker device and the personality questionnaire is included. The feature in the final experiment contains 26 variables, Logistic regression model was able to detect the engagement with 89 percentage accuracy with the precision:1, Recall:0.75, F1-score:086, Macro Average:0.92, and the weighted average:0.91.

To conclude, it is evident that the framework structured in this research project has the capacity to detect the online engagement in the online learning ecology with good accuracy. This study can be further extended by utilizing the point gaze video extracted post processing in the BeGaze tool. It has a good scope in the technical learning development research. Moreover, deep learning could be done on those videos to understand the layers involved in both engagement and disengagement from the tutorial by the students in the online learning.

There are a few points to be noted for the future extension of this project. There was an abnormal competitive attitude noticed with a few participants as they were informed about the MCQ post experiment. This is a very good point to note as the test at the end helps the student engage them well with the tutorial. There was a loss of a few experiments as there was absence of advanced features in the ETG mobile application like automatic screen locks inspite of providing the necessary settings, therefore this must be taken care properly. This research can be used for practical application additional to detecting engagement in online learning such as safe driving, medical field, android app development and in emotion engineering researches. Moreover a dynamic software or application can also be developed with the massive computational power of machine learning and eye tracker device.

References

Al-Hendawi, M. (2012). Academic engagement of students with emotional and behavioral disorders: existing research, issues, and future directions, *Emotional and Behavioural Difficulties* **17**(2): 125–141.
URL: *https://doi.org/10.1080/13632752.2012.672861*

Anderson, A. R., Christenson, S. L., Sinclair, M. F. and Lehr, C. A. (2004). Check connect: The importance of relationships for promoting engagement with school, *Journal of School Psychology* **42**(2): 95–113.
URL: *https://www.sciencedirect.com/science/article/pii/S0022440504000172*

Banfi, T., Coletto, E., d'Ascanio, P., Dario, P., Menciassi, A., Faraguna, U. and Ciuti, G. (2019). Effects of sleep deprivation on surgeons dexterity, *Frontiers in Neurology* **10**.
URL: *https://www.frontiersin.org/articles/10.3389/fneur.2019.00595*

Bhardwaj, P., Gupta, P., Panwar, H., Siddiqui, M. K., Morales-Menendez, R. and Bhaik, A. (2021). Application of deep learning on student engagement in e-learning environments, *Computers Electrical Engineering* **93**: 107277.
URL: *https://www.sciencedirect.com/science/article/pii/S0045790621002597*

Binh, H. T., Quang Trung, N., Nguyen, H.-A. T. and The Duy, B. (2019). Detecting student engagement in classrooms for intelligent tutoring systems, *2019 23rd International Computer Science and Engineering Conference (ICSEC)*, pp. 145–149.

Bosch, N. (2016). Detecting student engagement: Human versus machine, p. 317–320.
URL: *https://doi.org/10.1145/2930238.2930371*

Bulygin, A. and Kashevnik, A. (2022). Methodology for in-the-wild driver monitoring dataset formation, *2022 31st Conference of Open Innovations Association (FRUCT)*, pp. 30–36.

Calvi, C., Porta, M. and Sacchi, D. (2008). e5learning, an e-learning environment based on eye tracking, pp. 376–380.

Chicco Davide Tötsch, N. J. G. (2021). The matthews correlation coefficient (mcc) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation, *BioData Mining*, Vol. 14.

DeCoster, J., Gallucci, M. and Iselin, A.-M. R. (2011). Best practices for using median splits, artificial categorization, and their continuous alternatives, *Journal of Experimental Psychopathology* **2**(2): 197–209.
URL: *https://doi.org/10.5127/jep.008310*

Dewan, M. A. A., Lin, F., Wen, D., Murshed, M. and Uddin, Z. (2018). A deep learning approach to detecting engagement of online learners, pp. 1895–1902.

Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J. and Popovich, D. L. (2015a). The median split: Robust, refined, and revived, *Journal of Consumer Psychology* **25**(4): 690–704.
URL: *https://www.sciencedirect.com/science/article/pii/S1057740815000753*

Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J. and Popovich, D. L. (2015b). Toward a more nuanced understanding of the statistical properties of a median split, *Journal of Consumer Psychology* **25**(4): 652–665.
URL: *https://www.sciencedirect.com/science/article/pii/S1057740814001107*

Lim JZ, Mountstephens J, T. J. (2020). Emotion recognition using eye-tracking: Taxonomy, review and current challenges, *Sensors (Basel)*.

Liu, N. and Wang, H. (2006). Feature extraction with genetic algorithms based nonlinear principal component analysis for face recognition, *18th International Conference on Pattern Recognition (ICPR'06)*, Vol. 3, pp. 461–464.

Mangalathu, S., Hwang, S.-H. and Jeon, J.-S. (2020). Failure mode and effects analysis of rc members based on machine-learning-based shapley additive explanations (shap) approach, *Engineering Structures* **219**: 110927.
URL: *https://www.sciencedirect.com/science/article/pii/S0141029620307513*

Matusz, V. I., Rácz, M., Rokai, J., Jorgosz, N. E., Molnár, T., Maráki, D., Ulbert, I. and Márton, G. (2020). Head-mounted, wireless eyetracker for real-time gaze prediction utilizing machine-learning, *2020 IEEE International Conference on Human-Machine Systems (ICHMS)*, pp. 1–4.

Patel, A. (2021). A machine learning based eye tracking framework to detect zoom fatigue.
URL: *http://norma.ncirl.ie/5207*

Wang, Y., Kotha, A., Hong, P.-h. and Qiu, M. (2020). Automated student engagement monitoring and evaluation during learning in the wild, *2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, pp. 270–275.

Zhang, Y., Wang, H., Xu, F. and Jia, K. (2017). A deep face recognition method based on model fine-tuning and principal component analysis, *2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 141–146.