National College of Ireland

A Machine Learning framework to Detect Student's Online Engagement

MSc Research Project
Data Analytics

Komal Riddhish Bharadva
Student ID: x19213051

School of Computing
National College of Ireland

Supervisor: Paul Stynes, Anu Sahni, and Pramod Pathak

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Komal Riddhish Bharadva……………………………………………………………………

Student ID: x19213051………………………………………………………………………………………

Programme: Data Analytics…………………………………………… **Year:** 2020-21…………

Module: MSc Research Project…………………………………………………………………………

Supervisor: Paul Stynes, Anu Sahni, and Pramod Pathak…………………………………….

Submission Due Date: 16/08/2021………………………………………………………………………………

Project Title: A Machine Learning framework to Detect Student's Online Engagement

Word Count: 5055…………………………… **Page Count** 15………………………………………

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: Komal Riddhish Bharadva……………………………………………………………………

Date: 15/08/2021…………………………………………………………………………………………

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):	

A Machine Learning Framework to Detect Student's Online Engagement

Komal Riddhish Bharadva
x19213051

Abstract

Detecting student's engagement in online lectures involves monitoring eye movement as they learn concepts or complete tutorials. The challenge is to detect if a student is engaged or distracted. This research proposes a machine learning framework to identify if the students are engaged in learning. The framework combines a machine learning model and an Eye-tracker device. Students must wear an eye tracker device and are then shown a video lecture of approximately 25 minutes. This is followed by a questionnaire that assesses the student's cognitive processes, transfer, and retention learning. The video from the eye tracker device is analysed and then this data is processed by machine learning models. The results show that we were able to detect student's online engagement using Random Forest Classifier which outperformed other models with 88.9% accuracy. This research can benefit the teaching industry to understand student's engagement in an online learning environment as this directly impacts their learning outcomes.

Keywords: Eye-tracker device, Machine learning, Engagement detection, Online learning, Cognitive processes

1 Introduction

Detecting online engagement is very important as student's success is highly dependent on their state of mind[1]. This is sometimes challenging as the students who are habitual of classroom learning may face some difficulties in adjusting to the online environment. It is important for a lecturer to evaluate the engagement of the students as this is directly associated with their learning outcomes. The student's feedback is important to understand their overall learning experience. Several studies have proposed different ways to detect online student engagement which includes facial expression recognition, head pose, gestures, postures, and so on. This study uses an Eye-tracker device that captures the eye movement of the students in an online learning process which will help lecturers to get a sense of student engagement. This can be implemented by developing a Machine Learning model and evaluating a test sample as carried out in the state of the art [3] [9].

The aim of this research is to investigate to what extent the proposed machine learning framework can be used to detect online student engagement. To address the research question, the following sets of research objectives were derived:

- Investigate the state of the art broadly around student engagement in an online environment using the proposed machine learning framework.

[1] https://edtechmagazine.com/higher/article/2020/12/3-ways-increase-student-engagement-online-learning

- Design a machine learning framework for engagement detection.
- Implement the machine learning framework using various models such as Logistic Regression, Random Forest Classifier, Linear SVC, Bernoulli Naïve Bayes, and K-Neighbors Classifier.
- Evaluate the machine learning model's performance to detect the student's online engagement using various evaluation metrics such as accuracy, balanced accuracy, AUC-ROC score, and F1-score.

The major contribution of this study is a machine learning model which is a novel combination of a machine learning framework and an eye-tracker device focused on detecting student's engagement in an online environment. This model is useful for lecturers to understand if the student is engaged during online lectures.

This paper discusses various tools and techniques used by researchers for student online engagement detection in section 2. The research methodology is discussed in section 3. Section 4 discusses the design components for the proposed machine learning framework. The implementation of this research is discussed in section 5. Section 6 presents and discusses the evaluation results. Section 7 concludes the research and provides some directions for future work.

2 Related Work

Over the recent years, numerous studies have been done to detect online student engagement using machine learning, deep learning, and computer vision techniques. Some researchers have used features like fixation (that is the time for which the eye is kept at a certain area of interest (AOI)) and a saccade (that is the eye movement from one AOI to another). The following paragraphs discuss how the eye-tracking device is actively used in online learning for research purposes and different aids that can be used to improve online learning.

People nowadays are adopting online learning from their usual method of classroom learning. Research shows that the use of visual cues combined with textual data helped students to gain more information as compared to traditional methods [1]. It was observed that this combination supported participant's transfer learning. Another investigation showed the use of hands for learning [2]. The results showed that the use of gestures creates a connection between the text and picture which in turn affects the learning performance of an individual. The ANOVA method for assessing learning performance, cognitive load ratings, and eye movements was used. A similar study records the eye movement of students and introduces some gaze variables which are stimuli-based [3]. In this study student's overall performance is predicted from their behaviour with less than a 5% error rate using Artificial Intelligence (AI) algorithms.

Another research suggests the scan paths (that is the sorted group of fixation points connected by saccades) can be used to analyse the participants learning with the help of eye-tracking technology [4]. Statistical methods like ANOVA and k-means clustering were used to detect the participants having the most effective scan paths. A similar investigation detects learner's engagement in the training of quadcopter controllers conducted in three scenarios: video, live, and simulation [5]. This study used ANOVA and Logistic Regression for the detection of low and high-level engagement. A comparable study uses eye-tracking technology in the Industrial design domain [6]. The participants performed a searching task in the tool board displayed. Statistical analysis like ANOVA was performed using SPSS.

Moreover, a study captured eye movement patterns for participants, and feedback was taken to understand their feeling towards the instructor's present and absent video [7]. The results show that the presence of the instructor highly affected transfer performance positively reduced cognitive load, and improved judgment of learning. An attempt has also been made to review the existing methods, datasets, and metrics used for the evaluation of student engagement level detection [8]. The testing based on computer vision was done and found to be more efficient although they have some drawbacks. Another experiment captured the eye movements using an eye tracker sensor while attendees were instructed to engage with the learning video [9]. The key features that contributed to the best fit of the model were fixations and saccade obtained using Support Vector Machine Recursive Feature Elimination (SVM-RFE). Different machine learning classification algorithms were applied out of which Naïve Bayes was the best model with 71% accuracy.

Another study shows that educational recommender systems (ERSs) can be used to recommend personalized educational resources to students [10]. Overall, there was a positive response on the student's engagement and their perspective while using ERSs with complementing interface and OLM. Although there were few instances where engagement and interest in participants were not seen much. One of the investigations shows that the eye-tracking data can reveal some information about an individual as well as social cognitive processes which is helpful for the diagnosis of various health issues [11]. Several parameters such as user's identity, personality traits, preferences, cognitive processes, and health condition can be inferred from eye-tracking data. A similar kind of study focuses on the cognitive load while modulating social attention [12]. It was determined that attention plays a very important role in social attention behaviours in both live and video recorded scenarios. Another study uses variables related to social cognition for the classification of psychosis (PSY) and found that they are highly associated [13].

Regardless of much previous work has been done to detect student engagement, we were unable to find any study that have the combination of an eye-tracker device with machine learning to detect student engagement in an online environment. There are many investigations done to analyse participant's involvement using eye-tracking devices as well as to classify their learning styles. Also, the use of deep learning for different applications including eye-tracker data and recommendation systems using eye-tracker data is noticed. This is the only study that focuses on student's online engagement detection with the help of a machine learning framework and cognitive processes.

3 Research Methodology

The research methodology followed for this study involves multiple steps like data gathering, data pre-processing, data transformation, data modelling, and results as shown in **Fig. 1**. The first step involves the collection of data with the help of an Eye-tracker device. 31 students (13 females and 18 males) participated in this experiment, and they had some Java background beforehand. Initially, a consent of students is taken before the experiment. After which a 25-minute video lecture is shown to the participants with an Eye-tracker device put on, which captures the student's eye movements. The video consists of a PowerPoint presentation with an instructor video at the bottom right corner describing the same. The topic for this video lecture is Android App Development. At the end of this process, a multiple-choice questionnaire is provided to all participants to answer questions based on video lectures followed by personality traits and social cognition questionnaires. These tests are done to assess participant's transfer and retention learning. All the data is recorded using an Eye-tracker device and converted into a readable format using BeGaze software. This completes the first step of data gathering.

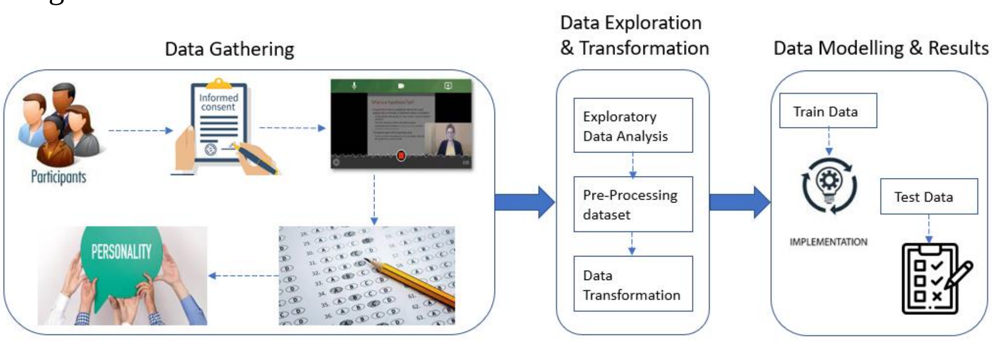

Fig. 1. Research Methodology

The second step is data exploration and transformation that involves exploratory data analysis (EDA), pre-processing dataset, and data transformation. **Fig. 2a** shows the target class distribution. The target variable indicates whether a student is engaged or not. A class imbalance can be observed in the target variable which needs to be rectified before applying machine learning models. Various sampling methods are present from which Synthetic Minority Oversampling Technique (SMOTE) is used in this study.

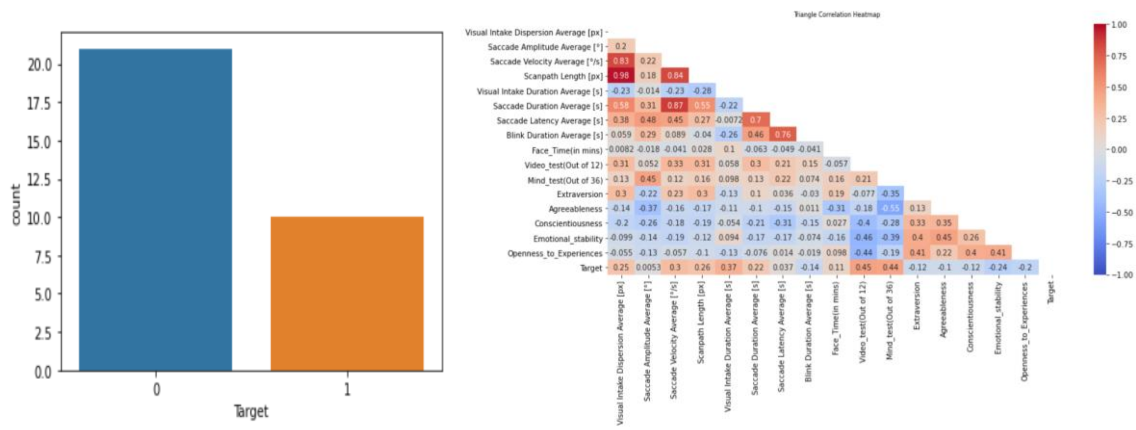

Fig. 2. a) Target class distribution b) Correlation Matrix plot

Fig. 2b shows the correlation matrix plot for all the variables present in the dataset. The correlation between dependent and independent variables can be seen from this graph. A high correlation (p-value greater than 0.90) between Visual Intake Dispersion Average and Scan path Length can be observed. Therefore, the Scan path Length feature is eliminated as this is redundant data. One of the important steps is to check to feature importance as not all the features present in the dataset would contribute to predict the outcome variable. This can be done using any feature selection techniques available. In this study, Recursive Feature Elimination and Cross-Validation Selection (RFECV) is used for feature selection.

The third step is data modelling and results in which the recorded data is converted to an excel sheet and read in as a dataframe. After collecting this data, the entire data is divided into an 80-20 ratio. 80% of the data will be used for training the Machine Learning model and 20% of data to test the trained machine learning model which is the data modelling phase. The different machine learning models used were Logistic Regression, Random Forest Classifier, Linear SVC, Bernoulli Naïve Bayes, and K-Neighbors Classifier. Finally, all the applied models are assessed using evaluation metrics.

4 Design Specification

This section discusses the design specification of this project. **Fig. 3** shows the process flow diagram followed for this research project. Initially, the required data will be gathered and aggregated which is followed by the data pre-processing step. This is followed by data transformation which includes feature scaling that brings all the data in a common scale.

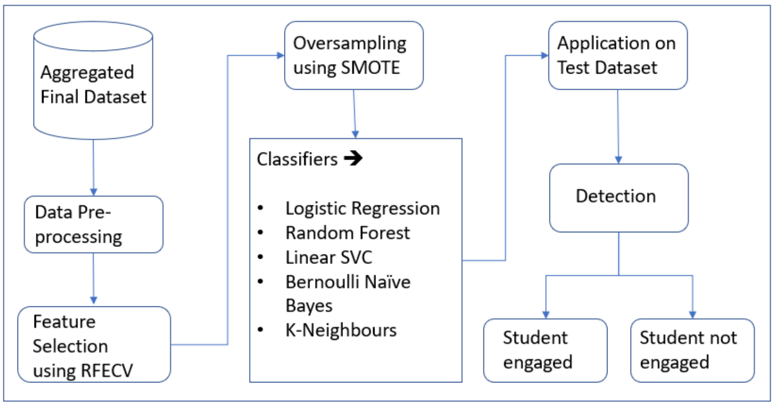

Fig. 3. Process Flow Diagram

One of the studies suggests that the feature selection should always be performed before sampling as the sampling biases the classification towards minority class [14]. Therefore, feature selection is applied first followed by oversampling. The SVM-Recursive Feature Elimination (SVM-RFE) method will be used for feature selection as this method proved to be very useful in the previous study [9]. This is followed by SMOTE which overcomes the overfitting problem by random oversampling of the minority class. The processed data is then fed to five different machine learning classifiers namely, Logistic Regression, Random Forest

Classifier, Linear SVC, Bernoulli Naïve Bayes, and K-Neighbors Classifier. These models were chosen as they gave good results in the related investigations [5] [6] [9] [13]. Finally, the model performance is evaluated on the test dataset for the detection of the outcome variable.

5 Implementation

This section discusses the implementation of the machine learning models for the detection of student's online engagement. Also, the selection of relevant features and oversampling of the dataset are discussed in this section. The Integrated Development Environment (IDE) used for the implementation of this project is Jupyter notebook (v.6.0.3) and the programming language used is Python (v.3.9.0). The python programming language is used for the implementation of this project because it is very powerful and easy to use. It has a wide database of libraries that can be used for almost every task involved in building an end-to-end data science project and is highly scalable.

The dataset used for the implementation of this project is taken from the eye tracker with the help of BeGaze software. The different parameters taken from the eye trackxer are as described in **Table 1** below. All these variables are described in detail in BeGaze software

Parameter	Dimension Unit	Description
Fixation Duration Average	ms (milliseconds)	Sum of duration of all fixations divided by the number of fixations
Fixation Dispersion Average	px (pixels)	Sum of dispersion of all fixations divided by the number of fixations
Saccade Duration Average	ms (milliseconds)	Sum of all saccade duration divided by the number of saccades
Saccade Amplitude Average	degrees	Sum of all saccade amplitude divided by the number of saccades
Saccade Velocity Average	degrees/second	Sum of all saccade velocities divided by the number of saccades
Saccade Latency Average	ms (milliseconds)	Total saccade latency for all saccades divided by saccade count
Blink Duration Average	ms (milliseconds)	Sum of duration of all blinks divided by number of blinks

Table 1: Eye Tracker parameters

documentation[2]. The visual intake is also termed fixation. Apart from the eye tracker parameter, some other parameters are also used as shown in **Table 2**. The 5 personality traits questionnaire variables were derived from 10 variables as per the Ten item personality measure (TIPI) scale which uses reverse scoring[3]. The video test results, mind test results and face time (that is the time for which a student looked at the instructor's face during the whole video) was also used. The final dataset was imported as dataframe in Python and EDA was performed. The

[2] https://psychologie.unibas.ch/fileadmin/user_upload/psychologie/Forschung/N-Lab/SMI_BeGaze_Manual.pdf
[3] https://gosling.psy.utexas.edu/scales-weve-developed/ten-item-personality-measure-tipi/

dataset was scaled using the Standard Scaler library to bring all data on a common scale. After this, the feature selection method was carried out using RFECV which removes irrelevant features based on validation scores with the help of sklearn.feature_selection library.

Parameter	**Units**	**Description**
Video Test	Score out of 12	Test based on video shown
Face Time	Minutes	Time for which student looked at instructor's face
Mind Test	Score out of 36	Reading the mind from the eyes test for cognitive process evaluation
Extraversion	Ranging between 0 to 7	Personality questionnaire variable 1
Agreeableness	Ranging between 0 to 7	Personality questionnaire variable 2
Conscientiousness	Ranging between 0 to 7	Personality questionnaire variable 3
Emotional stability	Ranging between 0 to 7	Personality questionnaire variable 4
Openness to Experiences	Ranging between 0 to 7	Personality questionnaire variable 5
Target	Either 1 or 0	Indicates whether student is engaged or not.

Table 2: Other parameters used

SMOTE technique is used for oversampling which is implemented using imblearn.over_sampling library in python. Finally, five different machine learning classifiers were applied to the final sampled dataset. The different models applied were Logistic Regression, Random Forest Classifier, Linear SVC, Bernoulli Naïve Bayes, and K-Neighbors Classifier. These algorithms are available in different packages in the sklearn library in the Python programming language. Various evaluation metrics were considered like accuracy, balanced accuracy, AUC-ROC score, and F1-score. All these metrics are available in sklearn.metrics library in Python. Detailed implementation and evaluation of these models and metrics are presented in Section 6.

6 Evaluation

The evaluation metrics basically assess the performance of applied machine learning algorithms. Various metrics such as accuracy, balanced accuracy, AUC-ROC score, and F1-score are used for the evaluation of the machine learning models. These metrics are highly suitable for binary classification[4].

6.1 Experiment 1

The aim of this experiment is to observe the eye movements, eye-opening, and closure patterns using the data obtained from the eye tracker device. There are 38 features obtained from the eye tracker device out of which only 7 variables are used for the analysis. **Fig. 4** shows the dataset considered for experiment 1. The feature selection is not performed in Experiment 1 as there are fewer variables and it will not influence the outcome significantly. The SMOTE is used for oversampling since the target variable is highly imbalanced. The dataset is divided

[4] https://thedigitalskye.com/2021/04/19/6-useful-metrics-to-evaluate-binary-classification-models/

into 80-20 ratios for training and testing purposes. Different machine learning models like Logistic Regression, Random Forest Classifier, Linear SVC, Bernoulli Naïve Bayes, and K-Neighbors Classifier were applied.

	Visual Intake Dispersion Average [px]	Saccade Amplitude Average [°]	Saccade Velocity Average [°/s]	Visual Intake Duration Average [s]	Saccade Duration Average [s]	Saccade Latency Average [s]	Blink Duration Average [s]	Target
0	15.300	4.800	76.400	0.372	0.052	0.526	0.206	0
1	26.200	2.700	130.700	0.257	0.047	0.570	0.269	0
2	16.800	2.600	53.600	0.304	0.043	0.569	0.209	0
3	25.200	5.400	53.100	0.643	0.043	0.711	0.216	1
4	917.400	11.400	5910.600	0.255	0.198	1.375	0.613	1

Fig. 4. Dataset considered for Experiment 1

Fig. 5 shows the result of experiment 1. Out of all models, the Logistic Regression model performed very well in terms of all the evaluation parameters considered except for the time taken by the model. The K-Neighbours classifier was the fastest among all the models. **Fig. 6** shows the learning curve of the Logistic Regression model and scalability plot.

Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Time Taken
LogisticRegression	0.889	0.900	0.900	0.889	0.251
RandomForestClassifier	0.889	0.875	0.875	0.886	0.678
LinearSVC	0.778	0.775	0.775	0.778	0.062
BernoulliNB	0.444	0.450	0.450	0.444	0.052
KNeighborsClassifier	0.444	0.450	0.450	0.444	0.027

Fig. 5. Experiment 1 results

The learning curve shows the model learning performance on training and testing datasets over a number of varying parameters[5]. Finding the model complexity with low bias and variance is the key. The higher the score, the better is the performance of the model. The Scalability plot shows the scalability of the model with respect to fit times. The eye movements, eye-opening, and eye-closure patterns were observed using the data obtained from the eye tracker device. These patterns are very useful in detecting student engagement in an online environment. The outcome of the model was verified, and the results are as shown in Fig. 5. The Logistic Regression and Random Forest outperformed with the accuracy of 88.9% though Logistic Regression was fastest among both. Even though good results were observed, but this was not satisfying to identify the student engagement. Hence, there was a requirement of a variable that could lead us to more satisfying results.

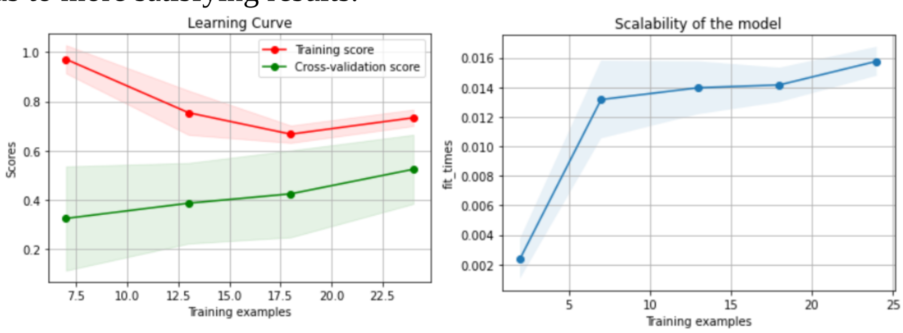

Fig. 6. Logistic Regression Learning Curve and Scalability plot (Experiment 1)

[5] https://www.ritchieng.com/machinelearning-learning-curve/

6.2 Experiment 2

The aim of this experiment is to observe the eye movements patterns as well as assess student's retention learning with the help of a test score based on the video. This experiment also includes the face time variable that can contribute to detect student engagement. This variable is obtained manually for each student. **Fig. 7** shows the preview of the dataset used in this experiment.

	Visual Intake Dispersion Average [px]	Saccade Amplitude Average [°]	Saccade Velocity Average [°/s]	Visual Intake Duration Average [s]	Saccade Duration Average [s]	Saccade Latency Average [s]	Blink Duration Average [s]	Video_test(Out of 12)	Face_Time(in mins)	Target
0	15.300	4.800	76.400	0.372	0.052	0.526	0.206	7	5.920	0
1	26.200	2.700	130.700	0.257	0.047	0.570	0.269	6	6.020	0
2	16.800	2.600	53.600	0.304	0.043	0.569	0.209	6	3.720	0
3	25.200	5.400	53.100	0.643	0.043	0.711	0.216	8	3.530	1
4	917.400	11.400	5910.600	0.255	0.198	1.375	0.613	12	4.380	1

Fig. 7. Dataset used for Experiment 2

The SMOTE technique is used followed by the application of the machine learning model same as applied in Experiment 1. **Fig. 8** shows the results for Experiment 2. We can observe that the Bernoulli Naïve Bayes model gave the best results when all the evaluation parameters are considered.

Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Time Taken
BernoulliNB	0.889	0.900	0.900	0.889	0.028
LinearSVC	0.889	0.900	0.900	0.889	0.029
LogisticRegression	0.889	0.900	0.900	0.889	0.046
RandomForestClassifier	0.889	0.875	0.875	0.886	0.391
KNeighborsClassifier	0.778	0.800	0.800	0.772	0.040

Fig. 8. Experiment 2 results

Fig. 9 shows the learning curve and scalability plot for the Bernoulli Naïve Bayes model. The red line shows the training score, and the green line shows the cross-validation score in the learning curve. The scalability is the plot between the fit times and the training examples. The eye movements patterns and student's retention learning were observed in this experiment. The result shows that the students are able to retain their learning efficiently which also shows their engagement in the learning. This is because after considering the fixation and saccade, it was observed that the student was actually reading the presentation. The test results helped in identifying how much information the student was able to retain.

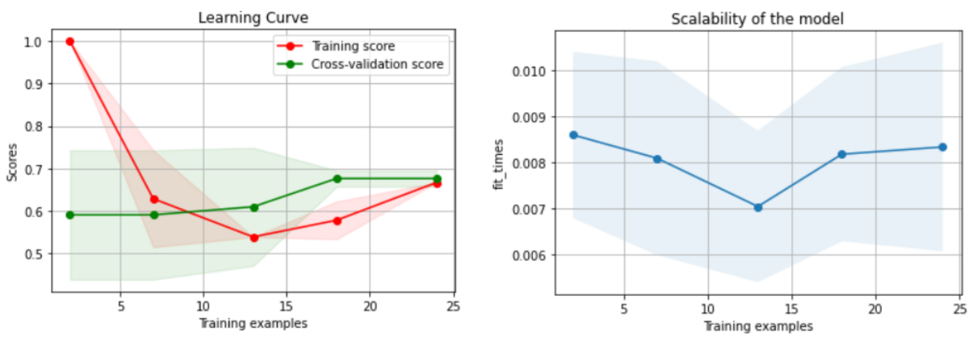

Fig. 9. Bernoulli Naïve Bayes Learning Curve and Scalability plot (Experiment 2)

6.3 Experiment 3

The aim of this experiment is to analyse the social cognitive processes of the students who participated in this research. This experiment includes the data from the eye tracker, video test result, face time, personality traits of students, and social cognition test result. There are total of 17 variables included in this experiment: 7 features from eye tracker, 1 video test variable, 1 face time variable, 5 personality trait variables, and 1 social cognition test variable. The dataset used for Experiment 3 is as shown in **Fig. 10**. The feature selection technique RFECV is used followed by SMOTE technique.

Intake Dispersion Average [px]	Amplitude Average [Â°]	Velocity Average [Â°/s]	Intake Duration Average [s]	Saccade Duration Average [s]	Latency Average [s]	Duration Average [s]	Face_Time (in mins)	Video_test(Out of 12)	Mind_test (Out of 36)	Extraversion	Agreeableness	Conscientiousness	Emotional_stability	Opennes_s_to_Experiences	Target
15.3	4.8	76.4	0.372	0.0518	0.526	0.2056	5.92	7	14	4	4	4	4	4	0
26.2	2.7	130.7	0.2573	0.0473	0.5703	0.2688	6.02	6	21	4.5	1.5	5.5	2	3.5	0
16.8	2.6	53.6	0.3036	0.0433	0.5687	0.2087	3.72	6	22	2.5	1.5	1.5	2	3.5	0
25.2	5.4	53.1	0.6433	0.0426	0.7111	0.2159	3.53	8	16	2	4	4.5	2	2.5	1
917.4	11.4	5910.6	0.2552	0.1978	1.3747	0.6134	4.38	12	18	5.5	2.5	2.5	2.5	3.5	1
37.8	4.5	200.1	0.3908	0.06	0.5729	0.5096	3.87	10	21	1	5	2.5	3	1	1
54	37.9	1104.9	0.3371	0.1272	1.645	4.4389	3.17	7	26	1.5	1	2.5	2	3.5	0
14.2	2.1	42	0.339	0.0428	0.4386	0.2742	3.52	8	14	3.5	4	4.5	4	4.5	1
22.3	6.2	118.9	0.3926	0.0515	0.4528	0.6452	4.28	8	19	2.5	1	3.5	2	2.5	1

Fig. 10. Dataset used for Experiment 3

The machine learning models are then applied similarly to the above experiments. **Fig. 11** shows the result of Experiment 3. We can see that the Random Forest Classifier model outperforms the rest in terms of all the evaluation parameters except for the time taken by the model. The model which took the least time to complete is Logistic Regression though its accuracy is not the best.

Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Time Taken
RandomForestClassifier	0.889	0.900	0.900	0.889	0.481
KNeighborsClassifier	0.778	0.800	0.800	0.772	0.041
LinearSVC	0.778	0.800	0.800	0.772	0.032
BernoulliNB	0.778	0.775	0.775	0.778	0.037
LogisticRegression	0.667	0.675	0.675	0.667	0.031

Fig. 11. Experiment 3 results

Fig. 12 shows the learning curve and the scalability plot for the Random Forest model. A constant training score and fluctuation of cross-validation can be observed in the learning curve. Also, there is not much variation seen in the scalability plot of the model.

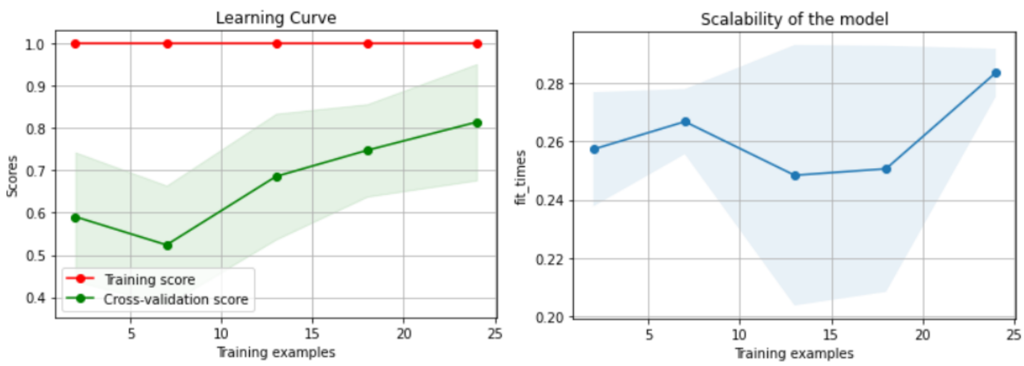

Fig. 12. Random Forest Learning Curve and Scalability plot (Experiment 3)

In this experiment, we were successfully able to capture student's state of mind with the help of personality traits and social cognition questionnaire. In addition to these, data used in Experiment 2 was used in this experiment which enabled us to better understand the student's engagement in the online learning process.

6.4 Discussion

The metrics used for the evaluation purposes were Accuracy, Balance Accuracy, ROC AUC score, F1-score, and time taken by the model. For Experiment 1, the Logistic Regression model gave the best results for the evaluation parameters considered except for the time taken by the model. However, from the learning curve shown in **Fig. 6**, it can be observed that the training score decreases to some point and then increases a bit. This shows the model is slightly underfitting and there is high bias present in the model. The cross-validation score shows that the model is trying its best to learn from the dataset provided. The Scalability plot shows the fit times increasing with the increase in the training examples with a sudden increase at the start. The eye patterns were observed in this experiment and the results show that these patterns are highly functional in detecting student's online engagement. The results obtained from this experiment are somewhat similar to the base papers that emphasize learner engagement [5] [9].

For Experiment 2, the Bernoulli Naïve Bayes model performed well while the other models namely, Linear SVC, Logistic Regression, and Random Forest Classifier also gave the same results as the former model except for the time taken. On the other hand, the learning curve of the Bernoulli Naïve Bayes model shown in **Fig. 11**, indicates that the model is clearly under-fitted and is short of data. There is not much variation seen in the scalability plot of this model. The student's transfer and retention learning were noted in this experiment. This experiment is comparable to one of the previous studies that focus on instructor presence [7]. The outcome shows that the students were able to remember most of the things which confirm their engagement in the video lecture. Though there was a limitation in this experiment, as the face time variable required to be calculated manually.

The model that gave the best results for Experiment 3 is Random Forest Classifier while the Logistic Regression model was the fastest of all. **Fig. 12** shows the Random Forest Classifier learning curve. A clear sign of underfitting can be seen in this learning curve. Therefore, the number of training samples needs to be increased in order to improve the overall model performance. This experiment is similar to previous investigations that concentrate on social cognition aspects for classification [12] [13]. This experiment was focused on the social cognitive processes of the students which gave decent results and hence indicates good student engagement. After conducting all the three experiments, we were able to identify student's state of mind more accurately in terms of their retention learning and overall engagement in the online learning process.

7 Conclusion and Future Work

Detecting student engagement is important nowadays as the mode of learning is shifting towards online learning from traditional methods. In this research, a novel combination of machine learning model and eye-tracker device is used for the detection of student's online engagement. This can be used by the instructors to monitor student's activities during online lectures. The data from the eye-tracker device and post questionnaires is aggregated and fed into a machine learning model. Student's cognitive processes are also assessed during this process. It can be concluded that we were able to successfully detect student's online engagement using machine learning framework and data from the eye tracker, test results, and social-cognitive processes. The parameter that highly contributed to predicting the target variable was fixation duration average, saccade duration average, video test, and mind test results. We have investigated the state of the art, designed and implemented a machine learning framework, and evaluated the machine learning models for engagement detection.

Different machine learning models were applied to detect student's engagement in an online learning environment. The results show that the Logistic Regression model gave the best results in Experiment 1 where only the data from the eye-tracker device was considered. The Bernoulli Naïve Bayes model performed well overall in Experiment 2 while Random Forest outperformed in Experiment 3. We have successfully captured student's engagement using data from eye tracker, personality traits and social cognition questionnaire and few other parameters. Though there were few limitations observed in these experiments. The target variable is highly imbalanced, and the sample size of the dataset is too small for the machine learning model to learn the behaviour of the dataset.

The future work for this research project can include a good size of the dataset. It can also include providing live student engagement level rather than only detecting it. The student's behaviour can also be explored to get a greater sense of their engagement. Various other machine learning and deep learning models can also be applied and evaluated.

References

1. Wang, X., Lin, L., Han, M., & Spector, J. M. (2020). Impacts of cues on learning: Using eye-tracking technologies to examine the functions and designs of added cues in short instructional videos. Computers in Human Behavior, 107, 106279.
2. Korbach, A., Ginns, P., Brünken, R., & Park, B. (2020). Should learners use their hands for learning? Results from an eye-tracking study. Journal of Computer Assisted Learning, 36(1), 102-113.
3. Sharma, K., Giannakos, M., & Dillenbourg, P. (2020). Eye-tracking and artificial intelligence to enhance motivation and learning. Smart Learning Environments, 7(1), 1-19.
4. Saiz Manzanares, M. C., Rodriguez Diez, J. J., Marticorena Sánchez, R., Zaparain Yanez, M. J., & Cerezo Menendez, R. (2020). Lifelong learning from sustainable education: An analysis with eye tracking and data mining techniques. Sustainability, 12(5), 1970.

5. Carroll, M., Ruble, M., Dranias, M., Rebensky, S., Chaparro, M., Chiang, J., & Winslow, B. (2020). Automatic Detection of Learner Engagement Using Machine Learning and Wearable Sensors. Journal of Behavioral and Brain Science, 10(3), 165-178.
6. Rojas, J. C., Cañizares, J. C. M., Higuera-Trujillo, J. L., & Muniz, G. (2020, April). An eye-tracking project in industrial design education: a case study for engaging in the research process. In 2020 IEEE Global Engineering Education Conference (EDUCON) (pp. 127-132). IEEE.
7. Wang, J., Antonenko, P., & Dawson, K. (2020). Does visual attention to the instructor in online video affect learning and learner perceptions? An eye-tracking analysis. Computers & Education, 146, 103779.
8. Dewan, M. A. A., Murshed, M., & Lin, F. (2019). Engagement detection in online learning: a review. Smart Learning Environments, 6(1), 1-20.
9. Pritalia, G. L., Wibirama, S., Adji, T. B., & Kusrohmaniah, S. (2020, September). Classification of Learning Styles in Multimedia Learning Using Eye-Tracking and Machine Learning. In 2020 FORTEI-International Conference on Electrical Engineering (FORTEI-ICEE) (pp. 145-150). IEEE.
10. Abdi, S., Khosravi, H., Sadiq, S., & Gasevic, D. (2020, March). Complementing educational recommender systems with open learner models. In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (pp. 360-365).
11. Kröger, J. L., Lutz, O. H. M., & Müller, F. (2020). What does your gaze reveal about you? On the privacy implications of eye tracking. IFIP Advances in Information and Communication Technology, 576, 226-241.
12. Bianchi, L. J., Kingstone, A., & Risko, E. F. (2020). The role of cognitive load in modulating social looking: a mobile eye tracking study. Cognitive research: principles and implications, 5(1), 1-21.
13. Antonucci, L. A., Raio, A., Pergola, G., Gelao, B., Papalino, M., Rampino, A., ... & Bertolino, A. (2021). Machine learning-based ability to classify psychosis and early stages of disease through parenting and attachment-related variables is associated with social cognition. BMC psychology, 9(1), 1-15.
14. Blagus, R., & Lusa, L. (2013). SMOTE for high-dimensional class-imbalanced data. BMC bioinformatics, 14, 106. https://doi.org/10.1186/1471-2105-14-106