



An Intelligent Examination Monitoring Tool for Online Student Evaluation

Rashidul Hasan Nabil^{*1}, Aneem-Al-Ahsan Rupai¹, Mimun Barid², Adnan Sami¹ and MD. Nazmul Hossain¹

¹ Dept. of Computer Science, American International University-Bangladesh. Dhaka, Bangladesh.

² Dept. of Computer Science and Engineering, University of South Asia. Dhaka, Bangladesh.

KEYWORDS

*Behavioural Biometrics
Monitoring System
Education Technology
Online MCQ
Machine Learning*

ARTICLE HISTORY

*Received 24 July 2022
Received in revised form
15 August 2022
Accepted 15 August 2022
Available online 16 August
2022*

ABSTRACT

The global reach of online education has increased due to a pandemic or other unique circumstances. As online education got more popular, it became crucial to ensure the quality of evaluation. This study's goal is to find a solution to the issue of monitoring during online exams. We have used behavioural biometrics through students' interaction with an Intelligent Examination Monitoring Tool (IEMT), which was developed, even though many studies concentrate on using video analysis. The test-taking prototype uses mouse, touch, and keyboard interfaces to administer multiple-choice questions with a variety of information and events. Students who used additional sources to answer questions were later discovered during an online interview. We built a prediction model that can determine if a student is answering on his own or using any other sources using the events through input interaction when these students are sorted. The Machine Learning (ML) techniques Decision Tree, Random Forest, K-Nearest Neighbour, and Naive Bayes were used to generate a few models. After evaluating the performance of the models, we find that random forest performs best, with an accuracy of about 91 percent.

© 2022 The Authors. Published by Penteract Technology.

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

1. INTRODUCTION

Exams allow instructors to evaluate a student's level of academic knowledge. These examinations can come in various forms like written, in person interviews, presentation, assignments and so on. Traditional ways of evaluation require students to be present physically in a room where invigilators carry out the evaluation process. In course of time, both for adverse situations like pandemic and technological advancement, evaluations for distance learning are getting popular [29]. That is, technology has created new platforms using concepts of asynchronous and synchronous distance learning that helps invigilators to take the examinations even if the students are far apart. Over recent years several online platforms are introduced to conduct students' evaluation [18] [19] [20].

One concern is that online examinations hardly provide an environment for monitoring [28]. Some students use this as an advantage to do unethical actions during an examination; especially for the written and Multiple-Choice Question (MCQ) based formats. These actions might include direct copying of answers from course mates, using the internet or books during a closed book exam, and discussing with course

mates. Several studies work behind these unfair means and outline two main reasons: one is the students think passing will be impossible for them without these activities [24], and peer pressure from concerns (guardians and teachers) for grades [25]. The Open Education Database shows that 68% college students disclose they are engaged in conducting these activities during examination [23] and Open Education Database suggest that these activities start at high school [26]. Studies have shown that students who pass their examinations in this way are likely to show unsatisfactory performance when hired as an employee leading them to perform unethical actions again [21]. On the other hand, students passing out in this way might also put a bad reputation on the institution where they study [22]. So, it is certain that students getting involved in such unfair means in the online examinations is a potential threat that needs to be resolved. Few solutions propose to solve this problem through various studies which are based on using cameras to capture student activities and detect unethical actions using image processing [16] [17], using students log generated during giving an evaluation [14] and using fingerprint, and screen in and out time recording [15]. To the best of our knowledge behavioural biometrics can

*Corresponding author:

E-mail address: Rashidul Hasan Nabil <rashidul@aiub.edu>.

2785-8901/ © 2022 The Authors. Published by Penteract Technology.

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

be used to investigate such behaviours. This study therefore analyses students' behavioural biometrics during such actions.

Behavioural biometric defines different ways to authenticate users based on their physical and psychological aspects [1]. The authentication is done leveraging the use of methods like voice recognition, gait recognition, keystroke, and mouse dynamics, and so on [2]. Over recent years several studies are conducted to identify the significance of these methods for authentication. A study investigates smartphone users' touch behaviour as they read their emails using different applications [3]. Later this behaviour is used for user authentication. In another study human walking pattern is utilized to extract biometric patterns and it is used for authentication [4]. There is also another study that read digits created by hand gestures. There is a limit of n digits which can then be fed as the verification code for authentication [5]. The concept of using fingerprints along with keystroke dynamics is also proven to work more efficiently for authentication purposes [6]. From the studies mentioned above it can be said that human interaction with different machines can be used to generate behavioural models which can be used for authentication.

As a part of this study, we would like to use the students' interactions with a computer or a smartphone during an online examination to create a behavioural pattern. Then with the supporting features that suggest that a student might have conducted unethical actions during the examination, we will create a behavioural model that can later be used to detect such activities. Key contributions of this study are mentioned herewith-

- A tool was developed that can extract students' interaction behaviors during an online examination.
- Students' interactions were used to identify their behaviors that can predict unethical actions during an online examination with the help of ML.

2. LITERATURE REVIEW

2.1 Behavioural Biometrics in Computer Interaction

It is vital to understand that humans' interaction with machines can be utilized for creating behavioural patterns that can assist in biometric authentication. However, humans' interactions may provide more insights if these biometrics can be used to obtain human behaviour during distinctive events. Recent studies focus on making predictions based on human behaviour during specific events. All these studies leverage the use of mouse and keyboard interactions. Monaro et al. [6] conducted a study to detect fake star ratings using mouse movements. They used mouse dynamics for this purpose and concluded that participants who have given fake ratings reflected greater response times and wider mouse trajectories as they have taken more time to start and move the mouse slowly on the screen [6]. Katerina et al. [7] conducted a study to identify the correlations between mouse behavioural patterns or keystroke dynamics and a set of End-User Development (EUD) behavioural attributes. Following a thorough investigation, it was discovered that mouse pattern metrics, such as random and straight motions, mouse hovers, etc., can be linked to self-efficacy, ease of use, perceived utility, and risk perception. Similar to this, certain keystroke characteristics, such as key press speed and down-to-down duration, might be linked to self-efficacy or perceived ease of

use. Nishiuchi et al. [8] conducted a study to measure users' interest level using web access logs. In the experiment, manual measurements are made of the access logs' page transitions, transition times, number of transitions, and task times using video capture software. According to the research, the frequency of page transitions and the average page transition duration during the information integration process can be used to estimate a user's level of interest. In another study, Balen et al. [9] conducted a study to do online gender classification. Based on mouse biometrics, they have presented a naturalistic technique for classifying gender that includes detailed guidance for which metrics (features) of movements are important for gender categorization. In their study, the gender of a participant is correctly classified at a rate of 89.4–100% for the labelled data and 72.4–75.9% for the unlabelled data. Based on the studies above it can be realized that behavioural biometrics is not only limited to user authentication. Further investigation can be done for its use to monitor students' behaviour in the education sector.

2.2 Interaction Tools used in Online Examinations

A considerable number of studies been seen on the use of mouse and keyboard dynamics to extract students' information in education. Tzafilkou et al. [10] conducted a study to identify students mouse behaviour during a web development course. Based on mouse coordinates (x , y), timestamps (in milliseconds), and JavaScript events like mouse hovers, clicks, and moves, a system was created to record mouse interactions. The outcome demonstrates that measures of mouse clicks and hovers can be connected to students' perceptions of usability and usability. Lim et al. [11] carried out research to develop a customized e-learning system that can deliver adaptive learning materials based on a student's cognitive effort and efficiency. Through the acquisition and analysis of task performance, mouse behaviour, and keystroke behaviour, this study carried out automated evaluation of cognitive stress. The conclusion is that combining mouse and keyboard dynamics analysis can be more beneficial than doing so individually because there are links between mouse behaviour and keystroke behaviour. Carneiro et al. [12] proposed a way to measuring student stress levels during online exams that is based on mouse dynamics. Results reveal that mouse dynamics alter consistently as stress builds, enabling its calculation from an examination of mouse behaviour. Salmeron et al. [13] proposed a study to Identify affective states and behavioural changes in an e-learning platform using non-intrusive and inexpensive ways by evaluating the use of mouse and keyboard dynamics. An undetectable mouse tracker and key logger are created in Java using the kSquared.de package in order to log keyboard and mouse movements in real time. According to findings from the study of data from 17 participants, these indications may be helpful in automatically and inexpensively identifying affective states from changes in a participant's behaviour during an engagement with an e-learning platform. All the studies above are evidence on the use of mouse and keyboard interaction to identify students' biometric behaviour when they are interacting with a computer.

2.3 Evaluation Techniques for Online Examinations

According to Martin et al., 2020, one of the most popular study themes in online learning from 2009 to 2018 is course assessment. This trend has been demonstrated. Due to the lack of direct control over students and instructors, course

evaluation in online learning is exceedingly difficult [34]. As part of digitalization, students' evaluation techniques have also been transferred from traditional pen and paper-based examinations to online platforms [35]. The detection and mitigation of online cheating may be more difficult even when the motivations for cheating in offline and online exams are not noticeably different [36]. In this way, it is easy to take the evaluation for a large number of students, do the grading which can sometimes be automated and most importantly record the evaluation information that can be accessed and updated in the simplest way. However, there is one concern that students conduct unethical actions to do better performance. These include taking help from books, internet sources and even from other students which cannot be administered by the invigilator. There have been several studies to find a solution to this problem. Bilen et al. [14] presented a study of online exam cheating of university students that take place during Covid-19. They use the students log information and find that an average time of 30 seconds is required to answer a question. As a result, they suggest adding timestamps to each question which must not exceed 30 seconds. Bawarith et al. [15] implemented an E-exam management system in order to prevent cheating on online exams. Authors used Eye Tribe Tracker to continuously detect a real student and a fingerprint reader to authenticate him. Their implemented system show success with 98% accuracy. Atoum et al. [16] propose a multimedia system which automatically proctors online exams. To do this, they used a webcam, wearcam, and microphone in the system. The system also show success in detecting and preventing cheating. Jalali et al. [17] proposed a method to prevent cheating based on students' webcam images. They invited 10 students and clustered 50 webcam images before and after exams. Authors detect cheating based on distance between an image and a reference image. In another study, Tion et al. [37] utilised an e-cheating intelligence agent for detecting cheating in online examinations, that is comprised of two modules - The Internet Protocol (IP) detector and the behavioural model. They used Deep Neural Network, long-short term memory (LSTM) and recurrent neural network on various data and achieved significant results.

The studies above are a clear indication that behavioural biometrics is not limited to authenticating users for security reasons [16][17]. The concept can be extended further to study human behaviour during different events. It is noticeable that mouse and keyboard dynamics can be cross linked with other factors that together help to study the behaviour of a particular individual. The studies conducted on the educational sector mentioned above serves as evidence to this statement. The literature on unethical action detection mentioned earlier mostly depends on the use of cameras to keep the students under observation. This requires a camera as the mandatory hardware. It is possible that a student's camera might not be working during the examination which might exempt him from the exam, or the invigilator may have to arrange another evaluation. Also, the students might feel uncomfortable as they are forced to keep their cameras on all the time. On the other hand, the invigilator must continuously monitor the students which might get difficult as there are a lot of students in a class. The existing smart monitoring system also requires complex analysis and modelling that might not be time efficient and must be done with the help of high-performance hardware components. A simpler but efficient solution that

uses a student's navigation along with mouse and keyboard interactions during an online examination to understand the type of behaviour they show during conducting unethical actions can be effective for online evaluation systems.

3. METHODOLOGY

This section outlines the procedures that was followed to conduct our study. As shown in Fig.1 multiple steps were executed to build a predictive model that can detect unethical actions during an online examination. These steps are described in the following sections-

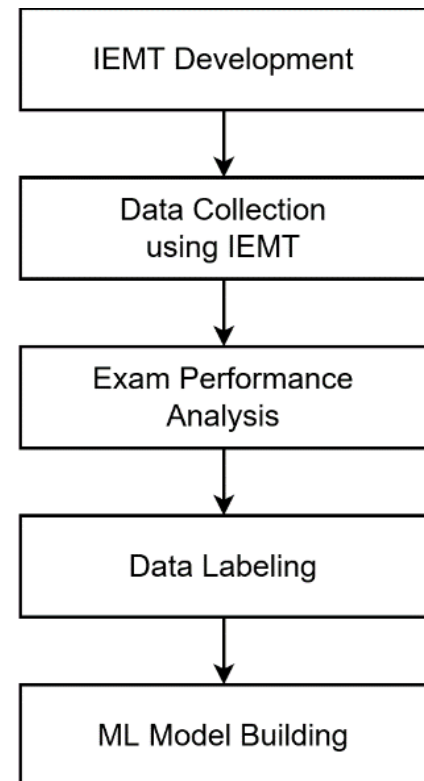


Fig. 1. Methodology Flowchart

3.1 IEMT Development

IEMT was developed in order to extracting students' navigation through mouse and keyboard. The system was created on the web with using HTML, JavaScript, and PHP. The interaction information is collected with HTML Document Object Model (DOM) events [27]. Two html forms were created to collect this information. A student enters into IEMT using his/her ID. All his/her interactions are stored under this ID as CSV files. The interactions are categorized into 2 parts with the help of two different forms: events and mouse or touch movements. Here, mouse movement is for desktop or laptop users and touch for smartphone users. In events, information is gathered like first response time (FRT), last response time (LRT), changes in an answer (CHCOUNT), selected answer (GA), total copy count (CPCOUNT) for each question, focus out count (FOUT), and total idle time (IDLE). On the other hand, in mouse/touch movement, we gathered x and y coordinates. Finally, the CSV files are stored under two different folders for each form in the web server. One folder is for storing events and another is for mouse/touch movements for each student. Here, when the first form is submitted then all the information of this form is stored and then the second

form is initiated. After the second form is submitted, IEMT stops collecting information from the user. Fig. 2 shows IEMT workflow.

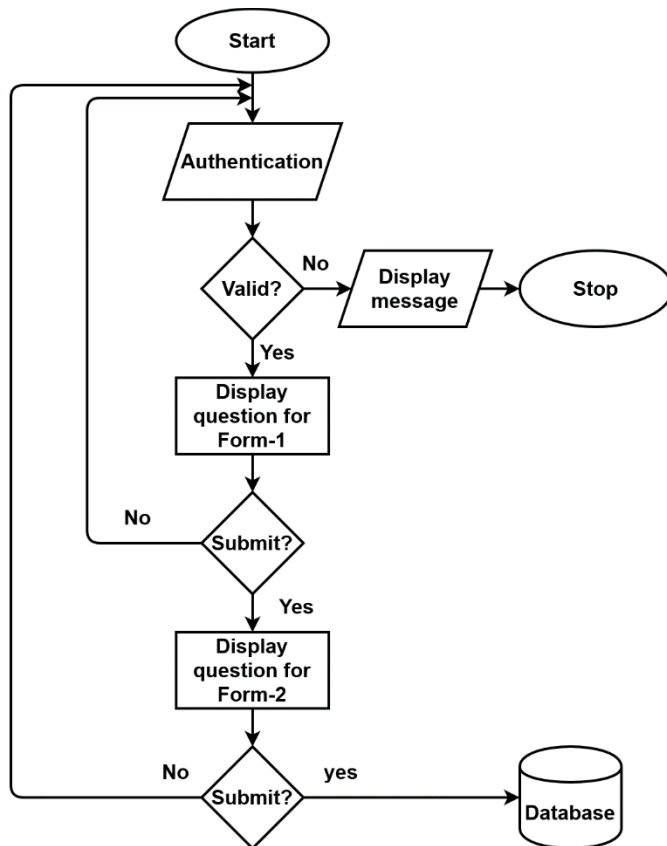


Fig. 2. Workflow Diagram of IEMT

IEMT¹ is available online. Figure 3(a, b, c, d) shows few screenshots of live system prototype.

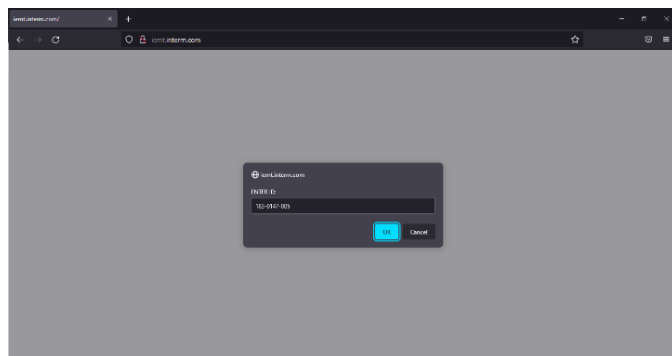


Fig. 3(a). Step 1 - IEMT User Authentication

The online system lands a student in the Authentication page as Fig. 3(a). The student enters his/her organizational ID and enters to the exam. After entering to the exam first the questions from Form-1 appear in the screen shows in Fig. 3(b) containing non-academic questions. From here students' interactions will be recorded. Submission of form-1 will lead the student to the next screen which containing the questions from the course in form-2. The interaction recording will be continued in this page as well. Exam will be ended by submission of form-2 and interactions will be saved to the databases along with the student's answer selection.

1. Which device are you using?
☐ Desktop
☐ Laptop
☐ Mobile
☐ Tablet

2. How is the internet connection in your area?
☐ Excellent
☐ Good
☐ Fair
☐ Poor

3. How are you feeling at this moment?
☐ Very Good
☐ Good
☐ Tensed
☐ Stressed

4. What is your current semester?
☐ 7th Semester
☐ 8th Semester
☐ 9th Semester
☐ Other

5. Do you think this course is relevant to your program (CSE)?
☐ Very Much Relevant

Fig. 3(b). Step 2 - Display question for Form-1

1. Who is one of the inventors of microprocessors?
☐ Stanley Mazor
☐ Bill Gates
☐ Jeff Bezos
☐ Steve Jobs

2. What is the biggest microprocessor company in the world right now?
☐ Intel
☐ Samsung
☐ AMD
☐ Qualcomm

3. Which is not a main part of the functional block?
☐ ALU
☐ Flag Register
☐ Execution Unit
☐ All of them

4. Which generation Intel 80386 belongs to?
☐ 4th Generation
☐ 3rd Generation
☐ 5th Generation
☐ 2nd Generation

5. The pipeline is divided into segments and each segment can execute its operation concurrently. Do you agree?
☐ Strongly Agree
☐ Agree

Fig. 3(c). Step 3 - Display question for Form-2

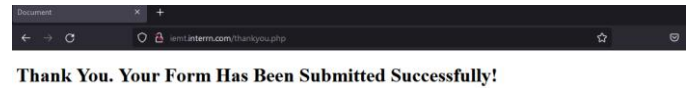


Fig. 3(d). Step 4 - Submission done Saved in Database

3.2 Data Collection using IEMT

For the purpose of the study, data of 81 students were collected from a private university in Bangladesh. The students are exposed to a closed book online examination environment where clear instruction about not doing any unethical action is given. A MCQ based examination was conducted for the "Microprocessor" course using the tool that was designed for the study. The examination consists of two sections. The first section asks a student to input common information like where he stays, about the internet connection of his/her area, and so on which does not need any unethical means to answer. The second part consists of questions where he/she must answer using previous knowledge of the course leading to the possibility of unethical actions. To detect such actions, some events were identified. Those were as follows-

- Total number of times a student opens another tab or application during the time of examination
- Total number of times a student copies a word or text.
- Total number of times a student changes the answer for a question.
- Total idle time of a student during the examination
- The answer to each question
- X and Y coordinates for the mouse or touch movement for each question obtained after every 20 ms.

¹<http://iemt.intern.com/>

The tool was to extract these matrices and later construct the dataset² which is as follows-

Table 1. Dataset Notation Description

Notation of Feature	Description of Feature
SID	Identity number for each student assigned by the university
QID	Unique identity number assigned for each question
FOUT	Total number of times a student opened another tab on the browser or another application
CPCOUNT	Total number of times a student copied a word or text for each question
FRT	The first time a student answered a question
LRT	The last time a student answered a question
CHCOUNT	The total number of times a student changes the answer to a particular question
GA	Final answer for each question
X	X coordinate of mouse movement obtained after every 20 ms for the entire length of the examination
Y	Y coordinate of mouse movement obtained after every 20 ms for the entire length of the examination

3.3 Exam Performance Analysis

This section describes some insights on collected data.

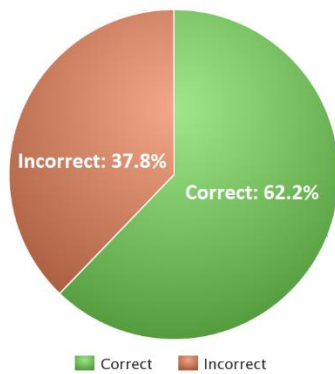


Fig 4(a). Percentage of correct and incorrect answer for the online MCQ examination

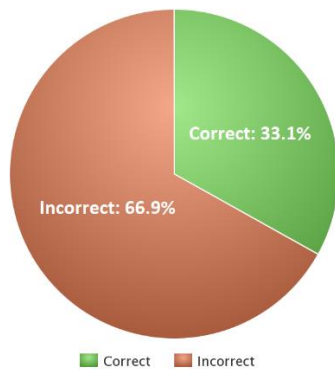


Figure 4(b). Percentage of correct and incorrect answer for the online face to face interview

As the dataset was prepared, the answers for each question were checked. It is seen that about 62% of the

answers are correct as shown in fig. 4(a). After the MCQ test using to developed tools, an online one to one interview session were conducted where the students are asked the same questions again. This time it is found that about 33% of the answers were correct as shown on fig. 4(b). This indicates the fact that some students might have applied unethical actions for the MCQ examination. Keeping that in mind, a comparative analysis was conducted between the events of the two parts of the MCQ examination considering different students' interaction events.

The first part named Form-1 consists of general information such as, Student Name, Student Id etc. which did not require the students to apply course specific knowledge. The second part named Form-2 contained questionnaires that required specific course-based knowledge. It is quite distinct that students who wanted to apply unethical action during the examination, tend to apply more interaction events during participating in Form-2. Thus, the occurrence of the unethical action during the examination can be determined by the equation –

$$E_{F1} < E_{F2} \quad (1)$$

where E_{F1} is the number of events occurred in the Form-1 and E_{F2} is the number events occurred in the Form-2.

From the findings of the interaction events of the conducted during MCQ examination, is plotted in fig. 5. And it was observed Form-2 has more occurrences than Form-1 in every interaction events. Hence, it can be concluded from the output, students tended to apply unethical action during the examination.

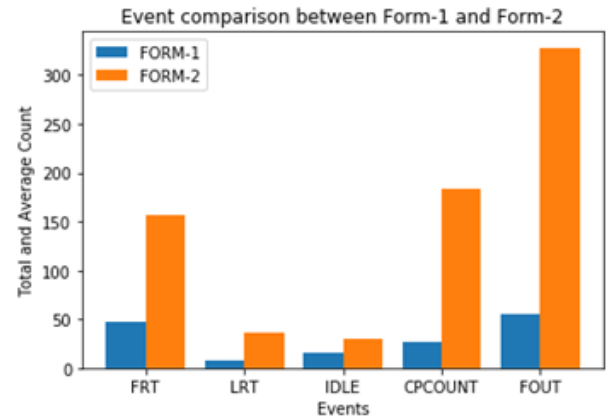


Fig 5. Comparative analysis of events for Form-1 and Form-2

3.4 Data Labelling

The analysis in the section above indicates that unethical actions are taken by some students during the examination. As a result, we label the events data for Form-2 to build a model that can predict such actions. Form-2 is chosen as the questions here are course based and students do such actions more in these situations. Three observations were identified which are the class values. These are as follows-

1. A student is using another tab of the browser or another application to answer, which we label as 'F'.
2. A student is using a book or another device to answer, which are we labelled as 'FD'.

²<https://www.kaggle.com/adnansami1/behavioural-biometric-data-for-student-interaction>

3. A student answers on his own without any help, which are we marked as 'N'.

MCQ and face to face interview marks were used to deduce these observations. If a student obtains a mark on a question in the MCQ assessment but fails to answer it during the interview and the count for FOUT and CPCOUNT is zero, then it can be considered that the second observation occurs. Moving on, if a student answer a question correctly during the MCQ and interview and count for FOUT and CPCOUNT is zero or if a student did not answer a question correctly during the MCQ and interview and count for FOUT and CPCOUNT is zero, then third observation occurs. The rest of the dataset observations can be said to be the first one. The mouse behaviour was also added with the dataset. During data extraction a large number of x and y coordinates are obtained for the students when they answer the questions. To reduce complexity, the mean, variance, and standard deviation of the x and y coordinates for each question were collected. Later the Student ID (SID) and Question ID (QID) to cross-link this information were used with the labelled dataset. At this point the dataset contains events (navigation) mouse or touch and keyboard movement data that can be used to predict illegal actions.

3.5 ML Model Building

Labelling the data to build a model for illegal action prediction during the examination was necessary. As mentioned earlier, the instances were labelled the dataset into three class values. After labelling the following result obtained

Table 2. Total number of instances for each observation

Class Value	Total Number of Instances
F	231
FD	82
N	92
Total Number of Instances	405

From Table-2 it is observed that the dataset is not balanced. The number of instances for 'FD' and 'F' is very low compared to that for 'N'. This is an under-sampling problem [30] and may lead to a degraded performance for predictions [31]. As a solution, a Synthetic Minority Oversampling Technique (SMOTE) [32] was applied through the python SMOTE package [33] and obtain a modified dataset containing the following number of instances for the different observations-

Table 3. Total number of instances for each observation after applying SMOTE

Class Value	Total Number of Instances
F	231
FD	231
N	231
Total Number of Instances	693

As the dataset is balanced, four ML algorithms were applied which are Decision Tree, Random Forest, K-Nearest

Neighbour and Naive Bayes to construct models that can predict unethical actions during an examination using student interaction.

4. DISCUSSION

The critical analysis on the various results and observations on number of interaction events and prediction model through the study is discussed in this section.

4.1 Events

Various observations were obtained during analysing the different events performed by each student while participating the MCQ test that helped us to visualize the findings.

4.1.1 Change in Answer

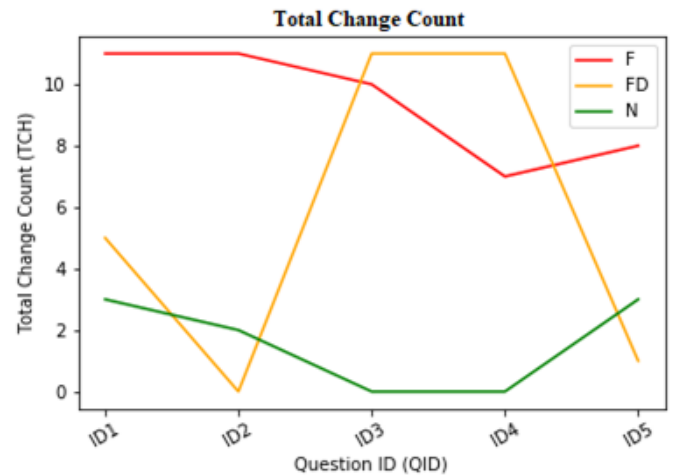


Fig 6(a). Question Wise Change Graph

The figure above shows the total number of times the students belonging to each class change the answer to a question. It is observed that the maximum number of changes occur when the students use another device to answer. However, overall, more changes are made when the students are using the same device to answer. When the students are answering on their own, the number of changes is considerably less compared to the other two class values.

4.1.2 Change in Response Time

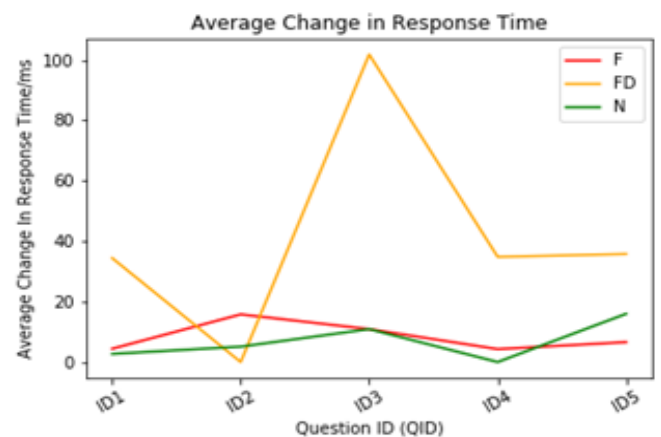


Fig. 6(b). Average Change in Response Time Graph

The information about the FRT was extracted which is the first time a student answered a particular question. If he/she changed the answer again, then LRT was collected which is the last time the student answered the same question. This information was used to find the change in response time. If the LRT of a particular question is 0, the change in response time was considered as 0. This is because no change is made after the first response. If the value of LRT is more than zero, then the FRT from LRT was subtracted to calculate the change in response time. Finally, the average change in response time was calculated for each question and plot the graph shown in Fig. 6(b). From the graph it is seen that the average change in response time is higher for students who used another device to answer and lower for students who answered on their own. For the students who answered using the same device, this time is higher than those who answer on their own but considerably lower than those who used another device.

4.1.3 Idle Time

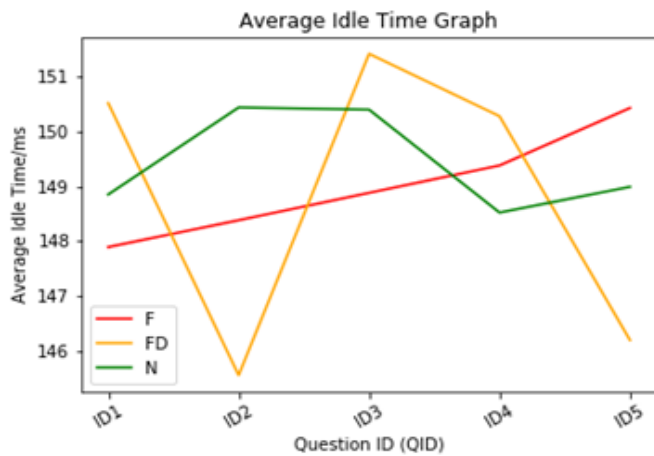


Fig 7. Average Idle Time Graph

The average idle time for each question belonging to the individual class values are calculated and the graph in Fig. 7 is plotted. It is seen that the graph depicting the students who answer by them have higher values compared to one that represents the students who answer using the same device. The graph for the students using another device shows variation.

4.2 Mouse or Touch Behaviour

This section discusses the mouse movement pattern that we observe during the study. The mean of x and y coordinates was calculated for each student based on the class values 'F', 'FD' and 'N'. Then scatter graphs were plotted to visualize how the mouse or touch movements vary. Finally, the same graph for Form-1 was plotted where it was considered without any unethical actions are present. The figures below show the mouse or touch behaviour for the different class values.



Fig. 8(a). Mouse Movement Coordinates for F

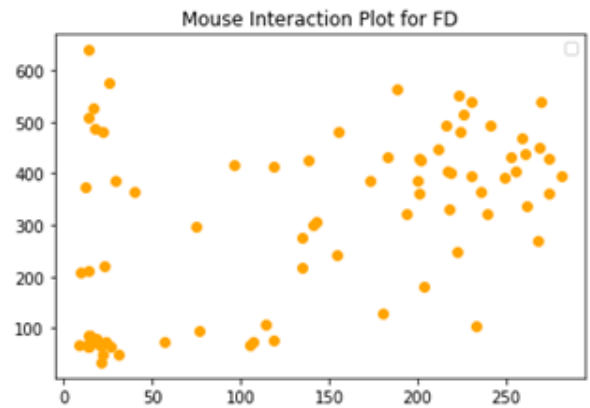


Fig. 8(b). Mouse Movement Coordinates for FD

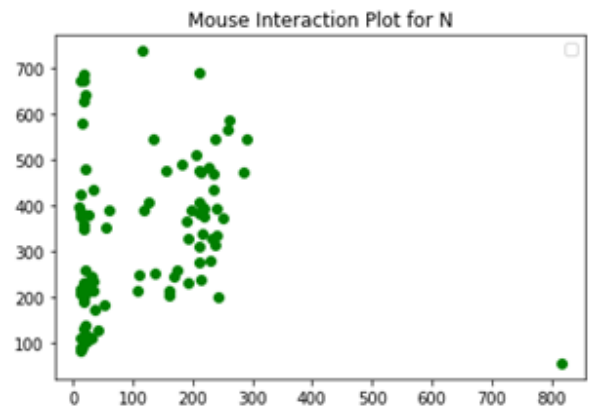


Fig. 8 (c). Mouse Movement Coordinates for N

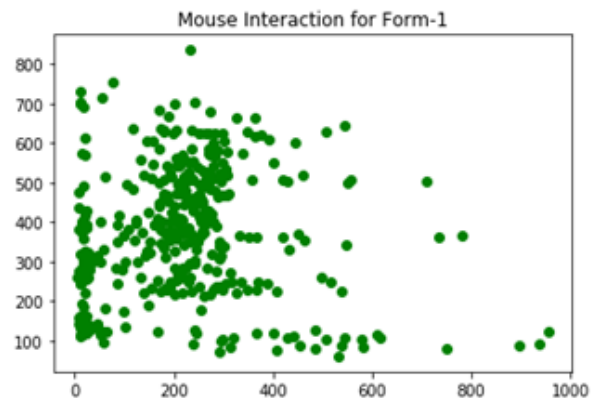


Fig. 8 (d). Mouse Movement Coordinates for Form-1

From Fig.8(a) and Fig.8(b) which depicts the presence of illegal actions it is seen that the students hovered their mouse or touch almost all over the screen. The probable cause might be they were confused about the answer, or they lack proper knowledge to answer. In this case they continuously hover their mouse or touch finally taking help from other sources to answer. On the other hand, from Fig.8(c) it can be observed that these movements are confined within a specific region. This specifies that when a student knows the answer to a particular question the mouse or touch movement is considerably less which is probably due to hovering the mouse or touch from one question to another. In Fig.8(d) it is also visible that most of these movements are confined within a particular region. This proves the fact that less movements are done when a student knows the answer to a question.

4.3 Model Analysis and Evaluation

This section discusses the result of the prediction model with and without SMOTE. Table 4 shows the results obtained without SMOTE and Table 5 shows the results after applying SMOTE.

Table 4. Results of Models built without SMOTE

	Decision Tree	Random Forest	KNN	Naïve Bayes
Accuracy	0.864	0.839	0.629	0.703
True Positive Cases	10	9	4	4
True Negative Cases	44	33	40	33
False Positive Cases	0	0	3	6
False Negative Cases	0	2	6	0

Table 5. Results of Models built with SMOTE

	Decision Tree	Random Forest	KNN	Naïve Bayes
Accuracy	0.841	0.906	0.726	0.654
True Positive Cases	31	37	35	29
True Negative Cases	48	38	37	31
False Positive Cases	2	3	7	5
False Negative Cases	2	0	1	1

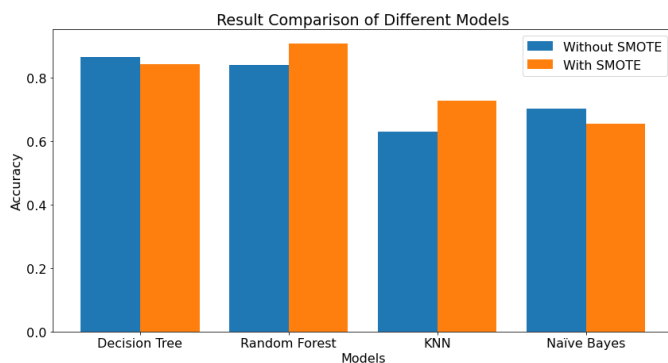


Fig 9. Result Comparison of Different Models

Fig. 9 illustrates the result comparison of 4 different ML prediction models including Decision Tree, Random Forrest KNN and Naïve Bayes. Since the dataset was not balanced as mentioned in section 3.5, SMOTE was used to balance the

data. In order to understand the performance, the prediction models were trained with balanced and unbalanced dataset. From the accuracy and confusion matrix it is in Table 4 and Table 5, it was observed Random Forest outperforms all the other models with an accuracy of almost 91% when SMOTE was applied and hence we propose it for model building.

4.3 Model

Based on the observations from the earlier parts of this section the following hypothesis can be made on the behavioural actions during an online examination-

1. Students who use the same device to answer usually change the answer to a question more and keep IEMT idle for a significant amount of time.
2. Students who use another device to answer show significant fluctuations for response and idle time and change in answer.
3. Students who answer on their own take significantly lower response time and change the answer less frequently.
4. The mouse or touch movement is scattered all over the screen for students who use the same or different device to answer. On the other hand, these movements are more confined within a particular region for students who answer on their own.

5. CONCLUSION AND FUTURE WORK

After completing all the analysis mentioned above, we conclude that the students show a significant change in the way they interact with a computer when using unethical means to give an examination. Using the findings and analysis, a prediction model can be built that can identify such actions during an online examination based on certain events and mouse or touch, and keyboard interaction specific to each student. Our proposed model can be integrated with an online learning environment that can detect whether cheating is practised during an examination.

In the future we would like to extend this work further for more analysis. It is quite determined from the literature given above that every student wants a better performance in class which sometimes forces them to show such behaviour. So just restricting them from these activities will not be enough. Rather we need to study in detail on how we can design the examinations so that the students are motivated to answer on their own. As a result, we hope to use this study and see how we can change the examination and the elements included in it to propose a better evaluation approach that can reduce the tendency of using unethical actions among students. Keeping this in mind, in the future, we would like to advance toward the development of a proper environment that will detect the potential attempts of unethical practice during an online examination.

REFERENCES

- [1] Revett, Kenneth. Behavioral biometrics: a remote access approach. John Wiley & Sons, 2008, ISBN: 978-0-470- 51883-0.
- [2] Alsaadi, Israa Majeed. "Study On Most Popular Behavioral Biometrics, Advantages, Disadvantages And Recent Applications: A Review." International Journal of Scientific & Technology Research, 10 (1) (2021): 15-21, ISSN: 2277-8616, pp. 15-21, Vol. 10, No.1, January 2021, Published by International Journal of Scientific & Technology Research.

- [3] Li, Wenjuan, Weizhi Meng, and Steven Furnell. "Exploring touch-based behavioral authentication on smartphone email applications in IoT-enabled smart cities." *Pattern Recognition Letters* 144 (2021): 35-41. DOI:10.1016/j.patrec.2021.01.019.
- [4] Mekruksavanich, Sakorn, and Anuchit Jitpattanakul. "Convolutional neural network and data augmentation for behavioral-based biometric user identification." *ICT Systems and Sustainability*. Springer, Singapore, 2021. 753-761, Online ISBN: 978-981-15-8289-9, DOI: 10.1007/978-981-15-8289-9_72.
- [5] Dayal, Aveen, et al. "Design and implementation of deep learning based contactless authentication system using hand gestures." *Electronics* 10.2 (2021): 182.
- [6] Monaro, Merylin, et al. "Spotting faked 5 stars ratings in E-Commerce using mouse dynamics." *Computers in Human Behavior* 109 (2020): 106348, DOI:10.1016/j.chb.2020.106348.
- [7] Katerina, Tzafilikou, and Protogeros Nicolaos. "Mouse behavioral patterns and keystroke dynamics in End-User Development: What can they tell us about users' behavioral attributes?." *Computers in Human Behavior* 83 (2018): 288-305.
- [8] Nishiuchi, Nobuyuki, and Seima Aoki. "Study on behavioral biometrics to predict user's interest level using web access log." 2017 International Conference on Biometrics and Kansei Engineering (ICBAKE). IEEE, 2017.
- [9] Van Balen, Nicolas, Christopher T. Ball, and Haining Wang. "A behavioral biometrics based approach to online gender classification." *International Conference on Security and Privacy in communication Systems*. Springer, Cham, 2016.
- [10] Tzafilikou, Katerina, and Nicolaos Protogeros. "Monitoring Mouse Behavior in E-Learning Activities to Diagnose Students' Acceptance Items of Perceived Usefulness and Ease of Use." *European Educational Researcher* 3.1 (2020): 21-27, ISSN: EISSN-2517-6323.
- [11] Lim, Yee Mei, Aladdin Ayesh, and Martin Stacey. "Using mouse and keyboard dynamics to detect cognitive stress during mental arithmetic." *Science and Information Conference*. Springer, Cham, 2014.
- [12] Carneiro, Davide, et al. "Using mouse dynamics to assess stress during online exams." *International Conference on Hybrid Artificial Intelligence Systems*. Springer, Cham, 2015, DOI: 10.1007/978-3-319-19644-2_29.
- [13] Salmeron-Majadas, Sergio, Olga C. Santos, and Jesus G. Boticario. "An evaluation of mouse and keyboard interaction indicators towards non-intrusive and low cost affective modeling in an educational context." *Procedia Computer Science* 35 (2014): 691-700.
- [14] Bilen, Eren, and Alexander Matros. "Online cheating amid COVID-19." *Journal of Economic Behavior & Organization* 182 (2021): 196-211.
- [15] Bawarith, Razan, et al. "E-exam cheating detection system." *International Journal of Advanced Computer Science and Applications* 8.4 (2017): 176-181.
- [16] Atoum, Yousef, et al. "Automated online exam proctoring." *IEEE Transactions on Multimedia* 19.7 (2017): 1609-1624.
- [17] Jalali, Khatereh, and Fakhroddin Noorbehbahani. "An automatic method for cheating detection in online exams by processing the student's webcam images." *Proc. 3rd Conf. Elect. Comput. Eng. Technol.(E-Tech)*. 2017.
- [18] Abdullah, Sarah Ali, Tariq Adnan Fadil, and Noor Ahmed. "Online Examination System (Electronic Learning)." *Research in Intelligent and Computing in Engineering*. Springer, Singapore, 2021. 309-32.
- [19] Mallik, Sharthak, et al. "Multi-factor Authentication-Based E-Exam Management System (EEMS)." *Proceedings of International Conference on Frontiers in Computing and Systems*. Springer, Singapore, 2021.
- [20] Hannan, Tufayl A., et al. "Designing and running an online Objective Structured Clinical Examination (OSCE) on Zoom: A peer-led example" *Medical Teacher* (2021): 1-7.
- [21] Syam, Mahmoud I. "The factors that lead to cheating in exams and how to avoid them." *GSTF Journal on Education (JEd)* 1.2 (2014).
- [22] Dusu, Paul Bot, et al. "Management of Re-Occurring Cases of Examination Malpractice in Plateau State Collage of Health Technology Pankshin, Nigeria." *Journal of Education and Practice* 7.6 (2016): 38-43, ISSN: ISSN-2222-1735.
- [23] Open Education Database, 8 Astonishing Stats on Academic Cheating, OEDB, Houston, TX, USA, 2010.
- [24] Elias, Rafik Z. "THE EFFECT OF MACHIAVELLIANISM ON BUSINESS STUDENTS PERCEPTION OF CHEATING." *Academy of Educational Leadership Journal* 19.1 (2015): 175.
- [25] Choi, Charles Q. "The pull of integrity." *ASEE prism* 18.7 (2009): 29.
- [26] Open Education Database, 8 Astonishing Stats on Academic Cheating, OEDB, Houston, TX, USA, 2010.
- [27] Jensen, Simon Holm, Magnus Madsen, and Anders Møller. "Modeling the HTML DOM and browser API in static analysis of JavaScript web applications." *Proceedings of the 19th ACM SIGSOFT symposium and the 13th European conference on Foundations of software engineering*. 2011.
- [28] Muzaffar, Abdul Wahab, et al. "A Systematic Review of Online Exams Solutions in E-Learning: Techniques, Tools, and Global Adoption." *IEEE Access* 9 (2021): 32689-32712.
- [29] Storozeva, S. P., E. G. Strukova, and R. D. Shilin. "Communication practices to maintain the student contingent in the distance education system." *Professional education in the modern world* 10.4 (2021): 4190-4201.
- [30] Yen, Show-Jane, and Yue-Shi Lee. "Under-sampling approaches for improving prediction of the minority class in an imbalanced dataset." *Intelligent Control and Automation*. Springer, Berlin, Heidelberg, 2006. 731-740.
- [31] Batista, Gustavo EAPA, Ronaldo C. Prati, and Maria Carolina Monard. "A study of the behavior of several methods for balancing machine learning training data." *ACM SIGKDD explorations newsletter* 6.1 (2004): 20-29.
- [32] Chawla, Nitesh V., et al. "SMOTE: synthetic minority over-sampling technique." *Journal of artificial intelligence research* 16 (2002): 321-357.
- [33] Kovács, György. "Smote-variants: A python implementation of 85 minority oversampling techniques." *Neurocomputing* 366 (2019): 352-354.
- [34] Martin, F., Sun, T., & Westine, C. D. (2020). A systematic review of research on online teaching and learning from 2009 to 2018. *Computers & Education*, 159, 104009. <https://doi.org/10.1016/j.compedu.2020.104009>
- [35] Curran, K., Middleton, G., & Doherty, C. (2011). Cheating in Exams with Technology. *International Journal of Cyber Ethics in Education*, 1(2), 54–62. <https://doi.org/10.4018/ijcee.2011040105>
- [36] Turner, S. W., & Uludag, S. (2013). Student perceptions of cheating in online and traditional classes. *Proceedings - Frontiers in Education Conference, FIE*, (October 2013), 1131–1137. <https://doi.org/10.1109/FIE.2013.6685007>
- [37] Tiong, L. C. O., & Lee, H. J. (2021). E-cheating Prevention Measures: Detection of Cheating at Online Examinations Using Deep Learning Approach--A Case Study. *arXiv preprint arXiv:2101.09841*.