UNIVERSITY OF THE EAST
Manila
College of Computer Studies and Systems

11/18/21

This is to certify that the proposal manuscript entitled **Tracking Input Devices to Detect Cheating Using Machine Learning Techniques** by the following students:

MR. LOUIS ANGELO D.R. ALTOVEROS
MR. VIRGILIO A. LOPEZ JR.
MR. RENZ DARRYL M. MANANGAT
MR. RALF MARTIN R. TABO
MR. JOHN CHRISTIAN B. TIOSAN

in compliance for the thesis defense are found to be complete and checked by the undersigned.

PROF. MELIE JIM SARMIENTO
Thesis Adviser

Tracking Input Devices to Detect Cheating Using Machine Learning Techniques

2021-1-PTCS-84

Table of Contents

I. Background of the Study

Problems and Opportunities

Online classes have been the temporary norm for education due to the state of the world brought by the pandemic. Lectures, training, seminars, and examinations are conducted through online means as a response to our current predicament. Advantages brought by this setup include the students, and teachers being in the comfort of their own homes, which gives them the luxury of taking these classes at their own pace. However, with upsides come downsides, and with that, the most prominent may be cheating on the exams where the teacher or proctor has limited surveillance over the student. In the academic setting, cheating refers to any type of illicit action for committing fraud (Bucciol et al., 2020, as cited by Sokout et al., 2020). Additionally, it involves the unauthorized use of information, materials, devices, sources, or practices in completing academic activities (Northern Illinois University, n.d.). This same problem has been in question for a long time as online tests conducted for online certificates are most likely being cheated upon due to its no surveillance nature.

Now, cheating may be a desperate last resort for a student who struggles with his studies, especially now for students who are not exactly favored to the current setup of classes. However, according to an article by middleearthnj (2014) entitled “Cheating in School: Facts, Consequences, and Prevention,” it is not simply just that. According to this article, cheating may lead to some consequences that may prove to be bigger problems later on. All it takes is for a student to cheat once, and he may find himself doing it again, and again in the future. Cheating also blocks out learning, and this not only refers to the learning of the examination's topic at hand. Since lessons are usually made so that they are progressive, students who failed to study a lesson, then resort to cheating, will find it even harder to study following lessons as they are usually a continuation of previous lessons. This problem not only stops in school, as experts have agreed that students who cheat in school are more likely to be dishonest in their future jobs, which will surely bring down the quality of workers as the years pass.

Statement of the Problem

Several kinds of research have been produced that aims to detect cheating in online assessments. The methods used by these studies vary. Some utilized the web cameras of examinees to monitor them while taking the exam (Bilen & Matros, 2021; Karim et al., 2014; *Webcam Proctoring for Online Exams,* n.d.), while other studies used eye-tracking algorithms to detect cheating behavior (Purnama et al., 2020; Javed & Aslam, 2013). These techniques have been widely studied and have a high level of support, however, they give the disadvantage of being too costly, they can be invasive, or they reduce the efficiency of online testing (Diedenhofen & Musch, 2016; Purnama et al., 2020). Diedenhofen and Musch (2016) state that the methods currently available are either limited to specific testing situations or result in significant and prohibitive costs.

A relatively unexplored area for detecting potential cheaters in online examinations is mouse tracking (Sokout et al., 2020). In this technique, the mouse behaviors of the participants are recorded to predict their actual behaviors. Tracking mouse behaviors has been utilized for various reasons. These include using it as an indicator of cognitive conflict in social dilemmas (Kieslich & Hilbig, 2014), quantifying and detecting difficulty indicators in web surveys (Horwitz et al., 2020), and improving data quality by assessing measurement error (Yan & Olson, 2013, as cited by Diedenhofen & Musch, 2017). However, only a few studies have used mouse tracking as an indicator of cheating behavior in online assessments.

One such study was conducted by Sokout et al. (2020) where they recorded specific mouse activities to predict whether an examinee cheated in an online examination. The reported sensitivity rate of their model was 94% while the specificity was 87%. These accuracy rates are already very impressive in terms of predictive power. However, there is still room for error in terms of their predictive power, and this is where the problem lies. Having a gap in the accuracy means there is a chance that the system might be wrong; an examinee who was flagged to be cheating may in fact be honest, or vice versa. Additionally, examiners become reluctant to report cheating in online exams when there is inconclusive evidence (Bilen & Matros, 2020). Therefore, there is a need to find a new method that will return higher accuracy metrics in terms of either sensitivity or specificity and thus, eliminate doubts on whether the system could be wrong.

As such, this study focuses on producing a model that will outperform the baseline in terms of the accuracy metrics: sensitivity and specificity. Currently, the model developed by Sokout et al. (2020) returns the highest prediction rates using mouse tracking as an indicator of cheating. Hence, this will be used as the baseline with which the accuracy of the model in the proposed study will be compared.

For this study, instead of using only the mouse behavior, the more general term “input devices” will be used. Computer Hope (2021) defines an input device as “any hardware device that sends data to a computer”. Examples of this include the mouse, keyboard, joystick, microphone, camera, etc. However, in this study, the term “input devices” is limited only to the following devices: mouse, keyboard, and trackpad.

Research Questions

The proponents have come up with a machine learning-based solution that detects suspicious activities of students while in an online examination.

Specifically, the study aims to answer the following questions:

1. How well can we detect cheating in online examinations based only on the examinees' input device data in terms of:
   a. Specificity
   b. Sensitivity
2. What specific input device data can be an indicator of cheating?
3. How useful will a system that uses input device data be for examiners?

Hypothesis

H_0: The model does not outperform the baseline in either sensitivity or specificity by at least 1 point.

H_1: The model outperforms the baseline in either sensitivity or specificity by at least 1 point.

Beneficiaries

The main beneficiaries of this study include academic institutions that are still having the online class setup, however, it is not limited to them as even though it is face-to-face classes, some schools still tend to have quizzes on an online platform. Therefore, this algorithm can still be used even after this pandemic is over and the state of all classes went back to the old norm. Furthermore, websites issuing online certificates after completing a certain test can also make use of this in order to prove the integrity of their takers. Ultimately, the true benefactors of this research are the students themselves as the presence of an anti-cheating function will make them resort to studying on a test rather than desperately thinking of ways on how to bypass the system.

II. Analysis of Related Literature and Studies and Conceptual Framework

Cheating in the online setting

With the COVID-19 pandemic affecting all countries globally, educational institutions were forced to move their operations online. Traditionally, educational institutions rely on proctored exams to keep the integrity of these examinations. However, with online education, exams are typically unproctored, allowing examinees to use any method of cheating, including but not limited to, looking at their notes, internet search, and assistance from other persons (Bilen & Matros, 2021). In fact, a study by Dendir & Maxwell (2020) compared the performance of examinees in unproctored online assessments to that in proctored online assessments. They found that the performance of the participants in the online assessments was significantly lower when proctoring was introduced, as compared to when there was no direct proctoring. Thus, they concluded that cheating is more prevalent in the absence of proctoring and suggested that some form of direct proctoring is perhaps the most effective way of mitigating cheating in online assessments (Dendir & Maxwell, 2020).

However, even when proctoring is introduced, cheating is still prevalent and poses a serious problem. In the US, many universities reported widespread cheating in the online Advanced Placement examinations that took place in 2020, when universities were forced to conduct the exams online due to the pandemic (Bilen & Matros, 2021).

Students themselves have conveyed that more cheating occurs in online examinations, typically by way of looking up answers on the internet, using social media, or seeking help from friends (Best & Shelley, 2018). When presented with questions that could be looked for on the internet, 22% of the participants indicated that they had used the Internet to identify the correct solution to at least one of the four questions presented (Jensen & Thomsen, 2013). During the aforementioned 2020 Advanced Placement examinations, many students took advantage of having immediate access to Google search (Bilen & Matros, 2021).

Examining the rampancy of cheating in online chess tournaments, Bilen & Matros (2021) predicted that online cheating in schools and universities will only get much worse. This prevalence of cheating in online assessments and examinations poses a serious problem to learning and education. With such inflated performance, it becomes much more difficult to gauge the learning of the exam-takers. Some even argue that scores on online assessments don't actually reflect learning because they are possibly tainted by cheating (Dendir & Maxwell, 2020; Arnold, 2016) and that the likelihood of cheating has a negative relation to academic progress (Arnold, 2016).

Methods used to prevent cheating

Realizing how much cheating occurs in the online setting, educational institutions have employed various methods to detect and prevent cheating in examinations. One method is by requiring exam takers to turn on their web cameras and having a test administrator supervise the test session (Bilen & Matros, 2021; Karim et al., 2014; *Webcam Proctoring for Online Exams,* n.d.). Using this method, the test administrator can monitor examinees and record them as they take the exam. The recording can then be viewed later by the administrator to ensure the integrity of the test. This has been proven to be effective in reducing the occurrence of cheating in online assessments. However, the problem with monitoring exam takers using their webcams is that it can produce negative reactions among the takers (Karim et al., 2014). Camera use can be perceived as conflicting with privacy rights, thus it may be viewed as invasive (Bilen & Matros, 2021). Furthermore, remote proctoring—having a test administrator supervise the test session over the Internet—is costly and reduces the desired efficiency of online testing (Diedenhofen & Musch, 2016).

Another approach is by administering an additional verification test to detect inconsistencies in the scores of the candidate and confirm whether the candidate engaged in illegal practices or not. This additional test is taken offline in a supervised environment (Tendeiro et al., 2013). The disadvantage of this method is that additional testing is necessary, thereby greatly reducing the cost-efficiency of unproctored Internet testing (Diedenhofen & Musch, 2016).

Arguably, one of the most accurate methods for recording the viewing activity of users is eye-tracking (Purnama et al., 2020). This method involves a camera analyzing the eye movements of a human and then determining whether they had cheated based on their eye movements (Javed & Aslam, 2013). Much of the problem with using eye-tracking is that the financial cost with this method is very high, thus the technology is confined to lab environments. It also requires separate eye-tracking hardware which is usually very expensive (Purnama et al., 2020). Furthermore, most online examinations and assessments do not require their students' cameras to be opened when taking the exam, such as in Massive Open Online Courses (MOOCs). As such, eye-tracking methods can't be used most of the time to detect cheating in online assessments.

Mouse tracking in similar literature

Fernández-Fontelo et al. (2021) computed nine (9) mouse-tracking measures and approached the task of predicting question difficulty in web surveys with machine learning.

Type	Measure	Definition
Time	Response time (RT)	Time from page load until response submission
	Initiation time	The duration from page load until the first recorded mouse movement occurred
Hovers	Number of hovers	Number of periods without movement exceeding a minimum duration threshold
	Overall duration of hovers	Total time of all periods without movement exceeding a minimum duration threshold
Distance	Total distance	Euclidean distance traveled by the mouse
Derivatives	Maximum velocity	Maximum movement velocity
	Maximum acceleration	Maximum movement acceleration
Flips	*x* flips (horizontal flips)	Number of changes in movement direction along horizontal axis
	y flips (vertical flips)	Number of changes in movement direction along vertical axis

Note. All hover measures exclude a potential initial phase without mouse movement (reflected in the initiation time).

Table 1: Mouse-tracking measures (Fernández-Fontelo et al., 2021)

For preprocessing, they filtered out questions for which no or incomplete mouse movements were recorded, considering intermittent connection problems. Page reloads also indicated that prior mouse actions recorded will be discarded. They considered logistic regression, tree-based models (classification trees, random forest, and gradient boosting), support vector machines, and single hidden layer back-propagation networks (a kind of neural network). Nested cross-validation was used for parameter tuning and model evaluation. It was also found that personalizing mouse-tracking measures resulted in higher predictive accuracy, which shows the importance of considering habitual behaviors across the respondents.

Ensemble methods such as gradient boosting and random forest were found to improve weaker models' predictive power by reducing variance and bias. Their best learning models showed moderately high accuracies, roughly between 59% and 65%. They noted that the limited classification performance might come from measuring only the summaries of the information in the mouse movements. They suggest the use of full mouse movement trajectories, or the collection of additional mouse information such as the respondents' click data and changes in the response options.

To detect and prevent cheating in online tests, Diedenhofen and Musch (2017) developed PageFocus, a Javascript plugin that detected when a participant switched to another window or tab while taking the test. To determine the sensitivity and specificity of the plugin, they conducted a study wherein 115 lab participants and 186 online participants were asked to complete a test. The test consisted of 16 difficult multiple-choice questions but was easy to look up. Half of the participants were invited to search for the answers online while the other half were not. After completing the test, participants were asked for each question whether they had looked up the solution or not. While taking the test, the PageFocus plugin recorded the occurrence of page switching along with the timestamp of when it occurred.

The following tables show the results of the study conducted by Diedenhofen & Musch (2017).

Instructions		Lab		Web	
Cheating	At least one page-defocusing event	54	(98.18 %)	84	(92.31 %)
	No page-defocusing event	1	(1.82 %)	7	(7.69 %)
	∑	55	(100.00 %)	91	(100.00 %)
Control	At least one page-defocusing event	2	(3.33 %)	18	(18.95 %)
	No page-defocusing event	58	(96.67 %)	77	(81.05 %)
	∑	60	(100.00 %)	95	(100.00 %)

Table 2: Summary of page-defocusing events registered

Table 2 shows the number of participants that triggered none or at least one page-defocusing event throughout the test. For the cheating group, 54 lab participants and 84 web participants triggered at least one page-defocusing event while 1 lab participant and 7 web participants did not trigger a single page-defocusing event.

For the control group, which was asked to answer the test honestly, 2 lab participants and 18 web participants triggered at least one page-defocusing event while 58 lab participants and 77 web participants did not trigger a single page-defocusing event.

Instructions		Lab		Web	
Cheating	Self-reported cheating	54	(98.18 %)	86	(94.51 %)
	No self-reported cheating	1	(1.82 %)	5	(5.49 %)
	∑	55	(100.00 %)	91	(100.00 %)
Control	Self-reported cheating	2	(3.33 %)	3	(3.16 %)
	No self-reported cheating	58	(96.67 %)	92	(96.84 %)
	∑	60	(100.00 %)	95	(100.00 %)

Table 3: Summary of Participants' Self-Reports

Table 3 shows the number of participants who reported that they cheated on at least one question and those that denied cheating. For the cheating group, 54 lab participants and 86 web participants reported cheating on at least one question while 1 lab participant and 5 web participants denied cheating.

On the other hand, for the control group, 2 lab participants and 3 web participants reported cheating while 58 lab and 92 web participants denied they cheated on the test.

Since the time stamp for each occurrence of page-defocusing events was recorded while the participants were taking the test, these were compared to the self-reports of the participants on whether they had cheated on a particular question or not. With this, they were able to determine the sensitivity and specificity of the page-defocusing events captured by PageFocus for both the lab sample and the web sample. For the lab, the sensitivity and specificity rates were 99.54% and 97.29%, respectively; for the web sample, the sensitivity and specificity were 96.64% and 94.56%, respectively.

However, the problem with relying only on window switching events is that in real-world situations, there could be a number of reasons for these defocusing events occurring aside from cheating. These could be in the form of misclicks, pop-ups appearing, or other important actions. In fact, Diedenhofen and Musch (2017) stated that some of the page-defocusing events produced by participants who switched to another browser tab were due to reasons unrelated to cheating.

Furthermore, the plugin doesn’t register an action as cheating when the total duration of a pair of page defocusing and refocusing events is less than 3 seconds. Thus, an examinee could repeatedly switch windows or tabs and the plugin wouldn’t detect these events as long as they return to the exam window within 3 seconds.

Another study was conducted by Sokout et al. (2020) where they also tracked the mouse behavior of students on a midterm examination. Using Support Vector Machine (SVM) as a supervised learning model, they were able to classify or predict when students are engaging in suspicious activities during the exam (potential cheaters) from the other honest students. The classification is mainly split into two, the students that only have a single tab from start to finish are labeled as honest students, while those with multiple tabs are labeled as potential cheaters.

Event Loggings	Tab-active	Active Duration (second)	Tab-inactive	Inactive Duration (second)	Left Clicks	Right Clicks	Copy	Paste	Double Clicks
Total	216988	41817	6673	2610	12798	15	30	18	83
Minimum	1411	158	3	0.03	11	0	0	0	0
Maximum	12984	1961	1041	647.2	2127	8	6	4	23
Mean	4931.5	950.4	151.7	59.3	290.9	0.34	0.68	0.4	1.1
Standard deviation	2528.8	414	280.1	144	573.9	1.3	1.6	0.9	4.8

Table 4: Summary of Collected Data

Table 4 indicates the events recorded for the study. These include the number of tabs active, the amount of time that a student is active or inactive in the exam tab (the tab that the quiz is on), and other mouse functions such as left-click, right-click, double-click, and copy and pastes. However, they only used four events as predictors for the model: *inactive duration*, *copy/cut*, *paste*, and *double-click*. They then used SVM with Linear Kernel to model and classify the data. The model was then evaluated using k-fold cross-validation, where they then measured various accuracy metrics to determine the performance of the model.

Measures	Fold 1 (%)	Fold 2 (%)	Fold 3 (%)	Fold 4 (%)	Sum of Folds (%)
Accuracy	0.82	0.91	1	0.82	0.89
Sensitivity	1	1	1	0.75	0.94
Specificity	0.71	0.89	1	0.86	0.87
Precision	0.67	0.67	1	0.75	0.77
F-score[a]					0.85

Table 5: Accuracy metrics of Sokout et al. (2020)

Table 5 indicates the details of the obtained accuracy metrics of the model. For accuracy, the average rate was 89%; the sensitivity was 94%; specificity was 87%; precision was 77%; the F-score was 85%.

Throughout this proposal, the baseline accuracy metric that will be used will be either the sensitivity or specificity since the former indicates the accuracy of examinees committing illicit behaviors while the latter indicates the accuracy of examinees *not* committing illicit behaviors.

Both researches have impressive accuracy in terms of predicting whether an examinee cheated or not. However, there is still room for error in terms of their predictive power, and this is where the problem lies. Having a gap in the accuracy means there is a chance that the system might be wrong; an examinee who was flagged to be cheating may in fact be honest, or vice versa. Additionally, examiners become reluctant to report cheating in online exams when there is inconclusive evidence (Bilen & Matros, 2020). Therefore, there is a need to find a new method that will return much higher prediction rates in terms of either sensitivity or specificity and thus, eliminate doubt on whether the system could be wrong.

In their study, Sokout et al. (2020) also considered only four events as their predictors. Results from the study of Diedenhofen and Musch (2017) show how *blur* and *focus* are effective in predicting whether examinees cheated or not. By adding these two events to the features that will be used in this study, it is possible to increase the predictive power of the model.

Logistic regression is a classifier that always maps its predictions between 0 and 1. This is the reason why this model is one the most popular choices for classification tasks (Kleinbaum & Klein, 2010).

Additionally, using different models and methods can also improve the prediction performance as compared to using only a single model. Methods and models such as ensemble learning and support vector machines can be useful here.

Ensemble learning combines different models in order to increase overall prediction performance. This method may also be used for models to compensate for the lapse of another (Sagi & Rockah, 2018). Although it can still be useful to apply it to a set of already well-performing models (scikit-learn, n.d.). Due to its promising results, it has also expanded to deep learning which similarly improved models (Ganaie et al., 2021). Thus, this method will also be employed in this study to further increase the overall prediction performance.

Support Vector Machines (SVM) are machine learning algorithms that are employed in regression and classification problems, which allows the determination of the relationship between the target and independent variables. SVM can work effectively on datasets of any size and is one of the most popular supervised learning algorithms because it constructs its analysis by looking at a very extreme case that is close to the boundary (Sokout et al., 2020).

Conceptual Framework

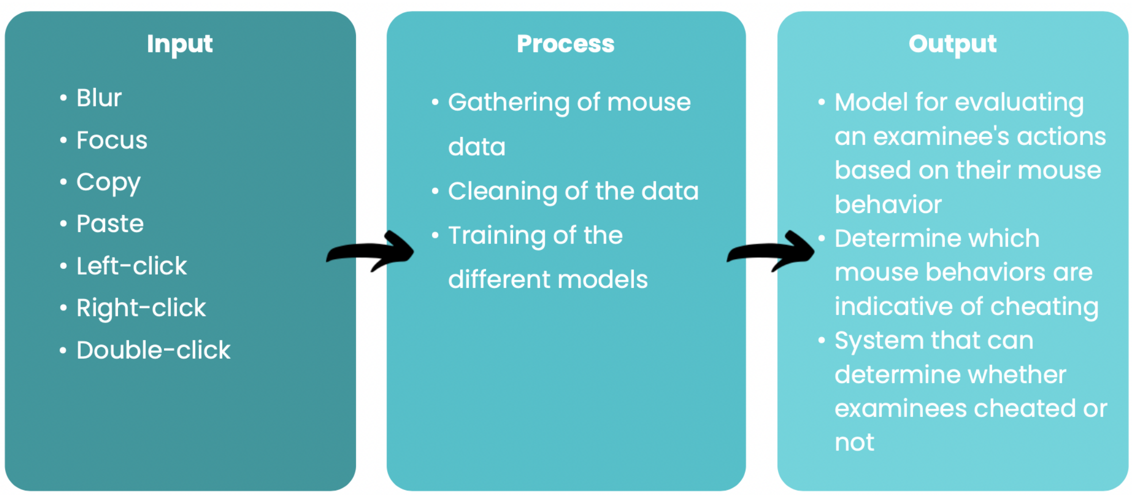


Figure 1: Conceptual Framework

Figure 1 presents the conceptual framework of the study. It consists of the independent variables or features as the input. These features pertain to the specific input device data gathered from participants. Blur refers to when the participant transfers to another tab or window while they are taking the exam; focus happens when the participant returns to the exam window/tab; copy and paste occur when the participant copies/pastes a content; left-click and right-click occurs when the participant uses the left and right buttons of their mouse; double-click is when the left button is clicked twice in quick succession.

As can be seen in the figure, these data will first be cleaned in order to improve quality and remove inaccurate data. A model will then be trained to obtain a classifier using the data gathered as features; the predicted class will either be cheating or not cheating. The primary models to be used in order to outperform the baseline are Logistic Regression and Support Vector Machine. Additionally, methods that improve performance, such as ensemble learning, will be explored if the previously mentioned models are not sufficient in beating the baseline.

Finally, the study will produce the best fitting model which can determine whether an examinee cheated or not based on their input device data. This model can be used in a more general way. Based on the final model, it can also be determined which specific input device data are indicative of cheating in an online exam.

The researchers will develop a simple system that uses this model where examinees can take their exams. The system will record input device data that are needed in the model and these will be used as features for prediction.

III. Methodology

Scope of the Study

The study focuses on developing a model that can predict whether an examinee cheated on an online examination. It is limited to the tracking of input devices data in identifying such activities of the examinees.

In this study, input devices refer to mouse, keyboard, and trackpad if they are using a laptop. However, for the mouse, it is limited to the traditional mouse having only the basic parts: the scroll wheel, left-click and right-click buttons. This means that other mice with programmable buttons will not be covered.

A web-based application will also be developed where examinees can take tests while being monitored for possible cheating. Since the study is primarily concerned with data from input devices such as mouse, keyboard and trackpad, it will not include mobile devices, which rely on capacitive touch.

Research Design

Subject and Locale

The proponents will use survey sampling in gathering data for the study by holding a mock online examination. There will be 112 participants chosen from students of the College of Computer Studies and Systems (CCSS) at University of the East – Manila, aged 18 and above. Voorhis and Morgan (2007) states that a general rule of thumb for choosing the sample size for regression analysis is by using Green's formula $N > 104 + m$, where m is the number of independent variables. For this study, 7 independent variables will be used; as such, 112 participants were determined as the sample size.

Participants will also come from the faculty of UE-CCSS Manila. 10 respondents will be selected using purposive sampling. These respondents will be professors who are conducting examinations online.

Data Gathering

Initially, 11 student and 10 professor participants will be selected to accomplish an interview questionnaire via Google Forms. The specific questions for both groups differ from each other, however, they all relate to the experiences of the professors and students when conducting and taking online examinations, respectively. There are 8 questions each for both groups, all of which are open-ended.

The results will then be analyzed and will help determine whether the state of cheating in online examinations in a specific setting is consistent with the studies discussed in Chapter 2.

After the interview questions, the 112 student participants from UE-CCSS Manila will randomly be split into two groups with equal members. Both groups will be asked to answer a question set consisting of 15 multiple-choice and fill-in-the-blank items. They will be answering the question set in a simple web application developed by the researchers.

The questions in the question set fall under the category of Information Technology; they are difficult questions but otherwise easy to find on the Internet. Both groups will be instructed to act as if the exam is a real test, and answer the questions as best they can. The only difference is that for the first group, they will cheat using any methods to obtain the answer; however, they can't use any devices other than the one they're using. On the other hand, for the second group they will be answering honestly and if they don't know the answer, they must choose the answer which they think is the best one. The items will be randomly arranged for each participant to further mimic the real-world testing environment.

A JavaScript plugin built into the web application will be used to gather input device data from each participant. This will include seven Document Object Model (DOM) events: blur, focus, copy, paste, left-click, right-click, and double-click. The focus and blur events will be used to tell the number of times a student changed browser tabs or applications during a test using any input devices as well as the duration in seconds that the student is inactive. Similarly, the number of times that copy and paste are used will be counted, as well as the frequency of the left, right, and double clicks occurring.

Modeling

Salehi and Gholampour (2021) found that exam difficulty is one of the reasons for cheating in exams, which is related to the study predicting question difficulty in web surveys based on mouse-tracking measures using logistic regression, tree-based models, and SVM (Fernández-Fontelo et al., 2021). Using the data gathered, the proponents will apply the supervised machine learning algorithms logistic regression and support vector machines (SVM) as well as ensembling methods to obtain a binary classifier that will be able to predict whether an examinee cheated or not. The mentioned input device data will serve as the predictors or features. All of the following data analysis will be done using Python libraries and frameworks such as scikit-learn.

To determine how good the model is once it has been fitted to the data, the deviance, or -2 log-likelihood (-2LL) statistic, will be used. In simple terms, deviance is a measure of how much unexplained variation there is in the obtained logistic regression model; the higher the value of deviance, the less accurate the model is, with large values indicating poorly-fitting models.

Aside from determining the overall fit of the model to the data, it is also important to determine the individual contribution of the independent variables. This can be done by using the Wald statistic which indicates whether or not an explanatory variable is significant to the model. In conjunction with the Wald statistic, the deviance statistic can also be used to see if adding or removing certain independent variables will improve the predictive power of the model. By comparing the deviance with each set of independent variables added to the model, the model can be improved—if the -2LL decreases to a statistically significant degree for each set of independent variables added, it means that the predictive power of the model is increasing. Thus, using these two methods, if an explanatory variable is found to be insignificant to the model, it will be removed from the final model.

To evaluate the performance of the model on new data, k-fold cross-validation will be used. This is a method that works by splitting the dataset into k-parts. To illustrate, if the dataset is split into k = 10 parts, the model will be trained 10 times and each time, 9 subsets will be included in the training group while the remaining subset will be the testing group. The testing group rotates each time the model has been trained. This will result in a more reliable estimate of the performance of the model on real data (Mayer, 2021).

The accuracy metrics that will result from k-fold cross-validation will also be used to compare the performance to the baseline that used mouse tracking. Thus, it is possible to determine whether the algorithm outperforms existing algorithms in terms of accuracy.

To obtain the sensitivity and specificity of the model, a confusion matrix will be used which is a performance measurement for machine learning classification problems (Narkhede, 2021). It is a table which describes the predicted and actual values, namely: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). With this, the sensitivity and specificity can be computed using the following formulas:

$$\textit{Sensitivity} = \frac{TP}{TP+FN}$$

Specificity = $\frac{TN}{TN+FP}$

Once the best fitting model has been determined, it can then be used for a more general use case. In light of this, the researchers will develop a simple web application where examiners can create and conduct their tests. The web application will have a built-in JavaScript script—which does not require examiners nor examinees to install additional software—that will record the required input device data. The recorded data will then serve as inputs to the model. Hence, it becomes possible to determine whether an examinee cheated or not.

Evaluating the Prototype

To evaluate the usefulness and effectiveness of the system when used by test administrators — specifically, professors of the College of Computing Systems and Studies (CCSS) at University of the East – Manila — a survey consisting of 12 questions will be given to 10 respondents who are all CCSS professors to determine how satisfied they are with the system. To obtain respondents, purposive sampling will be used since the target population is small.

The survey is based on the 12-item survey instrument developed by Doll and Torkzadeh (1988) that measures five components of end-user satisfaction: content, accuracy, ease of use, format, and timeliness. The survey uses a five-point Likert scale system, with “1 = Almost Never” and “5 = Almost Always”. The results will then be summarized using descriptive statistics.

Paper Prototype

The web application will have a landing page that will lead to the login or sign-up page. After signing up for a new account and successfully logging in, users will be directed to either of the two interfaces - depending on if their account is for a teacher or a student. Users with a teacher's account will be able to create exams and view their reports. On the other hand, users with a student's account will be able to take exams and view their activities.

The following images show the paper prototype of the system:

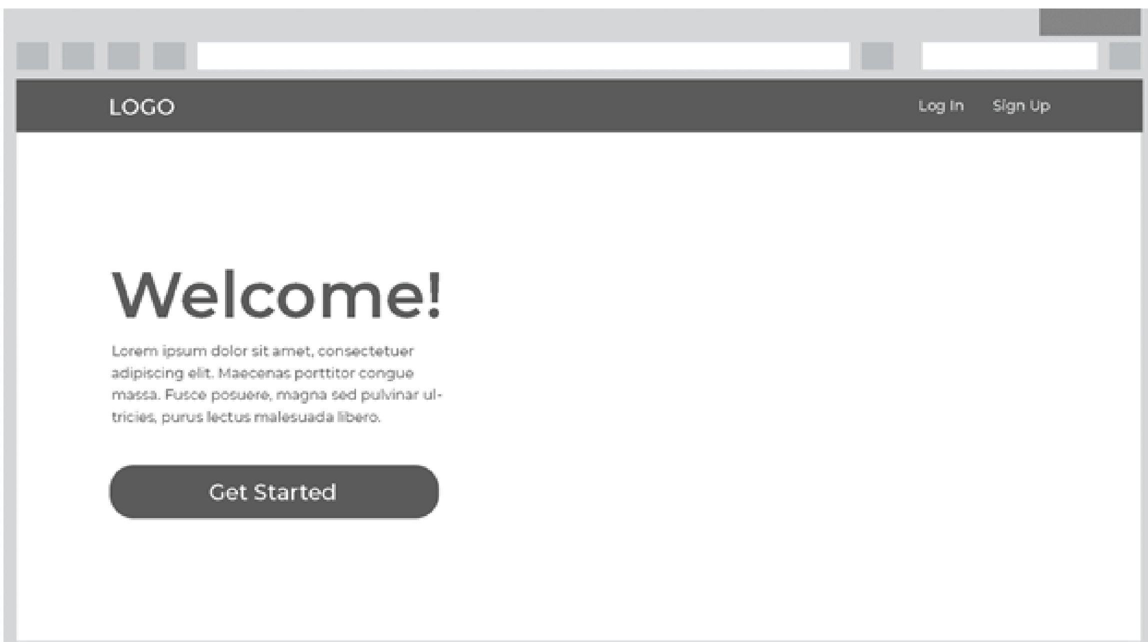


Image 1: Landing Page

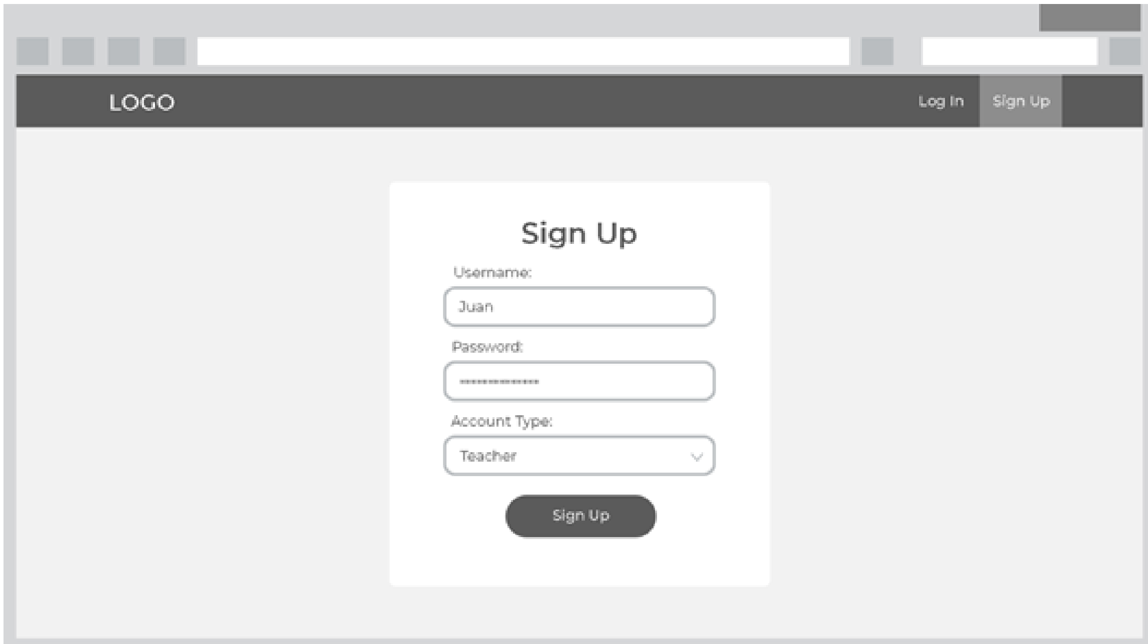


Image 2: Sign Up Page

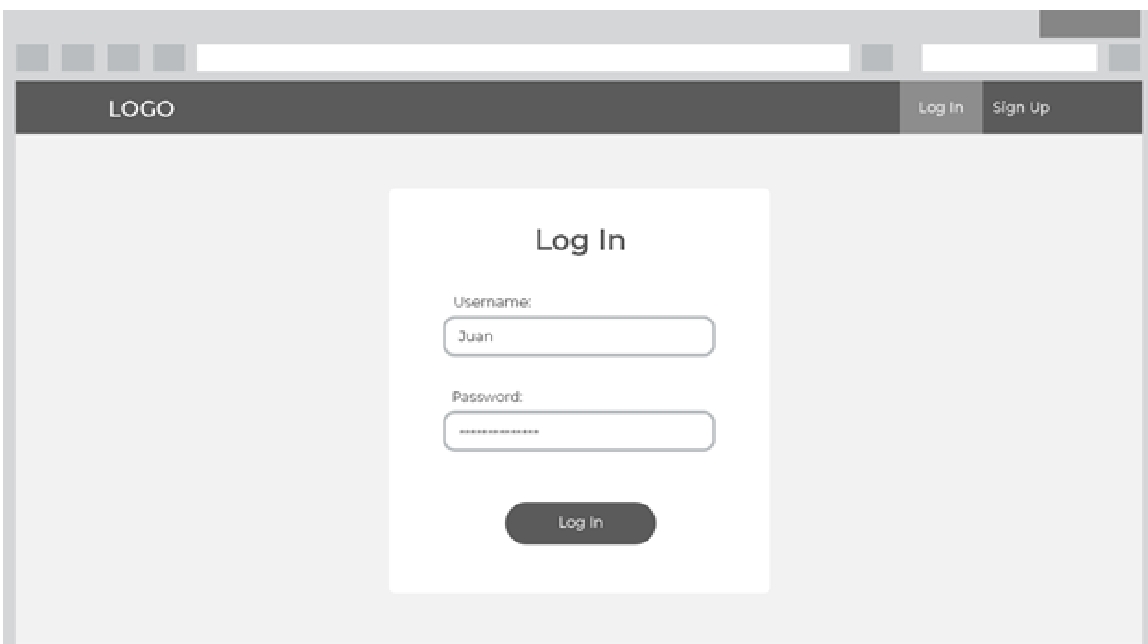


Image 3: Login Page

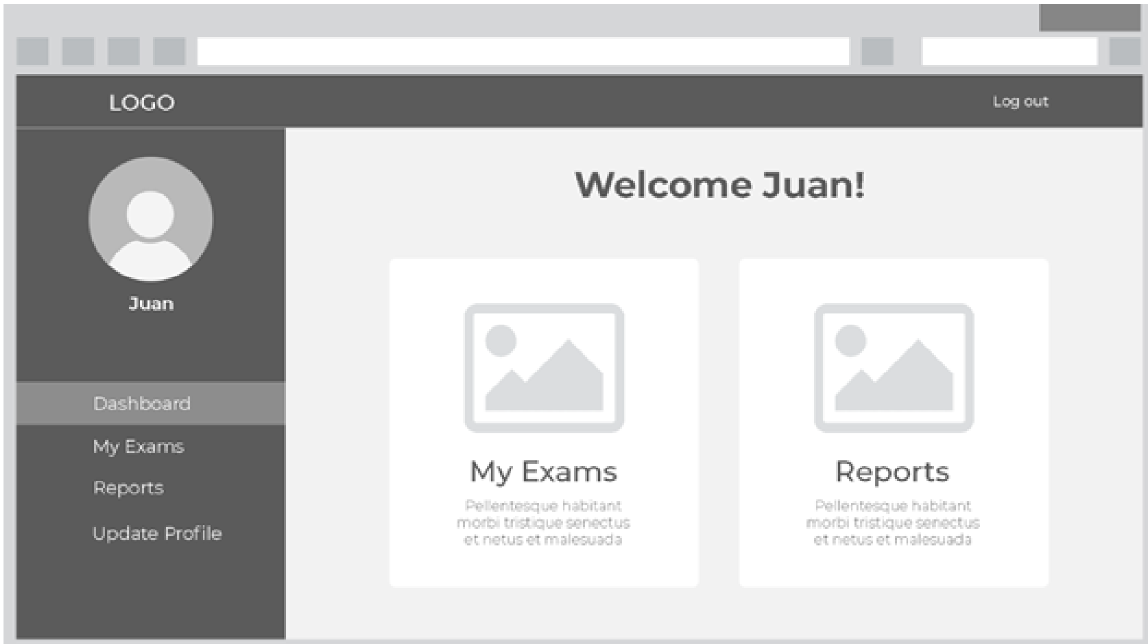


Image 4: Teacher's Dashboard Page

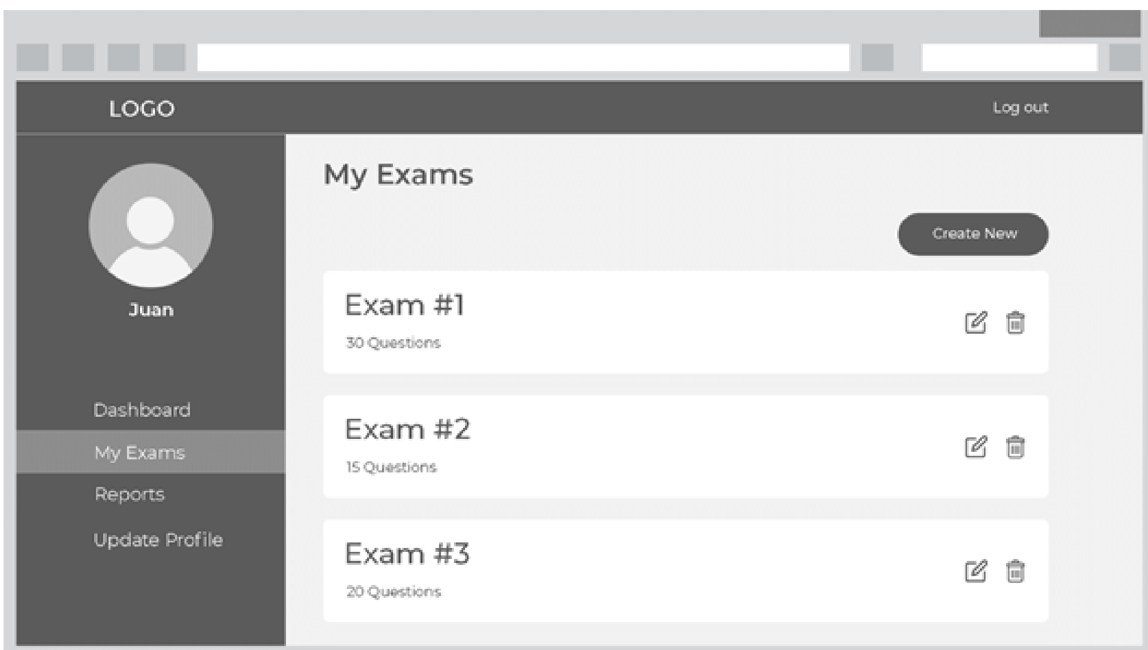


Image 5: My Exams Page

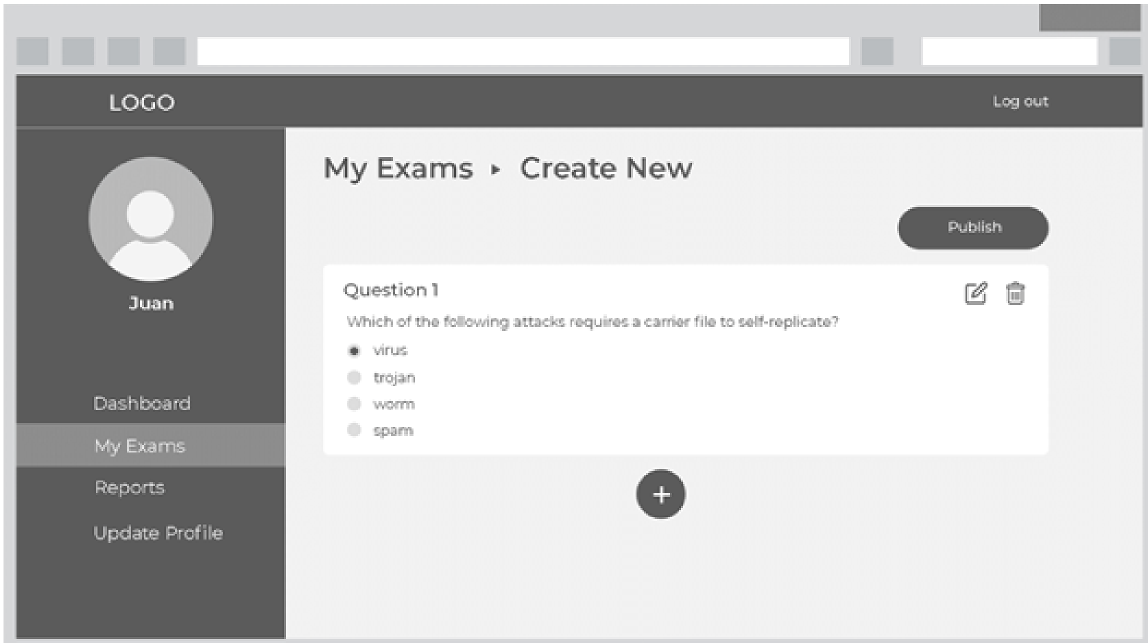


Image 6: Create New Exam Page

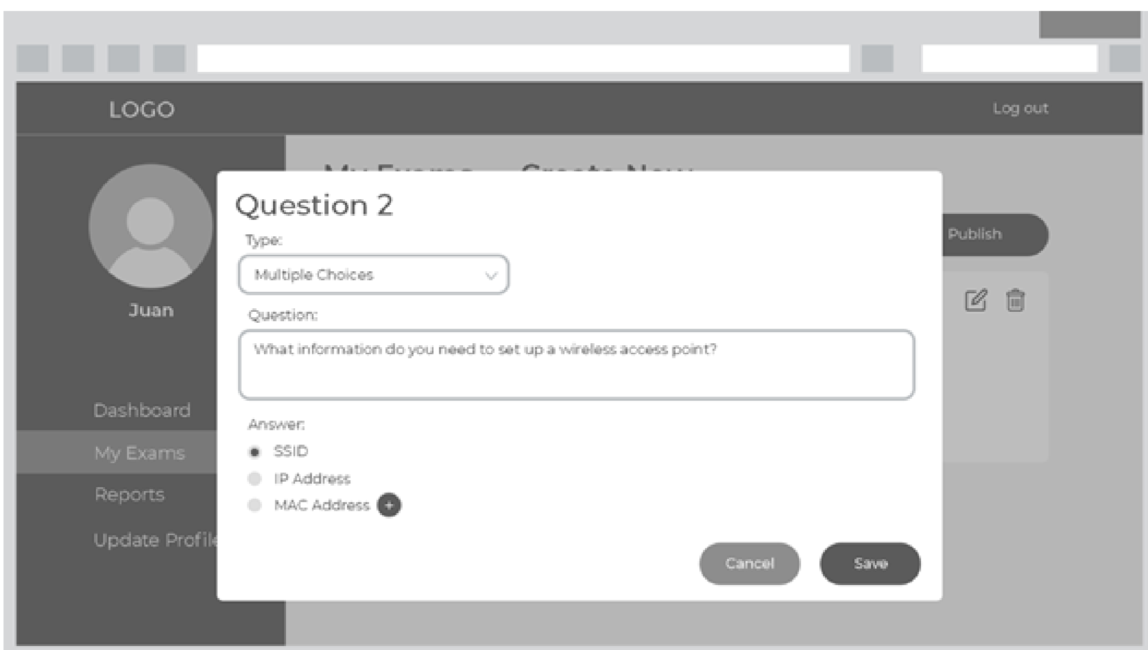


Image 7: Add New Question Modal Window for Multiple Choice Questions

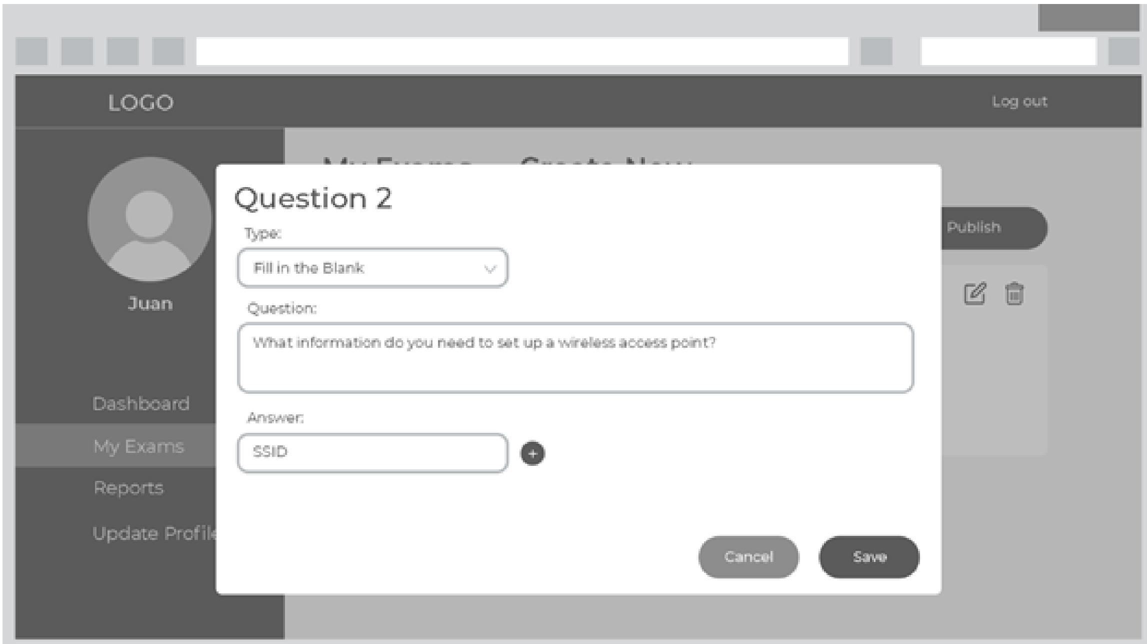


Image 8: Add New Question Modal Window for Fill-in-the-Blank Questions

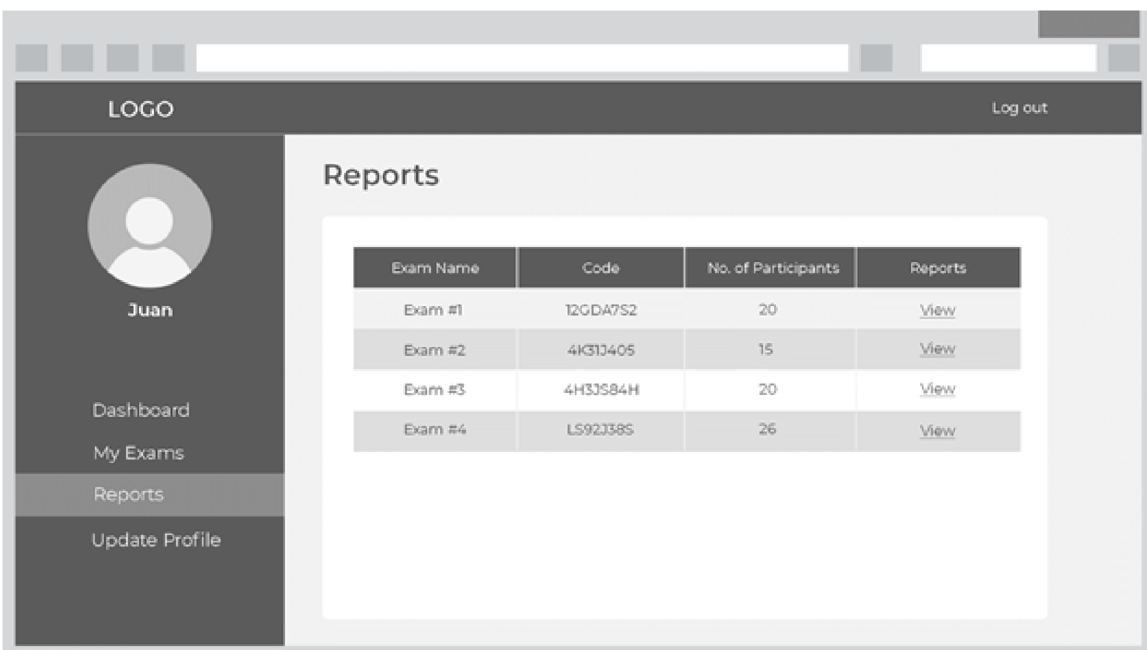


Image 9: Reports Page

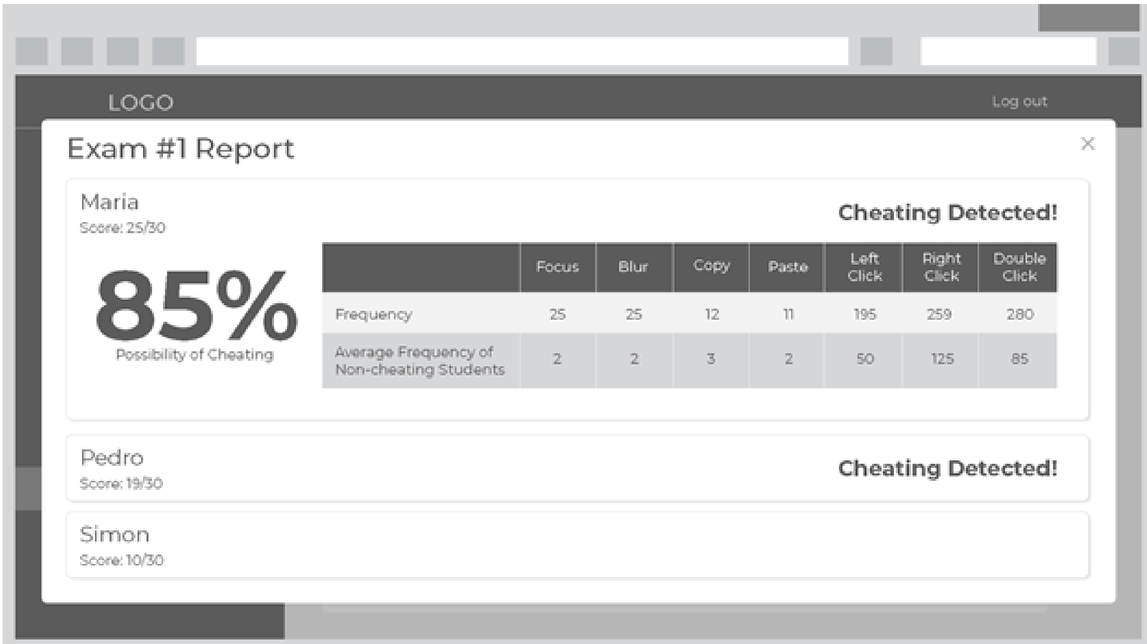


Image 10: View Reports Modal Window

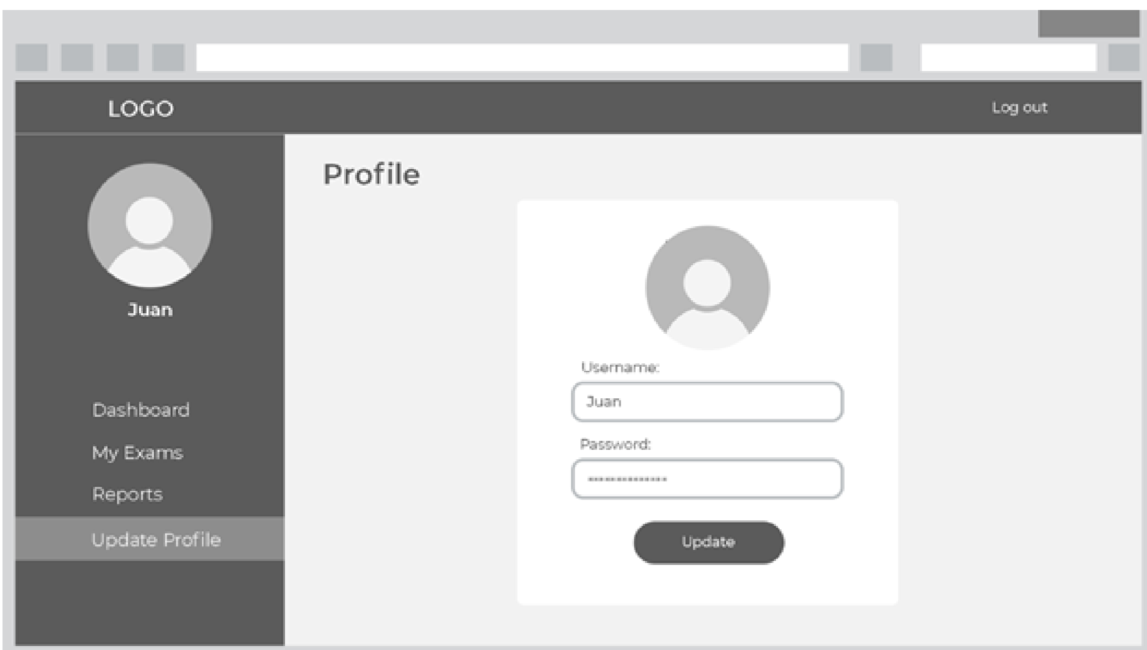


Image 11: Teacher's Update Profile Page

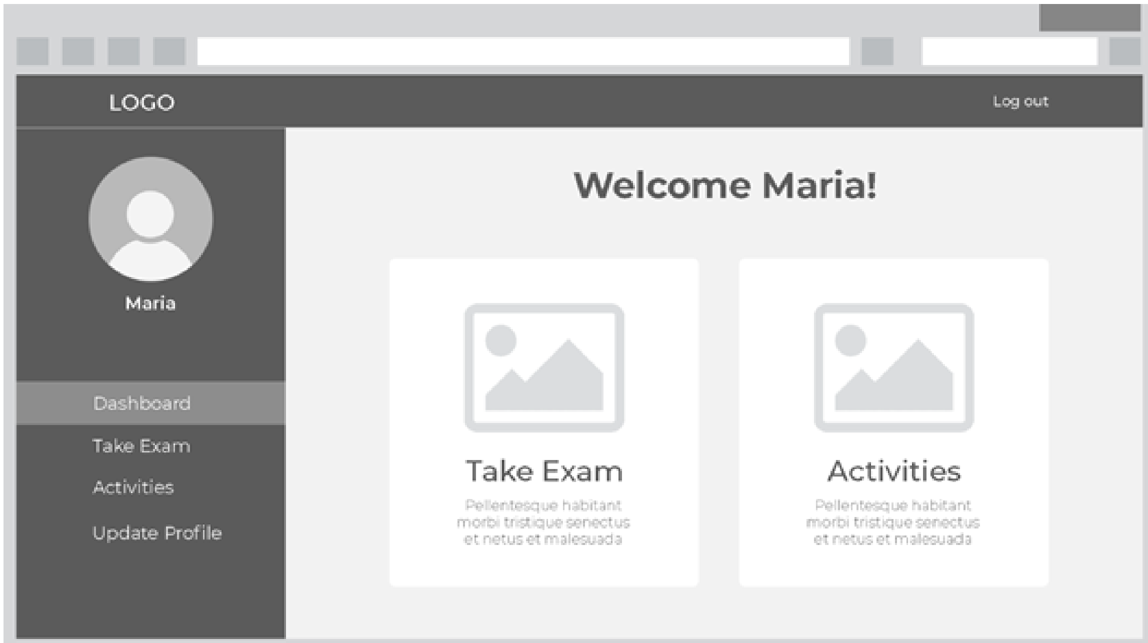


Image 12: Student's Dashboard Page

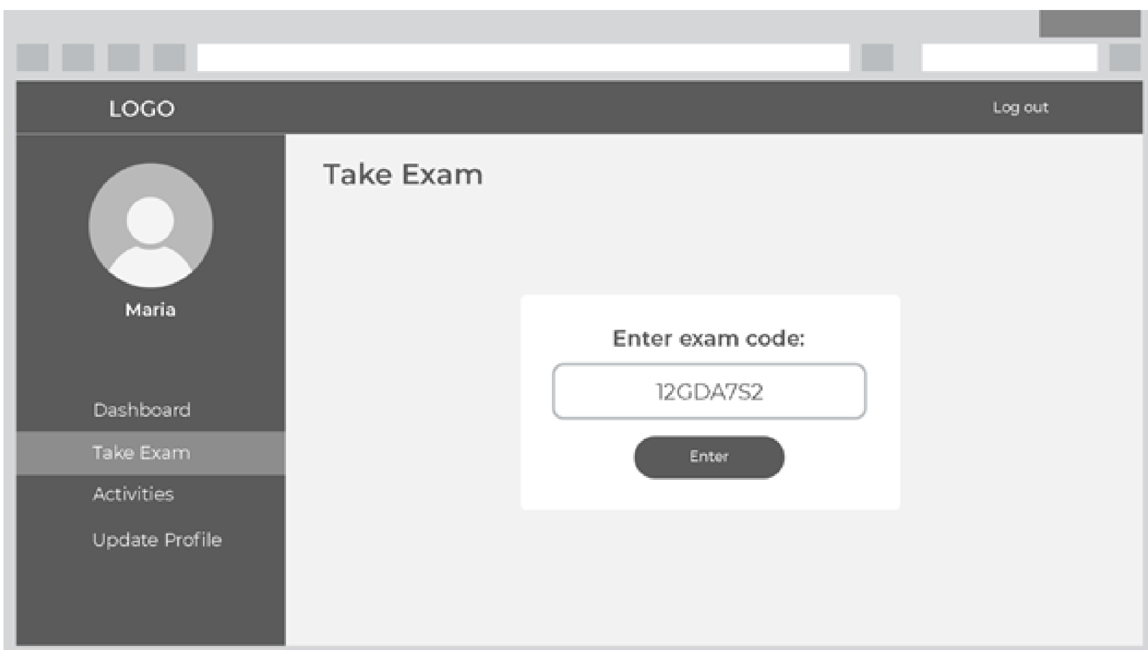


Image 13: Take Exam Page

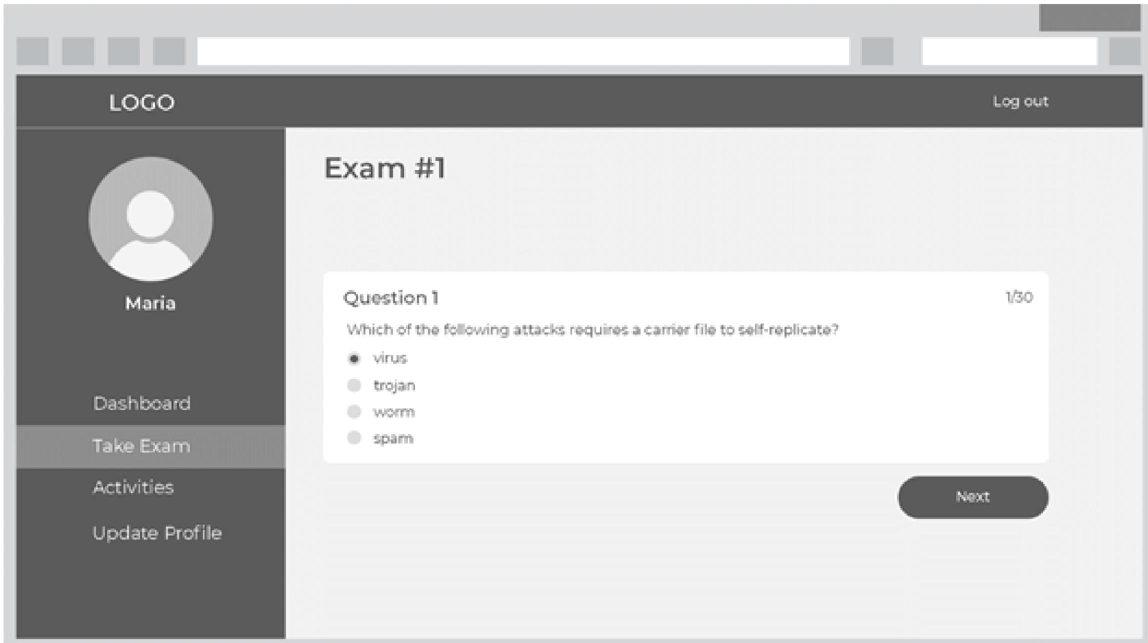


Image 14: Exam Page

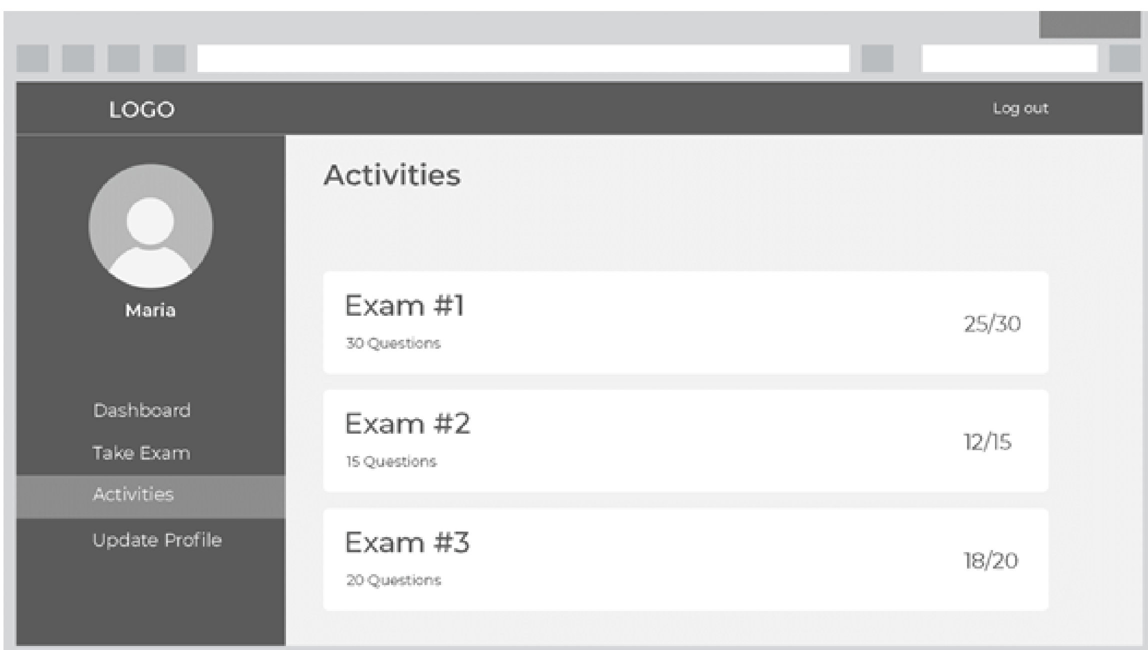


Image 15: Activities Page

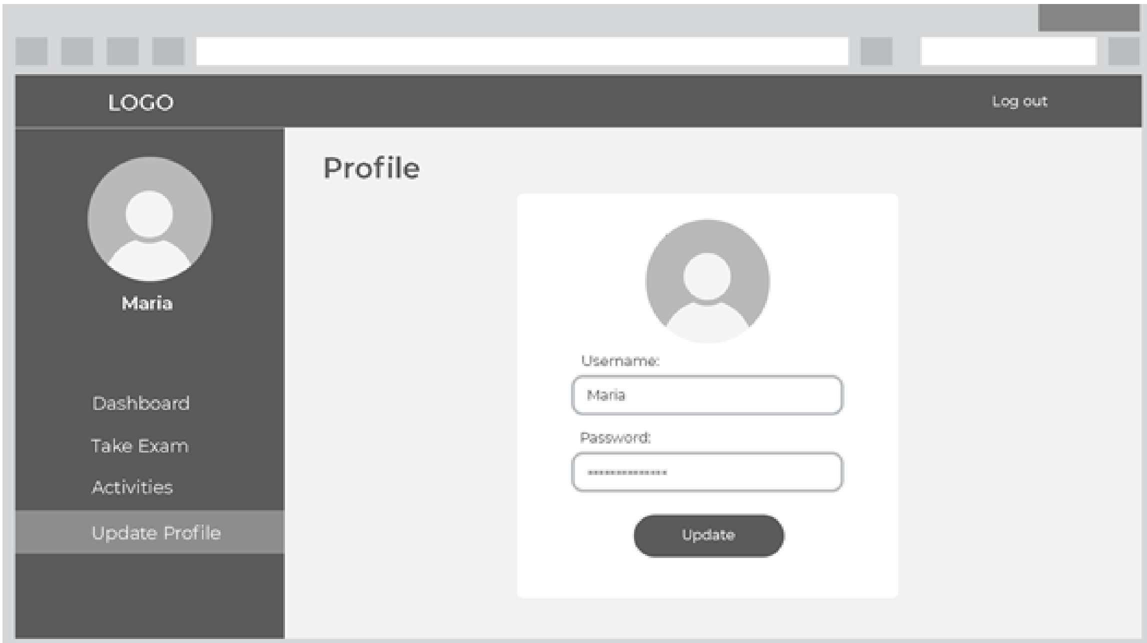


Image 16: Student’s Update Profile Page

References

- Arnold, I. J. M. (2016). *Cheating at online formative tests: Does it pay off?* ScienceDirect. Retrieved September 28, 2021, from https://www.sciencedirect.com/science/article/abs/pii/S1096751616300021
- Anonymous. (n.d.). Ensemble methods. Retrieved from https://scikit-learn.org/stable/modules/ensemble.html
- Best, L. M., & Shelley, D. J. (2018). *Academic Dishonesty: Does Social Media Allow for Increased and More Sophisticated Levels of Student Cheating?* IGI Global. Retrieved September 28, 2021, from https://www.igi-global.com/article/academic-dishonesty/205617
- Bilen, E., & Matros, A. (2021). *Online cheating amid COVID-19*. ScienceDirect. Retrieved September 28, 2021, from https://www.sciencedirect.com/science/article/abs/pii/S0167268120304510
- Computer Hope. (2021). *What is an Input Device?* Retrieved November 23, 2021, from https://www.computerhope.com/jargon/i/inputdev.htm
- Dendir, S., & Maxwell, R. S. (2020). *Cheating in online courses: Evidence from online proctoring*. ScienceDirect. Retrieved September 28, 2021, from https://www.sciencedirect.com/science/article/pii/S2451958820300336

- Diedenhofen, B., & Musch, J. (2016). *PageFocus: Using paradata to detect and prevent. . .* Springer Link. Retrieved September 28, 2021, from https://link.springer.com/article/10.3758/s13428-016-0800-7?error=cookies_not_supported&code=2cc8aa92-6e03-4d1f-b6db-b5574af2ba80
- Fernández-Fontelo, A., Kieslich, P. J., Henninger, F., Kreuter, F., & Greven, S. (2021). Predicting question difficulty in web surveys: A machine learning approach based on mouse movement features. Social Science Computer Review, 08944393211032950.
- Ganaie, M. A., & Hu, M. (2021). Ensemble deep learning: A review. arXiv preprint arXiv:2104.02395.
- Horwitz, R., Brockhaus, S., Henninger, F., Kieslich, P.J., Schierholz, M., Keusch, F. and Kreuter, F. (2020). Learning from Mouse Movements: Improving Questionnaires and Respondents' User Experience Through Passive Data Collection. In Advances in Questionnaire Design, Development, Evaluation and Testing (eds P. Beatty, D. Collins, L. Kaye, J.L. Padilla, G. Willis and A. Wilmot). https://doi.org/10.1002/9781119263685.ch16
- Javed, Ali & Aslam, Zeeshan. (2013). An Intelligent Alarm Based Visual Eye Tracking Algorithm for Cheating Free Examination System. International Journal of Intelligent Systems and Applications. 5. 86-92. 10.5815/ijisa.2013.10.11.

- Jensen, C., & Thomsen, J. P. F. (2013). *Self-reported cheating in web surveys on political knowledge*. Springer Link. Retrieved September 28, 2021, from https://link.springer.com/article/10.1007/s11135-013-9960-z?error=cookies_not_supported&code=753d9830-454a-46c5-9758-3ae5828b292b
- Karim, M. N., Kaminsky, S. E., & Behrend, T. S. (2014). *Cheating, Reactions, and Performance in Remotely Proctored Testing: An Exploratory Experimental Study*. Springer Link. Retrieved September 28, 2021, from https://link.springer.com/article/10.1007/s10869-014-9343-z?error=cookies_not_supported&code=317b7dc3-39ad-42be-b23b-47e8df04bcaa
- Kieslich, P. J., & Hilbig, B. E. (2014). Cognitive conflict in social dilemmas: An analysis of response dynamics. Judgment and Decision making, 9(6), 510.
- Kleinbaum, David G.; Klein, Mitchel (2010). [Statistics for Biology and Health] Logistic Regression || Introduction to Logistic Regression. , 10.1007/978-1-4419-1742-3(Chapter 1), 1–39. doi:10.1007/978-1-4419-1742-3_1
- Mayer, A. J. (2021). *Using K-Fold Cross-Validation to Evaluate the Performance of Logistic Regression*. Medium. Retrieved October 29, 2021, from

https://python.plainenglish.io/using-k-fold-cross-validation-to-evaluate-the-performance-of-logistic-regression-4439215f24c4

- middleearthnj. (2014). *Cheating in School: Facts, Consequences & Prevention*. Middle Earth. Retrieved October 2, 2021, from https://middleearthnj.org/2014/02/24/cheating-in-school/
- Northern Illinois University. (n.d.). *Academic Dishonesty Definition and Types - NIU - Academic Integrity Tutorial for Faculty*. Retrieved November 24, 2021, from https://www.niu.edu/academic-integrity/faculty/types/index.shtml
- Purnama, F., Sukhbaatar, O., Choimaa, L., & Usagawa, T. (2020, March 6). *Implementation of real-time online mouse tracking on overseas quiz session*. Springer Link. Retrieved September 28, 2021, from https://link.springer.com/article/10.1007/s10639-020-10141-3?error=cookies_not_supported&code=702cfb58-62e9-4eff-93f9-495d4da44754
- Sagi, O, Rokach, L. Ensemble learning: A survey. WIREs Data Mining Knowl Discov. (2018); 8:e1249. https://doi.org/10.1002/widm.1249
- Salehi, M., & Gholampour, S. (2021). Cheating on exams: Investigating Reasons, Attitudes, and the Role of Demographic Variables. SAGE Open, 11(2), 21582440211004156.

- Tendeiro, J. N., Meijer, R. R., Schakel, L., & Maij-de Meij, A. M. (2013). Using Cumulative Sum Statistics to Detect Inconsistencies in Unproctored Internet Testing. Educational and Psychological Measurement, 73(1), 143-161. https://doi.org/10.1177/0013164412444787
- Voorhis, Carmen & Morgan, Betsy. (2007). Understanding Power and Rules of Thumb for Determining Sample Size. Tutorials in Quantitative Methods for Psychology. 3. 10.20982/tqmp.03.2.p043.
- *Webcam Proctoring for Online Exams | Knowledge Base*. (n.d.). Drexel LeBow. Retrieved September 28, 2021, from https://www.lebow.drexel.edu/resources/technology/knowledge-base/webcam-proctoring-online-exams
- William J. Doll and Gholamreza Torkzadeh. 1988. The measurement of end-user computing satisfaction. *MIS Q.* 12, 2 (June 1988), 259–274. DOI:https://doi.org/10.2307/248851

Appendice

CCSS CERC Form No. 2, s2021

UNIVERSITY OF THE EAST
College of Computer Studies and Systems

RESULT OF THE CERC REVIEW

CCSS is collecting information for the purpose of evaluating your application for research proposal ethical review. By accomplishing this form, you are certifying that all information provided are true and correct and likewise authorizing this office to process your information. The accomplished form will be kept in a secured place and will be disposed of after one academic year.

2021-1-PTCS-84
CCSS CERC Code

This certifies that the CCSS- College Ethics Review Committee (CERC) after having reviewed the following proposal approves and recommends the study for implementation.

Type of Submission					
[✓]	Initial				
[ ]	Resubmission				
Research Proposal Title	**Tracking Input Devices to Detect Cheating Using Machine Learning Techniques**				
Principal Researcher	**Virgilio**		**A.**	**Lopez Jr.**	
	Given Name		Middle Initial	Surname	
Research Adviser	**Melie Jim**		**F.**	**Sarmiento**	
	Given Name		Middle Initial	Surname	
Type of Review					
[ ]	Exempt				
[ ]	Expedited with Minor or Major Modification				
[ ]	Full Minor or Major Modification Issues and concerns				
Initial action taken after the initial review:					
[ ] Approval	[ ] Major Modification	[ ] Minor Modification	[ ] Pending		
	[ ] Resubmit for another review	[ ] Resubmit for another review	[ ] Resubmit for another review	(Month Day, Year)	
Final action taken					
[ ] Approval granted after expedited review		[ ] Approval granted after full review			(Month Day, Year)

Recommendations:

CCSS CERC Form No. 1, s2021

UNIVERSITY OF THE EAST
College of Computer Studies and Systems
APPLICATION FORM FOR ETHICS REVIEW AND RESUBMISSION

CCSS is collecting information for the purpose of evaluating your application for research proposal ethical review. By accomplishing this form, you are certifying that all information provided are true and correct and likewise authorizing this office to process your information. The accomplished form will be kept in a secured place and will be disposed of after one academic year.

October 29, 2021

2021-1-PTCS-84
CCSS CERC Code

GENERAL INFORMATION			
Type of Submission:		***Research Category:***	
[✓]	Initial Review	[✓]	Research involving human (participants/respondents and/or subjects)
[ ]	Resubmission (This is a revised version incorporating comments and suggestions of the review committee)	[ ]	Research involving non-human living vertebrates
	NOTE: Version and date of revision must be inserted as document footer for all resubmissions.	[ ]	Research involving experiments on the environment
Study Site Address (Please indicate.):		[ ]	Others (Please indicate):
Type of Study Site:			
[✓]	within college		
[ ]	inter-college study		
[ ]	outside of the University campus		

STUDY DESCRIPTION		
Research Proposal Title	**Tracking Input Devices to Detect Cheating Using Machine Learning Techniques**	
Brief Description Describe the background, methodology, rationale/significance, hypothesis/research questions, dissemination plan, and the expected outcomes. (use additional sheet if necessary)	**Background of the Study** Online classes have become the temporary norm due to the state the world is in brought by the pandemic. With this, lectures and examinations are being conducted through online means. This may be advantageous with the students, and teachers being in the comfort of their own homes, however the disadvantage brought by this setup is also quite serious. The lack of surveillance in online exams is one such disadvantage that may prove to be quite a problem in the long run. *(continued on page 5)*	
Study Duration *(from proposal to deployment)*	**From** (start date)	**First Semester S.Y. 2021-2022**
	To (end date)	**December 2022**

Respondents / Research Participants: Students and professors from UE-CCSS Manila		
Use of special populations or vulnerable groups:		
[ ]	Children (under 18 years old)	
[ ]	Elderly (above 60 years old)	
[ ]	Indigenous people	
[ ]	People on welfare/social assistance	
[ ]	Displaced persons	
[ ]	Others (Please indicate):	
[✓]	Not applicable	

INFORMATION ABOUT THE RESEARCHER/S			
Principal Researcher	**Virgilio**	**A.**	**Lopez Jr.**
	Given Name	Middle Initial	Surname
Address	**82 Banaba South, Batangas City, Batangas**		
Landline No.	**N/A**		
Mobile No.	**0917-423-1378**		
UE Email Address	**lopez.virgiliojr@ue.edu.ph**		
Research course enrolled	***CCS 3102 - CSAE***	***Methods of Research for CS***	
	course code - section	course name	
Term	***1st***	***2021-2022***	
	Semester	Academic Year	

Submitted by:

Altoveros, Louis Angelo D.R.	**Lopez, Virgilio Jr. A.**
Manangat, Renz Darryl M.	**Cabo, Ralf Martin**
Tiosan, John Christian B.	

==

ADVISER'S ENDORSEMENT for ETHICS REVIEW
This is to certify that the proposed research topic has been reviewed and is endorsed for ethics review of the College Ethics Review Committee.

Endorsed by:

Melie Jim Sarmiento

Brief Description (additional sheet)	

The research is focused around this dilemma as the researchers think of a way to combat this using cheating detection on input device data.

Methodology

Initially, 11 students and 10 professors will be asked to answer an interview questionnaire consisting of 8 open-ended questions. These questions relate to the experiences of the professors and students in conducting and taking online examinations, respectively.

The researchers will then develop a simple web application where the participants will answer a questionnaire consisting of 15 multiple-choice and fill-in-the-blank questions. The web app has a built-in JavaScript plugin which is able to record the participants' input device data during the examination. Specifically, these data consist of seven Document Object Model (DOM) events, which are blur, focus, copy, paste, left click, right click, and double-click.

The researchers will also be creating a prototype for examiners to conduct their exams. To evaluate the effectiveness and usefulness of the prototype, it will be pilot-tested by professors from the College of Computer Studies and Systems in UE-Manila. At the end of the testing, they will be asked to accomplish a survey via Google Forms. The results will then be analyzed by the researchers using descriptive statistics.

Rationale and Significance of the Study

With how convenient it is now to cheat on online exams, an algorithm tracking input device data can, at the very least, act as a deterrent for students attempting to cheat. Also, previous study suggests that cheating has harmful effects in the long term, as a single attempt of cheating can ultimately lead to dependency on the act, furthermore in the field of employment, it sets the bar too high for the cheater due to their inflated marks. This algorithm can be useful, not only in school online examinations, but also in online certifications which require passing an online test to measure their capability.

Research Questions

1. How well can we detect cheating in online examinations based only on the examinees' input device data in terms of:
    a. Specificity
    b. Sensitivity
2. What specific input device data can be an indicator of cheating?
3. How useful will a system that uses input device data be for examiners?

Hypothesis

H_0: In terms of accuracy, the proposed algorithm does not outperform existing algorithms in detecting cheating in online examinations.

H_1: In terms of accuracy, the proposed algorithm outperforms existing algorithms in detecting cheating in online examinations.

Dissemination Plan

After the development of the software, proponents will invite 112 students aged 18 and above through Facebook posts on different active pages of College of Computer Studies and Systems (CCSS) of University of the East – Manila. They will be split into two groups, and each participant will take an online mock examination consisting of 15 multiple-choice and fill-in-the-blank questions. The first group will be instructed to use any methods to find the answer if they don't know the answer to a question. On the other hand, the second group will be instructed to just choose the best answer if they don't know the answer to a question.

	Expected Outcome The proponents expect to obtain a multiple logistic regression equation, with the specific mouse behaviors as its independent variables. This equation will determine the probability of cheating of a test taker based on their mouse behavior and tell whether the person was cheating or not on a particular question. Additionally, the equation will also determine which independent variable influences the probability the most, thus the proponents will be able to answer which specific input device data can be an indicator of cheating. By using k-fold cross validation, the performance of the algorithm in terms of accuracy, specificity, and sensitivity can be obtained. This will then be compared with the reported accuracy of other algorithms, thus it would be possible to determine if this algorithm is more accurate than other algorithms.

Prof Ma. Teresa F. Borebor
Dean
College of Computer Studies and Systems
University of the East
2219 C. M., Recto Ave., Manila

Dear Dean Borebor:

We are 3rd-year students taking up Bachelor of Science in Computer Science from the College of Computer Studies and Systems (CCSS). We would like to seek your endorsement as we perform data gathering for our research titled **"Tracking Input Devices to Detect Cheating Using Machine Learning Techniques"**

With your recommendation, we shall request the student and professor respondents to take part in an online interview questionnaire facilitated through Google Forms at the beginning of the study. During the study, we shall also request student respondents to participate in an online mock examination deployed via the Internet. Finally, the initial professor respondents will be asked to answer an evaluation questionnaire at the end of the study. All collected and processed information will be handled with the utmost confidentiality and will be used for academic and research purposes only. Copies of the questionnaires for the mock exam and survey instruments are attached for your perusal. Thank you for your support.

Sincerely,

Altoveros, Louis Angelo D.R.	Lopez, Virgilio Jr. A.
Manangat, Renz Darryl M.	Tabo, Ralf Martin R.
Tiosan, John Christian B.	

Noted by:

Melie Jim Sarmiento
Thesis Adviser

Signed:

Ma. Teresa F. Borebor
Dean, CCSS

Informed Consent (for students and professors)

We are 3rd year students taking up Bachelor of Science in Computer Science from the College of Computer Studies and Systems (CCSS). We would like to seek participation in our research titled **"Tracking Input Devices to Detect Cheating Using Machine Learning Techniques"** by answering the attached survey.

Please take the time to read the details of our study.

1. *Purpose of the study:* Our research aims to observe the behavior of mouse movements made by students who are taking an exam through dishonest means.
2. *Procedure:* As a respondent, you will be asked to accomplish an electronic form that may include personal and sensitive information.
3. *Confidentiality:* All collected information will be used for academic and research purposes only and will be handled with utmost confidentiality. Presentation of results will be in the form aggregate information, and will not identify a specific individual. Collected information in soft and hardcopy formats will be disposed of after printing of the final manuscript.
4. *Right to Decline:* Your participation is voluntary and you have the right to decline or withdraw at any time.
5. *Questions and Concerns:* We are open to answer any questions or concerns. Please contact us.

Consent Form

"By signing this form, I agree to be a respondent to this study. I give my consent in the processing of my information collected from the provided questionnaire for academic and research purposes."

NAME:		
	Given Name	Surname
Signature		

Thank you for your participation.
If you have any questions or concerns you may contact us.

Researcher	Email Address	Mobile Number
Altoveros, Louis Angelo D.R.	altoveros.louisangelo@ue.edu.ph	09064138199
Lopez, Virgilio Jr. A.	lopez.virgiliojr@ue.edu.ph	0917-423-1378
Manangat, Renz Darryl M.	manangat.renzdarryl@ue.edu.ph	0998-910-4419
Tabo, Ralf Martin R.	tabo.ralfmartin@ue.edu.ph	09260427627
Tiosan, John Christian B.	tiosan.johnchristian@ue.edu.ph	09284298596

Initial Interview Questions (for students and professors)

Tracking Mouse Behavior to Detect Cheating Using Multiple Logistic Regression

We are 3rd year students taking up Bachelor of Science in Computer Science from the College of Computer Studies and Systems (CCSS). We would like to seek participation in our research titled "Tracking Mouse Behavior to Detect Cheating Using Multiple Logistic Regression" by answering the attached survey.

Please take the time to read the details of our study:

1. Purpose of the study: Our research aims to observe the behavior of mouse movements made by students who are taking an exam through dishonest means.
2. Procedure: As a respondent, you will be asked to accomplish an electronic form that may include personal and sensitive information.
3. Confidentiality: All collected information will be used for academic and research purposes only and will be handled with utmost confidentiality. Presentation of results will be in the form of aggregate information, and will not identify a specific individual. Collected information in soft and hardcopy formats will be disposed of after printing the final manuscript.
4. Right to Decline: Your participation is voluntary and you have the right to decline or withdraw at any time.
5. Questions and Concerns: We are open to answering any questions or concerns. Please contact us.

Should you have any questions, please feel free to email or text:

Altoveros, Louis Angelo D.R. (altoveros.louisangelo@ue.edu.ph - 09064138199)
Lopez, Virgilio Jr. A. (lopez.virgiliojr@ue.edu.ph - 09174231378)
Manangat, Renz Darryl M. (manangat.renzdarryl@ue.edu.ph - 09989104419)
Tabo, Ralf Martin R. (tabo.ralfmartin@ue.ed - 09260427627)
Tiosan, John Christian (tiosan.johnchristian@ue.edu.ph - 09284298596)

Thank you!

Display a menu

lopez.virgiliojr@ue.edu.ph (not shared) Switch account   Draft restored

* Required

Part 1: Demographic Profile

Instructions: Please fill in the blank or check the box of your answer.

Full Name *

Your answer

Email Address *

Your answer

Age *

Your answer

Designation *

- ○ Student
- ○ Professor

College/Department *

Display a menu

College/Department *

- CBA
- CAS
- CCSS
- CEDUC
- CENG
- CDENT

Next Clear form

Never submit passwords through Google Forms.

This form was created inside of University of the East. Report Abuse

Google Forms

Display a menu

Tracking Mouse Behavior to Detect Cheating Using Multiple Logistic Regression

lopez.virgiliojr@ue.edu.ph (not shared) Switch account

* Required

Interview Questions for Students

Part 2: Interview Questions

The following questions are open-ended and pertain to your prior experiences when taking online examinations. Please answer the questions accordingly.

How often do you study for an online examination? *

Your answer

Have you ever cheated on an online examination? If yes, did you ever get caught? If no, has the thought ever crossed your mind? *

Your answer

What methods did you use, or planned to use to cheat on an online exam? Kindly describe them briefly. *

Display a menu

What methods did you use, or planned to use to cheat on an online exam? Kindly describe them briefly. *

Your answer

What do you think are the reasons why you don't get caught cheating? *

Your answer

Did your cheating produce good results on the examination? *

Your answer

Do you think it is easier to cheat in online examinations than traditional exams (i.e. offline face-to-face exams)? *

Your answer

In the case of an online examination being recycled for a future exam, did you still cheat on the latter? *

Your answer

Display a menu

What methods are used by your professors to prevent students from cheating in online exams? Are they effective? *

Your answer

Back Submit Clear form

Never submit passwords through Google Forms.

This form was created inside of University of the East. Report Abuse

Google Forms

Display a menu

Tracking Mouse Behavior to Detect Cheating Using Multiple Logistic Regression

lopez.virgiliojr@ue.edu.ph (not shared) Switch account

* Required

Interview Questions for Professors

Part 2: Interview Questions

The following questions are open-ended and pertain to your prior experiences in conducting online classes and examinations. Please answer the questions accordingly.

Is cheating rampant in your exams online? *

Your answer

How do you know if a student cheated on your exam? *

Your answer

What methods do you use to determine if a student is/was cheating? *

Display a menu

What methods do you use to determine if a student is/was cheating? *

Your answer

Were there times when you were not certain whether a student was cheating or not? *

Your answer

In the case of observed cheating, did you call the student's attention or had simply let it pass? *

Your answer

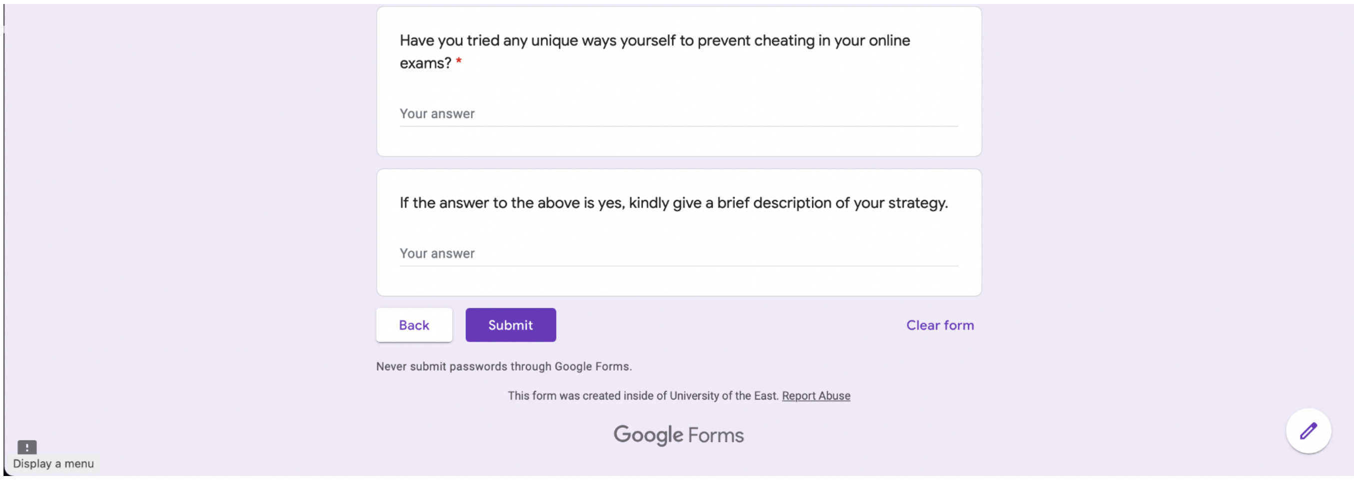
In the case of observed cheating, what behaviors have you observed from the student that gave away the hint that they are using dishonest means? *

Your answer

Have you tried any unique ways yourself to prevent cheating in your online exams? *

Your answer

Display a menu


Have you tried any unique ways yourself to prevent cheating in your online exams? *

Your answer

If the answer to the above is yes, kindly give a brief description of your strategy.

Your answer

Back Submit Clear form

Never submit passwords through Google Forms.

This form was created inside of University of the East. Report Abuse

Google Forms

Display a menu

Mock Exam Questionnaire (for students)

We are 3rd year students taking up Bachelor of Science in Computer Science from the College of Computer Studies and Systems (CCSS). We would like to seek participation in our research titled “Tracking Mouse Behavior to Detect Cheating Using Multiple Linear Regression” by answering the attached survey.

Please take the time to read the details of our study:

1. Purpose of the study: Our research aims to observe the behavior of mouse movements made by students who are taking an exam through dishonest means.
2. Procedure: As a respondent, you will be asked to accomplish an electronic form that may include personal and sensitive information.
3. Confidentiality: All collected information will be used for academic and research purposes only and will be handled with utmost confidentiality. Presentation of results will be in the form of aggregate information, and will not identify a specific individual. Collected information in soft and hardcopy formats will be disposed of after printing the final manuscript.
4. Right to Decline: Your participation is voluntary and you have the right to decline or withdraw at any time.
5. Questions and Concerns: We are open to answering any questions or concerns. Please contact us.

Should you have any questions, please feel free to email or text:

Altoveros, Louis Angelo D.R. (altoveros.louisangelo@ue.edu.ph - 09064138199)
Lopez, Virgilio Jr. A. (lopez.virgiliojr@ue.edu.ph - 09174231378)
Manangat, Renz Darryl M. (manangat.renzdarryl@ue.edu.ph - 09989104419)
Tabo, Ralf Martin R. (tabo.ralfmartin@ue.ed - 09260427627)
Tiosan, John Christian (tiosan.johnchristian@ue.edu.ph - 09284298596)

Thank you!

Instructions for first group:
The following questions are multiple-choice and fill-in-the-blanks type, which are typical in online examinations. We request you to answer the questionnaire as if it is a real graded examination. If you are uncertain with the question or you do not know the answer, you are free to use any method to obtain the correct answer, including but not limited to using Google search, asking others, etc. However, you may only use the device that you are currently using in answering the questionnaire.

Instructions for second group:
The following questions are multiple-choice and fill-in-the-blanks type, which are typical in online examinations. We request you to answer the questionnaire as if it is a real graded examination. If you are uncertain with the question or you do not know the answer, please choose the best answer you think is correct.

1. The distinguishing characteristic of a local area network (LAN) is that it ________.
    a. connects computers at multiple locations
    b. connects computers at a single location
    c. is a network of networks
    d. does not require a wired connection

2. Which of the following wireless protocols is designed for transmitting data over short distances?
    a. optical fiber
    b. coaxial cable
    c. twisted-pair cable
    d. bluetooth

3. What technology can collect information to make decisions, reach conclusions, and combine information in new ways?
    a. virtual reality (VR)
    b. embedded computers
    c. artificial intelligence (AI)
    d. robotics

4. On the following statements, which is one of the functions of encapsulation?
    a. ensures that data pieces can be directed to the correct receiving end device
    b. tracks delay between end devices
    c. enables consistent network paths for communication
    d. allows modification of the original data before transmission

5. What types of activities are ideal for a robot to perform?
    a. Creative design work
    b. Critical thinking
    c. Repetitive tasks
    d. Group interaction

6. How do delivery companies track packages?
    a. They use code-scanning technology.
    b. They use GPS.
    c. They use a learning management system.
    d. They use robotic arms to load packages.

7. During the encapsulation process, what occurs at the data link layer?
    a. No address is added.
    b. The logical address is added.
    c. The physical address is added
    d. The process port number is added.

8. What is the primary purpose of Layer 4 port assignment?
    a. to identify devices on the local media
    b. to identify the hops between source and destination
    c. to identify to the intermediary devices the best path through the network

d. to identify the processes or services that are communicating within the end devices

9. How will technology help people with disabilities become more transportation independent?
   a. Drones will deliver their packages.
   b. Scan codes can be used to track where a person is.
   c. GPS can be used to navigate traffic patterns.
   d. Automated vehicles will be developed.

10. Which option lists the computer components in order from oldest to newest?
    a. transistors, vacuum tubes, integrated circuits, personal computers, microprocessors
    b. integrated circuits, microprocessors, vacuum tubes, transistors, personal computers
    c. vacuum tubes, transistors, integrated circuits, microprocessors, personal computers
    d. microprocessors, vacuum tubes, transistors, personal computers, integrated circuits

11. A smart refrigerator can use ___ to detect when you are running low on milk, and then send a reminder to you on a wireless network. Answer: Sensors
12. The _______ layer encapsulates the segment into packets? Answer: Network
13. ________ is the use of computers to simulate a real or imagined environment that appears as a three-dimensional space. Answer: Virtual Reality
14. The only language understood by computer hardware is a/an _________ language. Answer: Machine Language
15. C, C++, and Java can be classified as _________ languages. Answer: High-level

User Satisfaction Survey (for the professor respondents)

Tracking Mouse Behavior to Detect Cheating Using Multiple Logistic Regression

We are 3rd year students taking up Bachelor of Science in Computer Science from the College of Computer Studies and Systems (CCSS). We would like to seek participation in our research titled "Tracking Mouse Behavior to Detect Cheating Using Multiple Linear Regression" by answering the attached survey.

Please take the time to read the details of our study:

1. Purpose of the study: Our research aims to observe the behavior of mouse movements made by students who are taking an exam through dishonest means.
2. Procedure: As a respondent, you will be asked to accomplish an electronic form that may include personal and sensitive information.
3. Confidentiality: All collected information will be used for academic and research purposes only and will be handled with utmost confidentiality. Presentation of results will be in the form of aggregate information, and will not identify a specific individual. Collected information in soft and hardcopy formats will be disposed of after printing the final manuscript.
4. Right to Decline: Your participation is voluntary and you have the right to decline or withdraw at any time.
5. Questions and Concerns: We are open to answering any questions or concerns. Please contact us.

Should you have any questions, please feel free to email or text:

Altoveros, Louis Angelo D.R. (altoveros.louisangelo@ue.edu.ph - 09064138199)
Lopez, Virgilio Jr. A. (lopez.virgiliojr@ue.edu.ph - 09174231378)
Manangat, Renz Darryl M. (manangat.renzdarryl@ue.edu.ph - 09989104419)
Tabo, Ralf Martin R. (tabo.ralfmartin@ue.ed - 09260427627)
Tiosan, John Christian (tiosan.johnchristian@ue.edu.ph - 09284298596)

Thank you!

Display a menu

lopez.virgiliojr@ue.edu.ph (not shared) Switch account Draft restored

lopez.virgiliojr@ue.edu.ph (not shared) Switch account

* Required

Part 1: Demographic Profile

Instructions: Please fill in the blank or check the box of your answer.

Full Name *

Your answer

Email Address *

Your answer

Age *

Your answer

College/Department *

- ◯ CBA
- ◯ CAS
- ◯ CCSS

Display a menu

Age *

Your answer

College/Department *

- CBA
- CAS
- CCSS
- CEDUC
- CENG
- CDENT

Next

Clear form

Tracking Mouse Behavior to Detect Cheating Using Multiple Logistic Regression

lopez.virgiliojr@ue.edu.ph (not shared) Switch account

* Required

Part 2: User Satisfaction Survey

This section pertains to your experiences in using our system. Please choose the answer which applies to you the most.

Content

Does the system provide the precise information you need? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Does the information content meet your needs? *

Does the information content meet your needs? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Does the system provide reports that seem to be just about exactly what you need? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Does the system provide sufficient information? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time

Display a menu

Does the system provide sufficient information? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Accuracy

Is the system accurate? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Are you satisified with the accuracy of the system? *

- 1 - Almost never
- 2 - Some of the time

Display a menu

Are you satisified with the accuracy of the system? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Format

Do you think the output is presented in a useful format? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Is the information clear? *

- 1 - Almost never

Display a menu

Is the information clear? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Ease of Use

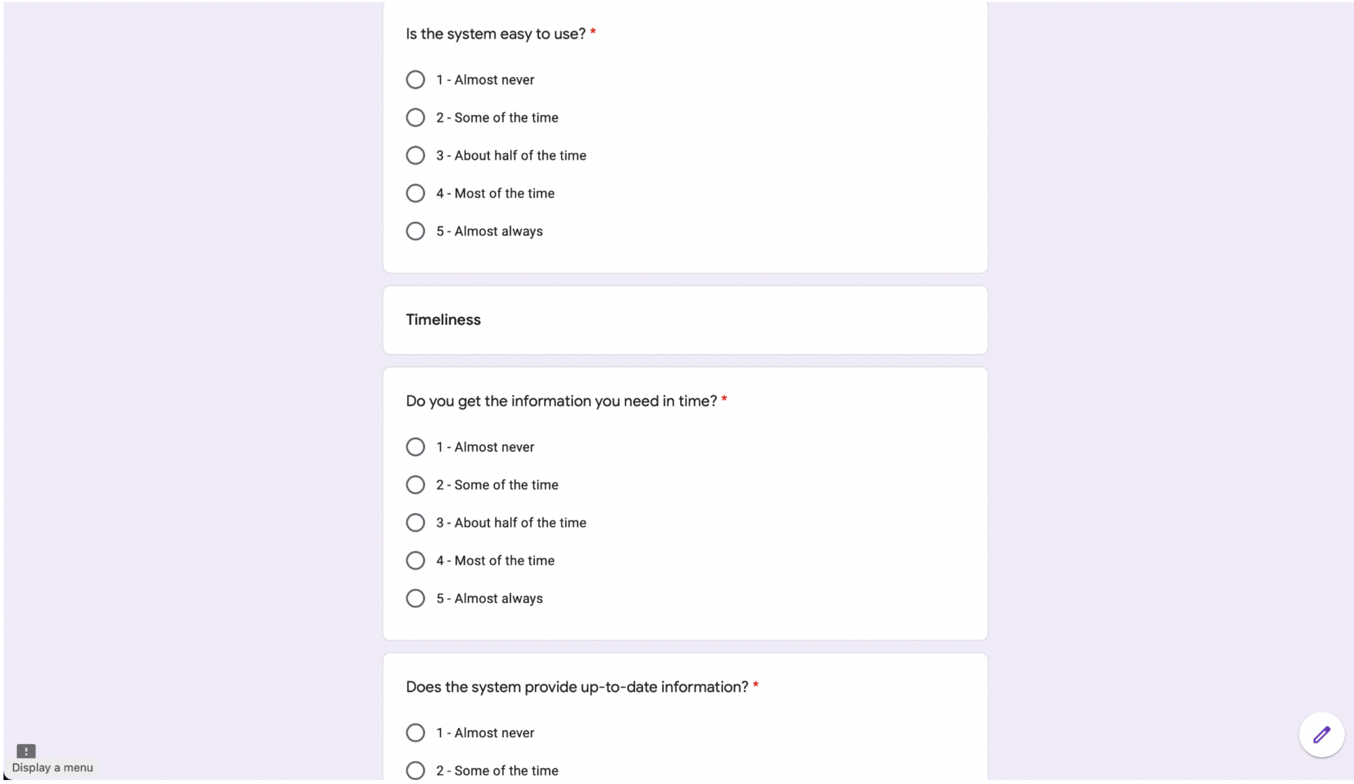
Is the system user-friendly? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Is the system easy to use? *

- 1 - Almost never
- 2 - Some of the time

Display a menu



Is the system easy to use? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Timeliness

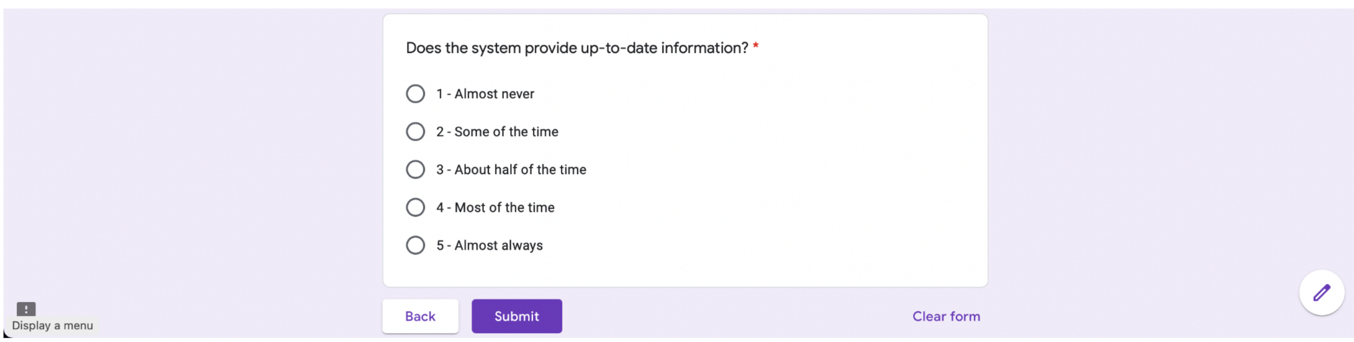
Do you get the information you need in time? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Does the system provide up-to-date information? *

- 1 - Almost never
- 2 - Some of the time

Display a menu



Does the system provide up-to-date information? *

- 1 - Almost never
- 2 - Some of the time
- 3 - About half of the time
- 4 - Most of the time
- 5 - Almost always

Back | Submit | Clear form

Display a menu

Comments of Panel Members

Comments of Panel	Actions Taken (Including the page number)
Show sample results of PageFocus and Sokout as baseline	Revised. See page 12-17. Literature review includes tables summarizing the performance of existing methods.
Consider keyboard as subject	Revised. See page 21. Scope generalized from ‘Mouse Behaviors’ to ‘Input Device Data’ (includes keyboard, and trackpad); Scope and Limitation.
Be specific with the mouse model	Revised. See page 21. Specified mouse to be used as a traditional mouse having only the basic parts: scroll wheel, left-click and right-click button; Scope and Limitation.
Be clear with the task of outperforming the baseline	Revised. See page 5 Hypothesis specifies the data to be surpassed, and the target increase; Hypothesis
Be able to show the difference between the two studies to show which is more effective	Revised. See page .
Include the additional modeling method algorithm to be used to show the relevance of the effectivity of the study	Revised. See page .
Either the specificity or the sensitivity should surpass Sokout’s metrics by at least 1 point	Revised. See page 5 Hypothesis specifies the data to be surpassed, and the target increase; Hypothesis