

I-Quiz: An Intelligent Assessment Tool for Non-Verbal Behaviour Detection

B. T. Shobana^{1,*} and G. A. Sathish Kumar²

¹Department of Information Technology, Sri Venkateswara College of Engineering, Sriperumbudur, 602117, India

²Department of Electronics and Communication Engineering, Sri Venkateswara College of Engineering, Sriperumbudur, 602117, India

*Corresponding Author: B. T. Shobana. Email: shobana@svce.ac.in

Received: 16 April 2021; Accepted: 29 May 2021

Abstract: Electronic learning (e-learning) has become one of the widely used modes of pedagogy in higher education today due to the convenience and flexibility offered in comparison to traditional learning activities. Advancements in Information and Communication Technology have eased learner connectivity online and enabled access to an extensive range of learning materials on the World Wide Web. Post covid-19 pandemic, online learning has become the most essential and inevitable medium of learning in primary, secondary and higher education. In recent times, Massive Open Online Courses (MOOCs) have transformed the current education strategy by offering a technology-rich and flexible form of online learning. A key component to assess the learner's progress and effectiveness of online teaching is the Multiple Choice Question (MCQ) assessment in most of the MOOC courses. Uncertainty exists on the reliability and validity of the assessment component as it raises a qualm whether the real knowledge acquisition level reflects upon the assessment score. This is due to the possibility of random and smart guesses, learners can attempt, as MCQ assessments are more vulnerable than essay type assessments. This paper presents the architecture, development, evaluation of the I-Quiz system, an intelligent assessment tool, which captures and analyses both the implicit and explicit non-verbal behaviour of learner and provides insights about the learner's real knowledge acquisition level. The I-Quiz system uses an innovative way to analyse the learner non-verbal behaviour and trains the agent using machine learning techniques. The intelligent agent in the system evaluates and predicts the real knowledge acquisition level of learners. A total of 500 undergraduate engineering students were asked to attend an on-Screen MCQ assessment test using the I-Quiz system comprising 20 multiple choice questions related to advanced C programming. The non-verbal behaviour of the learner is recorded using a front-facing camera during the entire assessment period. The resultant dataset of non-verbal behaviour and question-answer scores is used to train the random forest classifier model to predict the real knowledge acquisition level of the learner. The trained model after hyperparameter tuning and cross validation achieved a normalized prediction accuracy of 85.68%.

Keywords: E-Learning; adaptive and intelligent e-learning systems; intelligent tutoring systems; emotion recognition; non-verbal behaviour; knowledge acquisition level



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

Educational systems are looking for e-Learning programs [1] to address the challenges in the current traditional education system such as poor quality, high cost and limited access. E-Learning offers the possibility for the learners to learn irrespective of time and space. Today's world has far developed and hence the age of "standardization" and "One-size-fits-all" approach [2] no longer applies. As Technology and revolution are so ubiquitous in every aspect of our lives, "personalization" [3] has become very important. Personal Learning Environments [4] are becoming more relevant and demanding in today's academic world. Personalization refers to the teaching and learning process in which the stride of learning and the instructional strategies are customized for the needs of the individual learner. Learning objectives, instructional approaches, instructional content, and its sequencing differ for every learner, and hence personalization has a profound effect on the outcome [5].

Online courses [6] are becoming more prevalent in the educational field and have gained momentum. Massive open online courses (MOOCs) creates ample opportunities for knowledge sharing and helps in improving lifelong learning skills by providing easy access to global resources [7]. A darker side of MOOCs reveals that more than 90 per cent of those who sign up for MOOC courses do not complete the courses [8], so pitches up questions about the efficacy of the model. Although the MOOCs e-learning model is scalable, the low course completion rates and low effectiveness do not make it a sustainable model [9]. In order to cater to the needs of students with different backgrounds, personalization is especially essential for effective Learning. As a first leap towards personalization, understanding the learner's behaviour and evaluating the comprehension level of learners is important.

To achieve a greater impact, few pedagogical transformations in MOOCs [10] must be done in order to reduce the attrition rates and achieve a greater impact. A key component in the MOOC's platform for assessing the learner progress is the Multiple Choice Question (MCQ) assessment. MCQs are ubiquitous in all kind of education and are perceived as a time-efficient method of assessment. This type of assessment tool enables representative sampling of broad areas of course content [11]. However, the reliability of the assessment score reflecting the real knowledge level of the learner is still a million-dollar question.

Few points that play a key role in validating the effectiveness of MCQ assessments are

- a) The answer to the question is guaranteed in any one of the given options.
- b) The majority of the questions are fact-based or definition/principal based that certainly reduces the complexity of the assessment.
- c) Random guessing or smart guessing can also procure marks for the learner, which is definitely not possible in an essay type assessment.
- d) The confidence level of the learner is not taken into account, say the learner can choose/change the options any number of times. These would not affect the final score of a candidate.
- e) Time taken by the user to answer a question does not reflect in the final score.

This paper presents the I-Quiz system, a novel and intelligent online MCQ assessment tool for evaluating the real knowledge acquisition level of the learner by capturing the non-verbal behaviour of the learners during the assessment period. I-Quiz is developed as a part of an adaptive and personalized e-learning platform that can act as an intelligent agent and evaluate the real comprehension level of learners. The real knowledge acquisition level can be used by the E-learning platform to provide customized learning path to the learners and thus facilitate personalized learning.

2 Materials and Methods

E-learning facilitates acquisition of new skills and knowledge without the constraints of time, space and environment due to the enormous advancements in information technology [12]. Individuals learners differ

in their style/preferences and hence providing personalized and customized learning experiences would have a significant impact on learning outcomes [13]. Adaptive e-Learning aims to incorporate approaches and practices to online learners and thus provide a personal and unique learning experience, with the final goal of maximizing their performance [14]. Adaptive e-Learning addresses the fact that every learner is unique with different background, educational needs and learning styles [14]. MOOCs ascended from the incorporation of Information and communication technologies in the field of education, characterized by unrestricted participation and open access to the resources [15]. A major concern in the current MOOC system is the high attrition rate, survey says that there is a high drop out of about 90%. Personalized learning, customized learning path and sensing the learner emotions can increase the retention rate of learners [16,17].

Understanding learner's emotions play a vital role in providing feedback about them to the E-learning system [18,19]. The term 'affect' refers to the range of affective emotional states and feelings that one can experience while engaged in the learning process. Affective computing has emerged as one of the essential components in intelligent learning systems, since emotion plays an essential role in decision making, perception and learning [20]. Extensive research has been done on the methods and models of affect detection systems that are capable of handling conventional affect detection modalities such as facial expression, voice, body language and posture, physiology, brain imaging, multi-modality systems [21]. Research that conceptualizes the role of human emotions in learning intertwined the phases into four quadrants (curiosity, confusion, frustration, hopefulness) where the horizontal axis represents the emotions and the vertical axis represents the learning [22]. Intelligent systems that incorporate emotional intelligence using sensors were able to identify the four physiological signals of human emotions or their affective states. Several features were extracted over the obtained signals and by applying Sequential Floating Forward Search and Fisher Projection algorithms, eight emotional states were accurately recognized [23]. Affective AutoTutor, a smart tutoring system assists the learner by communicating with them using natural language similar to human dialogues and simulates a human tutor. The smart tutoring system can handle the specified instruction and their interaction to higher precedence by detecting and reacting to the learner's cognitive states and their corresponding emotional states, thus showing development in cognitive dynamics [24].

Non-verbal behaviour is a comprehensive term to denote any communicative behaviour not involving verbal conversation [25,26]. In an online learning context where learners are geographically remote and essentially use a computer, non-verbal behaviour includes the learner's facial expressions, Eye behaviour, different gestures, body posture, and body movement. MOOCs, the current trend in online learning facilitates only limited interaction between instructors and learners [10,27]. However, non-verbal communication indicates the emotion that often acts as a catalyst and a driver of learning success [28,29]. Attention to non-verbal behaviour in online learning will definitely lead to a profound effect on the learner's performance. Implicit non-verbal communication is a very powerful tool to identify the learner emotions. The implicit type captures the learner activity, analyzes and then process to extract the facial expressions, body language, gestures, postures, and blink rate. The implicit behaviour indirectly measures the required parameters using the Haar cascade classifier algorithm [30] for blink rate detection, Xception model [31] for emotion recognition, Manhattan distance [32] for body posture and gestures. Explicit type refers to the measure of direct communication that is clear and straightforward. It does not require any processing techniques to measure the parameters. With the Quiz interface, the explicit behaviour such as the number of click counts done while attempting to answer a question (changes in selecting an option before submitting the answer is a measure of uncertainty) and the time taken by the learner to attempt the question are measured.

3 I-Quiz System

I-Quiz System, an intelligent assessment tool aims to quantify the real comprehension level of learners after they attempt multiple choice assessment module as a part of e-learning. I-Quiz application was designed

and coded to present randomly ordered multiple choice questions on the topics of advanced C Programming. Participants were unaware that their implicit and explicit non-verbal behaviour is captured in real-time. There was no pre-set time limit for an individual question, but a time constraint of 25 minutes was set for the entire quiz to be completed. During each question slot, the participant and his behaviour were recorded using a front-facing web camera. The ground-truth MCQ answer along with the captured non-verbal behaviour such as the blink rate, manhattan norm and zero-norm for the body posture, facial expression, click count and clocking time for every individual question collectively forms the dataset.

3.1 Study Procedure and Participants

The study uses data collected from 500 undergraduate students from the Department of Information Technology in Sri Venkateswara College of Engineering. Participants have prior programming experience in the C language as they have learnt the language as part of the curriculum during the first year of the programme. There were 268 male participants and 232 female participants. Participants were aged from 18 and 21 with a median age of 19. The undergraduate students were asked to complete a 20 MCQ quiz using the I-Quiz system (Fig. 1a) on topics related to advanced C programming. Every question in the MCQ is mapped to the topic/chapter in the Learning Management System. While the learner attends the assessment, both the implicit and explicit behaviour would be simultaneously (without the knowledge of the learner) captured using a front-facing web camera attached to one of the preconfigured laptops provided for the assessment.

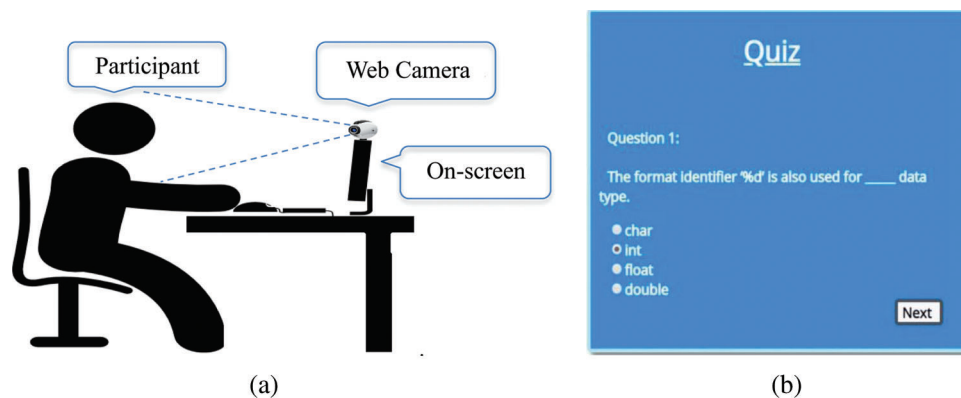


Figure 1: (a) Physical layout of the experiment (b) Screen shot of the quiz system (user interface)

I-Quiz system developed for the experiment (Fig. 1b) displayed a series of MCQs in random order. The questions included conceptual, code snippets, output prediction and error finding in advanced C programming. During each question period, the video stream from the webcam is captured and converted to frames in specified time intervals.

3.2 Non-Verbal Behaviour Extraction

I-Quiz system aims to classify and assess the performance of the learner in the assessment activity. The system architecture comprised of five modules and these modules simultaneously execute in the background to capture the Non-verbal behaviour of the learner as specified in Fig. 2. A background process activates the webcam as soon as the learner logs in to the quiz system as a registered user, records the video of the assessment question wise, simultaneously converts the video into image frames at regular intervals and feeds it to the other modules to extract the non-verbal behaviour of the learner. The image frames from the video stream are fed to the implicit behaviour capturing modules. Algorithm 1 represented below summarizes the steps involved in each of the modules to generate feature vectors from the non-verbal behaviour of learners.

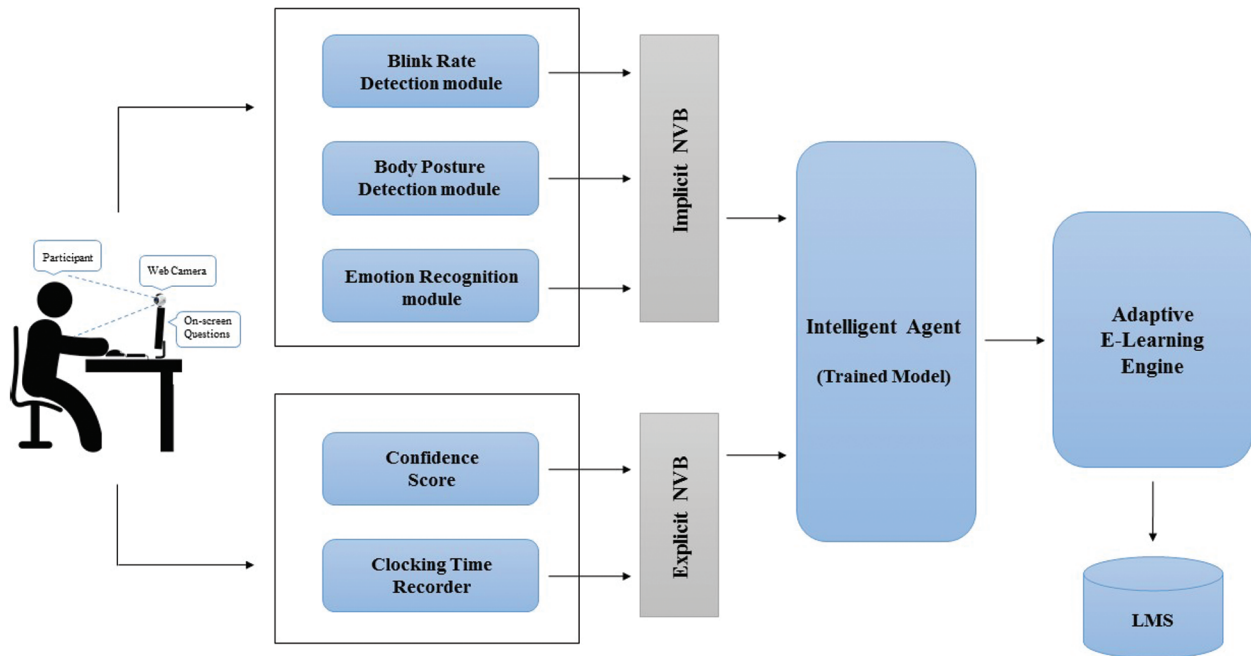


Figure 2: System architecture

Algorithm 1: Generating feature vector from non-verbal behaviour of learners

Set of learners $L = \{l_1, l_2, l_3, \dots, l_n\}$

Set of questions for individual learner $Q = \{q_1, q_2, \dots, q_m\}$

E_i – emotion of the learner

M_i – manhattan distance (maximum)

ZN_i – normalised zero norm

EBR_i – blink rate

C_i – confidence score

T_i – clocking time

$Score_{q_i}$ – marks awarded (0 – incorrect/1- correct)

Input : Continuous video stream of the learners L_i ($i=1$ to n) attending the quiz session

For each learner $l \in L$

1) Process the video stream V_l of each learner and convert it in to video chunks V_{q_i} for each individual question period

For each question $q \in Q$ process the video chunks V_{q_i} and execute the following steps 2 to 8 concurrently

2) Identify the emotion E_i of the learner using the image frames from the video by “Haar cascade classifier” and “Xception model”.

3) Measure the average manhattan distance M_i and the zero-norm ZN_i by processing the successive frames in the video.

(Continued)

Algorithm 1 (continued).

-
- 4) Calculate the eye blink rate EBR_i using facial landmark and eye aspect ratio for individual question period.
 - 5) Capture the click patterns of the learner while answering the questions and calculate the confidence score C_i .
 - 6) Record the clocking time T_i spent to answer every individual question.
 - 7) Evaluate the score for the answered questions and award $Score_{qi}$
 - 8) Generate the feature vector $F = \{E_i, M_i, ZN_i, EBR_i, C_i, T_i, Score_{qi}\}$ for individual question

end for

- 9) Add the feature vector to the dataset (Training and Testing)

end for**Output:**'m' feature vectors $\{\{F1\}, \{F2\}, \dots, \{Fm\}\}$ for an individual learner.

In order to analyze the non-verbal behaviour, the individual output parameters from the modules in the I-Quiz system are together collected to form a feature vector containing seven features (Time taken, Emotion, Manhattan norm, Zero norm, Eye blink rate, C score, Correct/Incorrect). The data set comprised of 20 feature vectors collected from individual learners.

3.2.1 Eye Blink Rate Detection

The comportment of an eye is one of the powerful non-verbal gestures and it has the ability to infer the non-verbal behaviour of learners. Blinking is a natural eye movement defined as the rapid closing and opening of the eyelid. The variations in the speed, frequency, and strength of blinks provide emotional evidence to the observer. Blink rate remains constant for any person and a study says that the rate can change due to external stimulus [33]. Average blink rates of a person range from 10 to 20 blinks per minute. Intellectual activities by a person can have a significant effect on the blink rate. The study says that there is a significant increase in blink rate of around 20 blinks per minute during conversation and verbal recall [33].

Factors related to cognitive, visual and memory tasks significantly influence the blink rate [34]. Other emotional states such as anger, anxiety or excitement affect the blink rate of a person. The blink rate detection module as depicted in Fig. 3, detects the number of blinks per minute of the learner, measured during the time taken to attend every question and formed as one of the features in the generated dataset.

3.2.2 Body Posture Estimation

One measure of emotional concentration is the body posture and the body movement during any learning or assessment activity. The learner is stable and concentrated during the assessment activity only when the understanding level and confidence in answering the question is high. Identifying the concentration level with the body posture and body movement helps to quantify the involvement and interest in the activity. Abnormal head rotation and frequent posture changes can definitely reflect the negative emotion of the learner during the activity. In the I-Quiz system, a background process activates the webcam as soon the learner starts the session and continuously records the assessment activity as a video. Over each question period, the video is converted into image chunks (frames) in pre-specified intervals. One metric that can find and measure the changes in frequent body posture movement is the manhattan distance. The measure helps in finding the difference in body posture at frequent intervals.

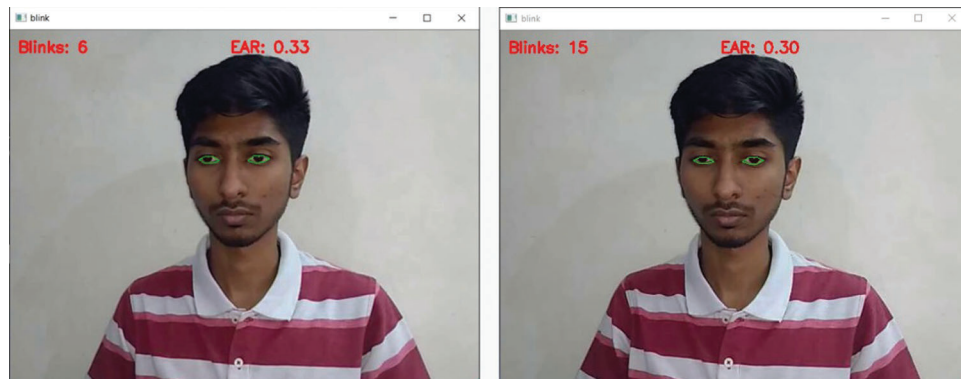


Figure 3: Blink rate detection

A metric that can quantify the similarity between two images is the manhattan distance [32], also termed as the L1 metric. The similarity (in our case the dissimilarity) is found by calculating the distance (difference) between the two pixels in images. Manhattan distance ‘d’ will be calculated using an absolute sum of the difference between its cartesian coordinates.

Let x and y be the image frames from the video at subsequent intervals. The manhattan distance ‘d’ is calculated as specified in Eq. (1).

$$d = \sum_{i=1}^n |x_i - y_i| \quad (1)$$

The sum of absolute differences (SAD) [35] is used to measure the similarity between images. In our case, it is calculated by taking the absolute difference between each pixel in the first image and the corresponding pixel in another image used for comparison. These individual pixel-wise differences are summed up to create a simple metric of image similarity, the L1 norm or manhattan distance of the images. Manhattan distance of consecutive frames for less body movement is depicted in Fig. 4a whereas manhattan distance of consecutive frames for a distracted learner is depicted in Fig. 4b.

In the proposed method, the intelligent agent in the assessment tool captures the body posture of the learner from the video in the form of image frames at pre-defined intervals and calculates the average manhattan difference score. A higher score infers the non-verbal behaviour that the learner is highly stressed.

3.2.3 Emotion Recognition

The emotions of an online learner during the MCQ assessment play a vital role in analyzing non-verbal behaviour. To understand and interpolate the non-verbal behaviour, it is vital to evaluate the emotions of the learner. According to the literature, one important way to detect emotion is to observe the symptoms through facial expression. Facial expressions of the learners are significantly correlated to their emotions that help to recognize their comprehension level. Four expressions of learners say happy, sad, angry and neutral are recognized using the Haar cascade classifier [26] (face detection) and the Xception model [27] (emotion recognition) is detected as illustrated in Fig. 5. The recognized expression from the module is added as a feature to the data set.

3.2.4 Confidence Score and Clocking Time

A learner is very confident if she/he attempts to answer the correct option only once. When the learner attempts to change/choose more than one option during the answering window, it reflects the



Figure 4: (a) Manhattan distance of consecutive frames for less body movement (b) Manhattan distance of consecutive frames for distracted learner

lower confidence level. Even if the learner lands up answering the right option after a series of changes from one answer to the other, the factor that he was not confident in the topic must be taken into account. Hence, a confidence score (in percentage) that reflects the knowledge level of the learner is validated from his clicking behaviour, which is non-verbal and explicit. If the learner has chosen the correct option only once, then he/she is awarded a confidence score of 100. Else if the behaviour is recorded that the learner has switched between different options before freezing one, his score is calculated using the [Eq. \(2\)](#).

$$\text{Confidence Score} = \frac{\text{No. of clicks on the correct option}}{\text{Total no. of clicks}} * 100 \quad (2)$$

The time spent by the learner (in seconds) to answer every question was recorded. The total clocking time of the user was also considered as a feature in the data set. The amount of time a learner spends on answering a question is an indicator of the confidence level of the learner.

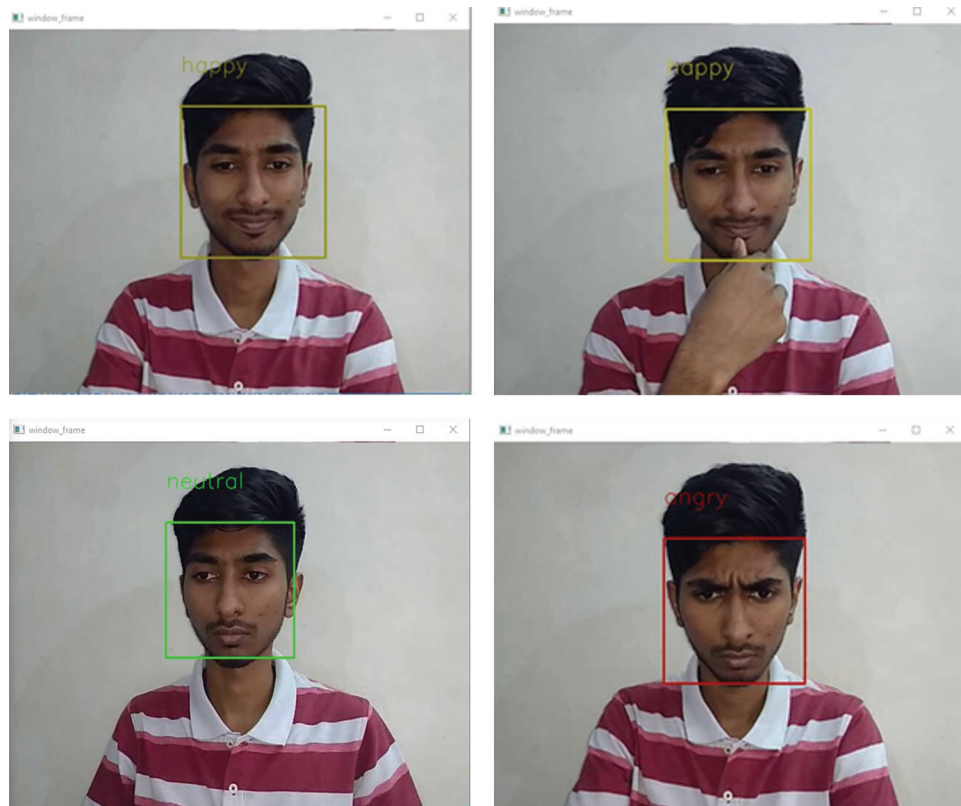


Figure 5: Emotion recognition of learner

4 Experimental Results

4.1 Data Set and Feature Selection

Ensemble learning is a powerful machine learning technique that combines several base machine learning models to produce an optimal predictive model for efficient classification and regression. Random forests [36] is a supervised ensemble learning algorithm used for both classifications and regression problems and is effective in reducing the problem of overfitting [37] the model.

The random forests machine learning model applies the method of bootstrap aggregating to decision tree learners. The random forest model trained with a training set $X = x_1, x_2, \dots, x_n$ with responses $Y = y_1, y_2, \dots, y_n$ bagging repeatedly (W times) selects a random sample with replacement of the training set and fits the decision trees generated to these samples:

For $W = 1, \dots, w$:

1. Sample with replacement, n training examples from X, Y (X_W, Y_W)
2. Train a classification tree T_W on X_W, Y_W .

Using the trained model, predictions for unseen samples can be made by taking the majority vote from the generated classification trees. The bootstrapping procedure in the random forest leads to better model performance as it leads to a decrease in the variance of the model, without increasing the bias and hence it was chosen for model training.

The backend of the I-Quiz system was implemented using python language due to its versatile nature and ability to support multithreading. The user interface and the quiz module were implemented using

JSP. The students attended the MCQ assessment using the I-Quiz system and the collected dataset was pre-processed. A random forest classification model was programmed (in python), optimized by tuning the hyperparameters, trained and tested with the collected data set. Binary classification is performed as the model predicts if a learner would answer the question correctly (1) or incorrectly (0). The dataset collected from 500 undergraduate participants each populating 20 feature vectors is split into two sets, training set ($Tr = 80$) and testing set ($Ts = 20$). The classifier was initially set with the parameter values hardcoded in the program. A snapshot of a few samples in the dataset collected from the learners during the assessment activity is specified in [Tab. 1](#). The summary of statistics on the collected dataset is depicted in [Tab. 2](#). As the summary applies only to numerical features, values are not generated for emotion (categorical feature).

Table 1: Sample of the dataset collected from learners

Qno	Time taken (Seconds)	Blink count	Blinks per minute	Manhattan norm	Max of norm	Click count	Emotion	Marks
1	28	5	11	0.101319	0.506274	2	Neutral	0
2	76	16	13	0.023523	0.046521	1	Surprise	0
3	50	4	5	0.025933	0.044861	1	Surprise	1
4	65	16	15	0.025622	0.047361	3	Neutral	0
5	75	22	18	0.02794	0.056398	2	Sad	0
6	90	42	28	0.024751	0.05833	1	Surprise	0
7	10	3	18	0.020306	0.023732	1	Surprise	1
8	59	13	13	0.028375	0.066599	1	Surprise	0
9	21	0	0	0.02356	0.034573	1	Happy	1
10	45	19	25	0.025324	0.044159	2	Surprise	0
11	11	4	22	0.019057	0.024769	1	Neutral	1
12	10	16	96	0.026078	0.032752	1	Sad	1
13	49	19	23	0.022465	0.035193	1	Surprise	0
14	8	4	30	0.019336	0.024052	1	Neutral	1
15	18	5	17	0.0284	0.062675	1	Surprise	0
16	45	17	23	0.025863	0.068666	1	Neutral	1
17	24	13	33	0.023532	0.036372	1	Surprise	0
18	20	8	24	0.025289	0.035949	1	Happy	1
19	14	4	17	0.023564	0.029783	1	Surprise	1
20	20	3	9	0.018416	0.035308	1	Neutral	0

Feature importance indicates the relative significance of each feature when making predictions from a trained model and has the class of techniques for assigning scores to input features. It is an important interpretation tool to identify relevant features that contribute to the performance and accuracy of the prediction process. Before training the model, the input features are tested for importance and the least significant features are not included in the training process as specified in [Fig. 6](#). It is observed that the emotion of the learner and the number of clicks attempted for each question (click count) during the answering period seemed less significant as compared to the other features.

Table 2: Summary of statistics-dataset

	Time Taken (Seconds)	Blink count	Manhattan norm (Average)	Max. of norm	Click count	Emotion (Categorical)
Mean	39.34	14.99	0.025201	0.054275	1.235	NA
Std dev	35.13	22.75	0.013075	0.057975	0.63806	NA
Min	4	0	0.005372	0.006688	1	NA
25%	17	1	0.017083	0.027930	1	NA
50%	30	7	0.022442	0.039584	1	NA
75%	50	20	0.029823	0.058244	1	NA
Max	418	209	0.101873	0.610723	9	NA

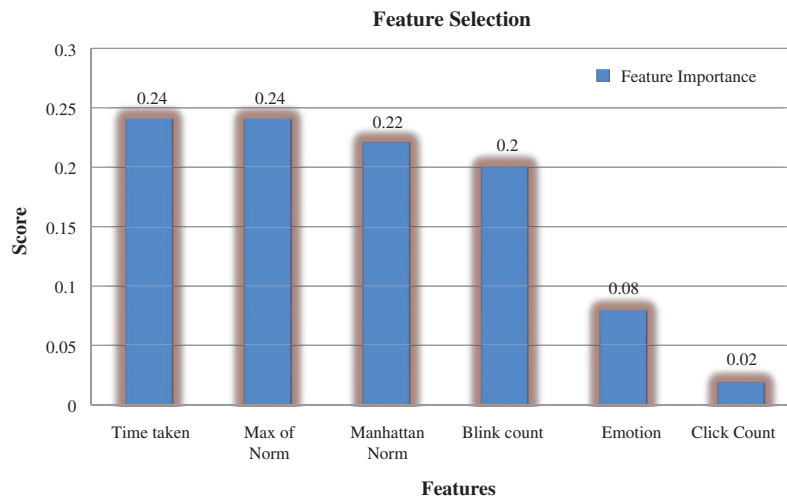


Figure 6: Feature importance score

4.2 Classifier Training and Prediction Accuracy

The classifier was initially trained using a dataset with the parameter values given in Tab. 1. The variable ‘mark’ forms the class label for the training and the rest of the variables form the feature set. The random forests model classified either if the student had answered the question correctly with knowledge acquisition (1) or had randomly guessed the answer (0). The model was evaluated with the classification accuracy score metric and found to have a classification accuracy of 76.8%. To further improve the accuracy of the trained random forests model, hyperparameter tuning was performed on the dataset. The technique of hyperparameter tuning is the key to optimize the random forests model trained with the dataset initially. Due to hyperparameter tuning during the training process, model accuracy and computational efficiency had a significant improvement. The initial parameters and the tuned parameter values are listed in Tab. 3. Hyperparameter tuning incorporated into the random forests achieved a considerable increase of 2.7 percent and had an accuracy score of 79.5%.

One good statistical method to estimate the skill of the random forest model is cross-validation. K-fold cross-validation is a powerful preventive measure against overfitting the model and generalizes to an independent dataset. A series of K fold validation with various values of K was performed (K = 5, 10,

15). K-fold cross-validation during the model training was done and the performance was measured. The model with cross-validation achieved an improved accuracy of 85.68% as specified in Fig. 7.

Table 3: Hyperparameter tuning results

S. No	Model parameters	Parameter value	Hyper tuned results
1.	n_estimators	100	400
2.	random_state	42	42
3.	max_features	Auto	Sqrt
4.	max_depth	None	6
5.	min_samples_split	1	10
6.	min_samples_leaf	2	5
7.	Bootstrap	True	False
8.	Cross validation	None	5
9.	Criterion	Gini	Gini

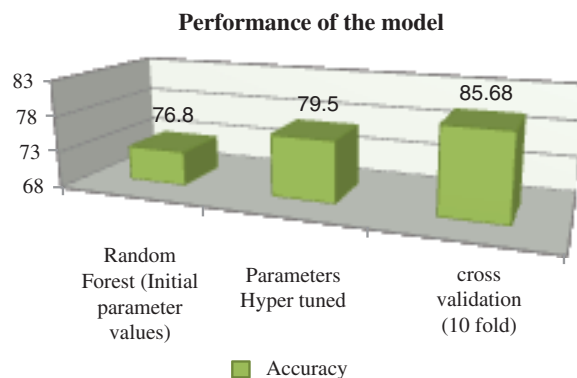


Figure 7: Performance of the classifier

During the assessment activity, when a new learner attempts the MCQ, the tuned, cross-validated model was effective in identifying if he/she had possessed the required knowledge level to answer the question. The agent does not award the marks directly for every correct answer, instead makes an inference from the predicted results. As given in Tab. 4, the agent takes the appropriate decision in awarding the marks. The trained model was effective in finding out if the learner had randomly guessed the answer. The agent does not award the marks to the learner for such attempts and the domain/topic pertaining to the question is reported to the adaptive e-learning engine for further intervention. Hence the trained model that acts as an intelligent agent is capable of assessing the real knowledge level of learners and aids in providing a personalized experience to the learners.

Table 4: Mark evaluation from the predicted results

Mark obtained for a question/answer in assessment	Mark predicted by the intelligent agent from the non-verbal behaviour	Action taken	Inference by the agent	Mark assigned by the agent
1	1	Marks awarded for the answer	Learner had the appropriate knowledge	1
0	0	Marks not awarded for the answer	Learner lacked knowledge	0
0	1	Marks not awarded for the answer	Learner made a very quick attempt to answer/over – confident	0
1	0	Marks not awarded for the answer.	Learner was not confident/ knowledgeable and had made a smart guess	0

5 Conclusion

This paper presented the design, development and evaluation of the I-Quiz system, a novel and intelligent MCQ assessment tool. The contribution to the literature is to exhibit that the non-verbal behaviour of learners could be analyzed to estimate the real knowledge level of learners during an assessment activity. The MCQ assessment tool can in real-time assess the learner and identify the specific area in which he/she lags in knowledge acquisition (even if the learner had randomly guessed and answered the question correctly). The intelligent agent as a part of the E-learning platform can help in providing a personalized learning experience to the learners. This is possible due to the ability of the agent to assess the learners with their non-verbal behaviour in a personalized manner to provide adaptive content delivery. The system can provide timely and personalized intervention points in an intelligent e-learning platform.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] N. S. Chen, I. L. Cheng and S. W. Chew, “Evolution is not enough: Revolutionizing current learning environments to smart learning environments,” *International Journal of Artificial Intelligence in Education*, vol. 26, no. 2, pp. 561–581, 2016.
- [2] E. J. Brown, T. J. Brailsford, T. Fisher and A. Moore, “Evaluating learning style personalization in adaptive systems: Quantitative methods and approaches,” *IEEE Transactions on Learning Technologies*, vol. 2, no. 1, pp. 10–22, 2009.
- [3] A. Garrido and L. Morales, “E-Learning and intelligent planning: Improving content personalization,” *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, vol. 9, no. 1, pp. 1–7, 2014.

- [4] U. Markowska-Kaczmar, H. Kwasnicka and M. Paradowski, "Intelligent techniques in personalization of learning in e-learning systems," in *Computational Intelligence for Technology Enhanced Learning*. Vol. 273. Berlin, Heidelberg: Springer, pp. 1–23, 2010.
- [5] L. A. M. Zaina, G. Bressan, J. F. Rodrigues and M. A. C. A. Cardieri, "Learning profile identification based on the analysis of the user's context of interaction," *IEEE Latin America Transactions*, vol. 9, no. 5, pp. 845–850, 2011.
- [6] O. Almatrafi and A. Johri, "Systematic review of discussion forums in massive open online courses (MOOCs)," *IEEE Transactions on Learning Technologies*, vol. 12, no. 3, pp. 413–428, 2019.
- [7] N. Zhu, A. Sari and M. M. Lee, "A systematic review of research methods and topics of the empirical MOOC literature (2014–2016)," *Internet and Higher Education*, vol. 37, no. 1, pp. 31–39, 2018.
- [8] T. Eriksson, T. Adawi and C. Stöhr, "Time is the bottleneck: A qualitative study exploring why learners drop out of MOOCs," *Journal of Computing in Higher Education*, vol. 29, no. 1, pp. 133–146, 2017.
- [9] K. S. Hone and G. R. El Said, "Exploring the factors affecting MOOC retention: A survey study," *Computers & Education*, vol. 98, no. 2, pp. 157–168, 2016.
- [10] E. B. Gregori, J. Zhang, C. Galván-Fernández and F. D. A. Fernández-Navarro, "Learner support in MOOCs: Identifying variables linked to completion," *Computers & Education*, vol. 122, no. 4, pp. 153–168, 2018.
- [11] C. P. M. Van Der Vleuten and L. W. T. Schuwirth, "Assessing professional competence: From methods to programmes," *Medical Education*, vol. 39, no. 3, pp. 309–317, 2005.
- [12] A. D. Dumford and A. L. Miller, "Online learning in higher education: Exploring advantages and disadvantages for engagement," *Journal of Computing in Higher Education*, vol. 30, no. 3, pp. 452–465, 2018.
- [13] R. M. Felder and L. K. Silverman, "Learning and teaching styles in engineering education," *Engineering Education*, vol. 78, no. 7, pp. 674–681, 1988.
- [14] C. A. Carver, R. A. Howard and W. D. Lane, "Enhancing student learning through hypermedia courseware and incorporation of student learning styles," *IEEE Transactions on Education*, vol. 42, no. 1, pp. 33–38, 1999.
- [15] T. Belawati, "Open education, open education resources, and massive open online courses," *International Journal of Continuing Education and Lifelong Learning*, vol. 7, no. 1, pp. 1–15, 2014.
- [16] L. Shen, M. Wang and R. Shen, "Affective e-learning: Using "emotional" data to improve learning in pervasive learning environment," *Educational Technology & Society*, vol. 12, no. 2, pp. 176–189, 2009.
- [17] M. Imani and G. A. Montazer, "A survey of emotion recognition methods with emphasis on e-learning environments," *Journal of Network and Computer Applications*, vol. 147, no. 2, pp. 102423, 2019.
- [18] R. Al-Shabandar, A. J. Hussain, P. Liatsis and R. Keight, "Analyzing learners behaviour in MOOCs: An examination of performance and motivation using a data-driven approach," *IEEE Access*, vol. 6, pp. 73669–73685, 2018.
- [19] J. C. Hung, K. C. Lin and N. X. Lai, "Recognizing learning emotion based on convolutional neural networks and transfer learning," *Applied Soft Computing*, vol. 84, 105724, pp. 1–19, 2019.
- [20] C. H. Wu, Y. M. Huang and J. P. Hwang, "Advancements and trends of affective computing research," *British Journal of Educational Technology*, vol. 47, no. 6, pp. 1304–1323, 2016.
- [21] R. A. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 18–37, 2010.
- [22] B. Kort, R. Reilly and R. W. Picard, "An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion," in *Proc. IEEE Int. Conf. on Advanced Learning Technologies*, Madison, WI, USA, IEEE, pp. 43–46, 2001.
- [23] R. W. Picard, E. Vyzas and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1175–1191, 2001.
- [24] S. D'Mello and A. Graesser, "AutoTutor and affective autoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back," *ACM Transactions on Interactive Intelligent Systems*, vol. 2, no. 4, pp. 1–39, 2012.

- [25] A. S. Won, J. N. Bailenson and J. H. Janssen, "Automatic detection of nonverbal behaviour predicts learning in dyadic interactions," *IEEE Transactions on Affective Computing*, vol. 5, no. 2, pp. 112–125, 2014.
- [26] A. D. Dharmawansa, Y. Fukumura, A. Marasinghe and R. A. M. Madhuwanthi, "Introducing and evaluating the behaviour of non-verbal features in the virtual learning," *International Education Studies*, vol. 8, no. 6, pp. 82–94, 2015.
- [27] K. Hew and W. Cheung, "Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges," *Educational Research Review*, vol. 12, no. 6, pp. 45–58, 2014.
- [28] M. Holmes, A. Latham, K. Crockett and J. D. O'Shea, "Near real-time comprehension classification with artificial neural networks: Decoding e-learner non-verbal behaviour," *IEEE Transactions on Learning Technologies*, vol. 11, no. 1, pp. 5–12, 2018.
- [29] T. Robal, Y. Zhao, C. Lofi and C. Hauff, "IntelliEye: Enhancing MOOC learners' video watching experience through real-time attention tracking," in *Proc. of the 29th on Hypertext and Social, Media, ACM*, New York, USA, pp. 106–114, 2018.
- [30] I. Gangopadhyay, A. Chatterjee and I. Das, "Face detection and expression recognition using Haar cascade classifier and Fisherface algorithm," in *Recent Trends in Signal and Image Processing, Advances in Intelligent Systems and Computing*. Vol. 922. Singapore: Springer, pp. 1–11, 2019.
- [31] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 1251–1258, 2017.
- [32] M. D. Malkauthekar, "Analysis of euclidean distance and manhattan distance measure in face recognition," in *Third Int. Conf. on Computational Intelligence and Information Technology (CIIT 2013)*, Mumbai, India, pp. 503–507, 2013.
- [33] A. R. Bentivoglio, S. B. Bressman, E. Cassetta, D. Carretta, P. Tonali *et al.*, "Analysis of blink rate patterns in normal subjects," *Movement Disorders*, vol. 12, no. 6, pp. 1028–1034, 1997.
- [34] A. D. Dharmawansa, K. T. Nakahira and Y. Fukumura, "Detecting eye blinking of a real-world student and introducing to the virtual e-learning environment," *Procedia Computer Science*, vol. 22, no. 1, pp. 717–726, 2013.
- [35] M. J. Atallah, "Faster image template matching in the sum of the absolute value of differences measure," *IEEE Transactions on Image Processing*, vol. 10, no. 4, pp. 659–663, 2001.
- [36] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [37] L. Dong, H. Du, F. Mao, N. Han, X. Li *et al.*, "Very high resolution remote sensing imagery classification using a fusion of random forest and deep learning technique—Subtropical area for example," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 113–128, 2020.