Journal of Advances in Computing, Communications and Information Technology (JACCIT)

Journal homepage: www.sciengtexopen.org/index.php/jaccit

An Implementation of K-NN Classification Algorithm for Detecting Impersonators in Online Examination Environment

I. O. Izu-Okpara *, O. C. Nwokonkwo, A. M. John-Otumu

Department of Information Technology, Federal University of Technology Owerri, Nigeria

ARTICLE INFO

Article history:

Received 24 March 2021

Received in revised form
30 March 2021

Accepted 21 April 2021

Available online
25 April 2021

Keywords:

Computer Networks
Impersonation
K-NN Classification
Machine Classification
Online Examination

ABSTRACT

The online examination platforms also known as computer-based testing (CBT) platforms for conducting mass-driven examinations over computer networks to eliminate certain issues such as delay in marking, misplacement of scripts, monitoring, etc., associated with the conventional Pen and Paper Type (PPT) of examination have also been bedeviled with the issue of impersonation commonly associated with the PPT system. The existing online examination platforms rely on passive mechanisms such as the CCTV system and the human invigilators for monitoring the examination halls against cheating and impersonation. The proposed model integrates some level of intelligence into existing online examination prototype by designing and developing an intelligent agent service that could assess students against impersonation threat in an online examination environment using the K-Nearest Neighbor (K-NN) machine learning classification technique considering the level of accuracy and response time in answering the questions. A total of 3,083 dataset was downloaded from an online repository; 80% (2,466) of the dataset was used for training the model, while 20% (617) dataset was used in testing the model to enable the model detect unseen data correctly. Results showed that the developed model has a 99.99% accuracy rate, precision, recall and f-score.

1. Introduction

Examination is simply seen as a regular use and standard factor for evaluating students understanding and reasoning capability [1]-[3]. Ismail and Shami [4] opined that examinations can also control the amount of which instructive purposes have stayed accomplished as well as the extent to which enlightenment has been establishments in order to assist the public in general. The rapid enhancement of information and communication technologies (ICT) has sharpened a new paradigm-shift for teaching and learning over the Internet and other network environments, which have undeniably compelled a thoughtful change from the conventional pen and paper type (PPT) of examination to online examination also known as computer-based testing (CBT), e-assessment, computer-based examination (CBE), e-examination, and web-based examination [4]-[5].

The application of information technology (IT) for assessment and related activities has paved way for a well-crafted and engineered online examination platform, which permits educators and instructors to program reviews, questions, experiments and examinations over the Internet and intranet [1]-[3], [6], [7]. Until now, the standard method for testing in an online examination environment has been the Triple-A model which stands for "Assembling, Administering and Appraising" as proposed by [8]. The Triple-A model was recommended for adoption as a reference point in the development of a standard online testing or examination application as necessitated by [9].

* Corresponding author
E-mail address: floxy83@yahoo.co.uk
https://doi.org/10.37121/jaccit.v1.153

It is noted by [10] that CBT applications can be installed either as a desktop-based application or a web-based application over the Internet or intranet using web-based technologies. Basically, the two types of CBT applications in existence for online examinations are the *Linear-test* and the *Adaptive-test* [1], [6], [11]. According to research conducted by [12]-[16], CBT systems offers several benefits: delivery of very few human resources and reduced paper work; examination can be scheduled at any time and can be done from anyplace; and improved testing capacity of partakers.

Online examination platforms are specifically developed for the conduct of mass-driven examinations over computer networks in order to eradicate known challenges such as delay in marking, misplacement of scripts, monitoring, etc., associated with the conventional PPT examination. The existing online examination application (Fig. 1) uses passive methods such as CCTV and human invigilators to monitor and detect impersonators during examination by physically moving from one system to the other in an examination hall to confirm the students' faces against their pictures on their respective system dashboard which takes so much time and effort and at the end of the day only very little or nothing is achieved due to the large crowd and the peculiar examination environment; many students still impersonate without being detected. However, this paper seeks to address the problem of examination impersonation associated with the existing online examination platforms, which are on the increase as observed by [6], [17] using machine learning technique.

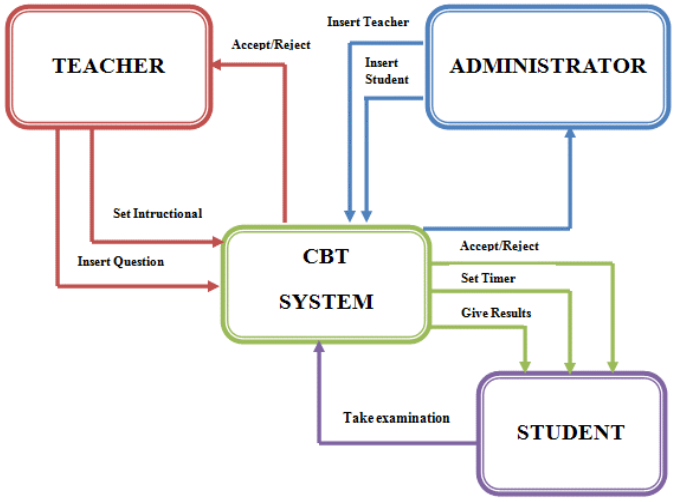

Fig. 1 Existing online examination data flow diagram [1], [5].

2. Related Work

Fagbola et al. [5] proposed an online examination system. The waterfall model was used as the software life cycle. Microsoft SQL Server 2008, Macromedia Dreamweaver 8.0, and Microsoft Visual Studio 2012 were used to implement the software prototype. The proposed CBT developed system was designed using the 3-tier architecture with basic security such as username/ password and users' privileges. The developed CBT system supported the linear-test administration and also supported the triple-A standard but lacked resumption capability, multi-level security measures and not intelligent to detect impersonation.

Also, Ajinaja [1] proposed a 3-tier online examination system using both waterfall and component-based software engineering models. PHP, JavaScript, HTML, MySQL, CSS and XAMP Server were used in the development. The proposed system has basic security features such as username/ password and users' privileges. The proposed system also supported the triple-A standard but lacked resumption capability, multi-level security measures and not intelligent to detect impersonation.

Suleiman and Nachandiya [11], proposed an online examination system using the agile software model. Development environments such as PHP, MySQL, JavaScript, CSS, HTML and XAMP Server were used. The proposed system supported security features such as username/ password, users' privileges, triple-A standard but also lacked resumption capability, multi-level security measures, and also not smart to detect impersonation.

Omotehinwa and Durojaye [18], proposed an online examination system using the 3-tier architecture; and modeled using sequence and use case diagrams. PHP, HTML, and MySQL were used to develop the proposed system with basic security features such as username/ password & users' privileges. However, the system lacked resumption capability, multi-level security measures and not smart enough to detect impersonation.

Ismail and Soye [3] designed and implemented a secure online examination system using biometric fingerprint and Advanced Encryption Standard (AES). Java programming language and MySQL were used to implement the system. The proposed system supported the triple-A standard, but lacked resumption capability, multi-level security measures, and not smart enough to detect impersonation.

Al-Saleem and Ullah [17] mitigated the threat associated with computer-based testing application for online examination. The system is a hybridization of the conventional username/password verification structure with palm-based biometric authentication was implemented. The system also supported the triple-A standard but lacked resumption capability, multi-level security measures, and not smart enough to detect impersonation. Also, Adebayo and Abdulhamid [2], mitigated the threat associated with computer-based testing application for online examination and the integrity of questions and results using biometric fingerprint technology, username/password and cryptographic technique. Java Applet, HTML, PHP, and MySQL were used to implement the proposed system. The system also supported the triple-A standard, but lacked resumption capability, multi-level security measures, and not smart to detect impersonation.

Kuyoro et al. [6] developed a 3-tier web-based online examination system for conducting online examination. The waterfall model was used as the system development life cycle. Cascading Style Sheets (CSS), JavaScript, PHP, HTML, MySQL and WAMP Server were used in developing the software prototype with basic security features such as username/ password & users' privileges. The system supported the triple-A standard but lacked resumption capability, multi-level security measures, but not smart to detect impersonation.

Oluwole [19] developed a desktop-based online examination system using Java programming language and MySQL with basic security features such as username/ password & users privileges only. The system supported the triple-A standard but lacked resumption capability, multi-level security measures, and not smart to detect impersonation.

Zheming et al. [20] also developed a novel electronic examination system using distributed Component Object Module Browser/server framework technologies such as Ajax, PHP, HTML and MySQL. The system supported the triple-A standard, cryptographic function to secure transmission, username/ password and users' privileges, questions encryption but lacked randomization of questions/distribution choice & resumption capability, multi-level security measures and not smart to detect impersonation.

Yuan-Lung et al. [21] developed a 3-tier web-based online examination system using Visual Basic Script in Active Server Page (ASP), Microsoft Access. The system supported username/ password & users' privileges, triple-A standard but lacked resumption capability, multi-level security measures and not smart to detect impersonation.

Indoria et al. [22] developed a web-based online examination system using ASP.NET, VB.NET, Microsoft Access with basic security features such as username/ password and users' privileges. System supported the triple A-standard but cannot generate random questions; teachers cannot login directly to the system, lacked resumption capability, multi-level security measures, and not smart to detect impersonation.

Rashad et al. [23] developed an Arabic web-based examination management system using AJAX, PHP, HTML and MySQL. System supported the triple-A standard, multiple questions type generation, instructors can directly login to the system, but lacked resumption capability, random questions generation & random choice distribution, multi-level security measures and not smart to detect impersonation.

Henke [24] proposed a web-based Test, Examination and Assessment System (WETAS) using Java Applet and PHP. The developed system supported multiple questions type generation, triple-A standard, eLearning but lacked resumption capability, random questions generation & random choice distribution, multi-level security measures and not smart to detect impersonation.

Ayo et al. [25] developed a model for an electronic examination system that supported the triple-A standard, multiple questions type generation, but lacked resumption capability, random questions generation & random choice distribution, multi-level security measures and smart to detect impersonation.

Akinsanmi et al. [26] proposed a web-based examination system using ASP.NET web server, C#, ADO.NET, and Microsoft SQL Server. The system supported the triple-A standard, multiple questions type generation, but lacked resumption capability, random questions generation & random choice distribution, multi-level security measures and not smart to detect impersonation.

Tasci et al. [27] developed an online examination system that supported the triple-A standard, multiple questions types generation, and an intelligent agent to detect shortage of time but lacked resumption capability, multi-levels security measures and not smart to detect impersonation.

Qiao-fang [28] developed an online examination system using JavaBean, JSP, Tomcat, and JavaScript. System supported multiple questions type generation such as Yes/No, MCQ, Fill-in gaps, numeric and essay questions, instructor's direct login to the system, but lacked resumption capability, random questions generation & random choice distribution, multi-level security measures and not smart to detect impersonation.

Adewale et al. [29] developed a 3-tier web-based online examination system using VB.NET, ASP.NET. System supported multiple questions type generation, but lacked direct login of lecturers to the system, resumption capability, random questions generation & random choice distribution, multi-level security measures and not smart to detect impersonation.

Mohammed and Maysam [30], proposed an application-based online examination system with resumption and randomization capabilities using spiral model software life cycle. The system was developed using Java programming language and Derby database. The system supports the triple-A standard, multi-instructor login, resumption capability, random question selection, distribution and choices selection but lacked multiple questions type generation, multi-level security measures and is not smart to detect impersonation.

3. Methodology

This section describes the different methods put together in achieving the design and development of the proposed system.

3.1. Data Collection Technique

The following data collection techniques were used in gathering information about the propose system:

(a) Key Information Interview Technique (KIIT) for interviewing seasoned experts in the IT industry.
(b) Observation of existing industry-based CBT software products such as Moodle LMS and JAMB CBT system.
(c) Download of published papers on CBT applications from Open Journal Access.
(d) Download of Dataset from online repository for training and testing the developed model.

3.2. Development Environment

The development environments used in this research work are as follows:

(a) Visual Studio Code: (HTML/CSS, PHP and JavaScript) used in coding the entire web platform
(b) Jupyter Notebook: for executing the Python language used in the dataset preparation, preprocessing and algorithm evaluation.
(c) XAMPP: the web server used in hosting the entire platform on a local host.

3.3. Proposed System Notation

The proposed system integrated intelligence into a prototype online examination platform and was modeled using notations defined by equation (1).

$$IOEP = \{OEP, IAS\} \quad (1)$$

Where, IOEP is the Intelligent Online Examination Platform, OEP (online examination platform) is the Computer-Based Testing System, and IAS is the Intelligent Agent Services interacting within the online examination platform in carrying out its specific function of detecting impersonators using K-NN classification algorithm.

3.4. Proposed System Architecture

The system architectural diagram shows the blueprint of the proposed model for detecting impersonators during the online examinations. Fig. 2 shows the architectural blueprint of the proposed model for detecting examination impersonators using an intelligent agent service designed with K-NN classification algorithm for analyzing and

classifying the students for any possible case for impersonation inclinations before giving the students control to write the actual online examination as compared to Fig. 1 that gives the students direct control to write the online examination without any form of active verification by the CBT application.

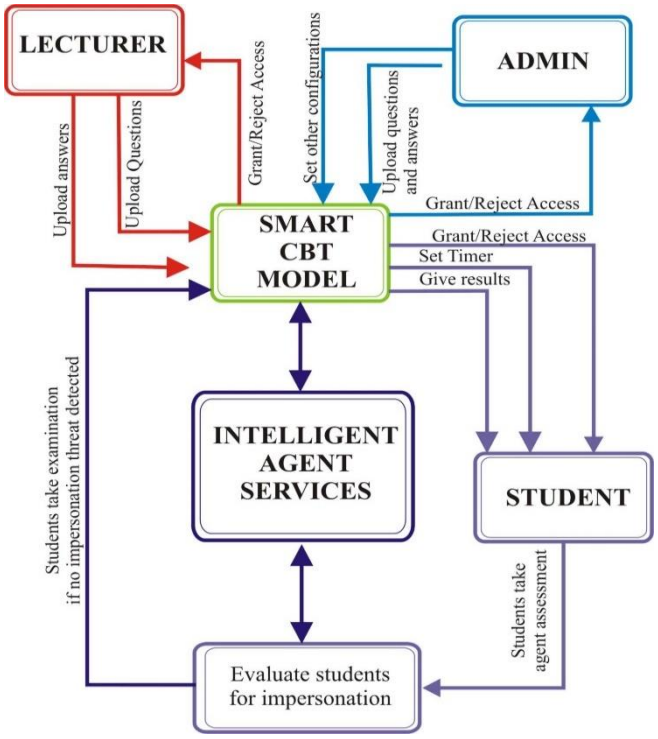

Fig. 2 Proposed system architecture.

3.5. Algorithm Adopted

The choice of adopting the K-NN machine learning classification algorithm for implementing the intelligent agent services responsible for detecting likely impersonation threats was based on the fact that the technique is well suited for many class labels and multi-modal classes, and finally, it has a low error rate when compared to other machine learning techniques such as naïve bayes. The K-NN implementation steps are as follows:

Step 1: For implementing the algorithm, a dataset is need. So, during the first step of K-NN, the training as well as test data must be loaded.
Step 2: Next, choose the value of K, i.e., the nearest data points. K can be any integer.
Step 3: For each point in the test data, do the following for (step 4 to step 7)
Step 4: Compute the distance between test data and each row of training data with the help of any of the method namely: Euclidean, Manhattan or Hamming distance. Though the most commonly used method to calculate distance is Euclidean.
Step 5: Based on the distance value, you can now sort them in ascending order.
Step 6: Next, it will choose the top K rows from the sorted array.
Step 7: Now, it will assign a class to the test point based on most frequent class of these rows.
Step 8: End

4. Results and Discussion

This section shows some of the user's responsive interfaces developed as a result of research work, and also evaluation done using confusion matrix to test the level of false alarm rates.

4.1. Responsive Interfaces

Fig. 3 shows the administrator's dashboard of the proposed system. The dashboard enables the administrator to carry out certain functions and also reveals certain events that have taken place within a single screen view. Fig. 4 shows the interface through which the administrator can activate the examination agent from the proposed system.

Fig. 5 shows the student's dashboard in order to enroll/take an examination, and also check their results through the proposed system.

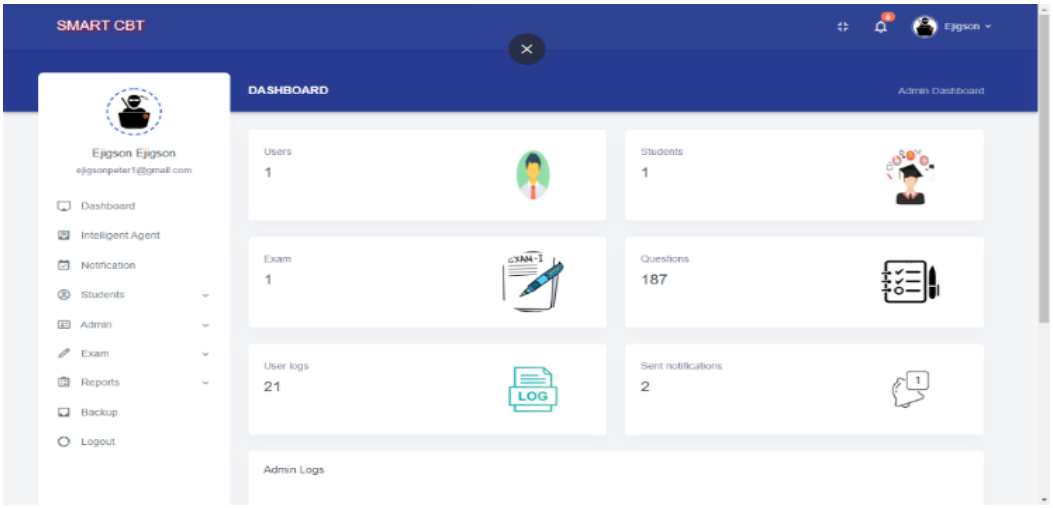

Fig. 3 Admin dashboard interface.

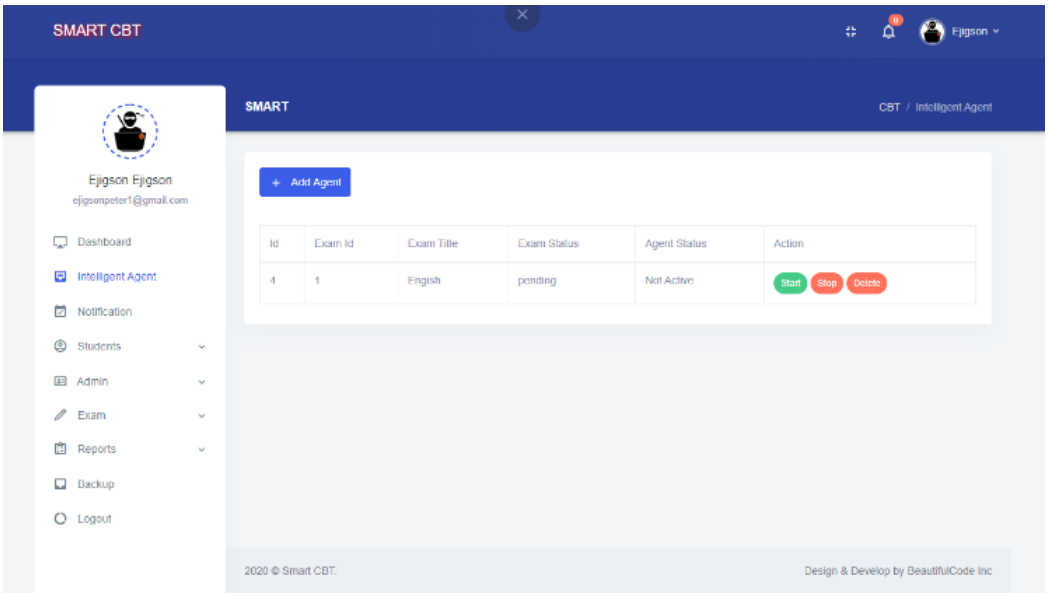

Fig. 4 Admin dashboard to activate the agent services.

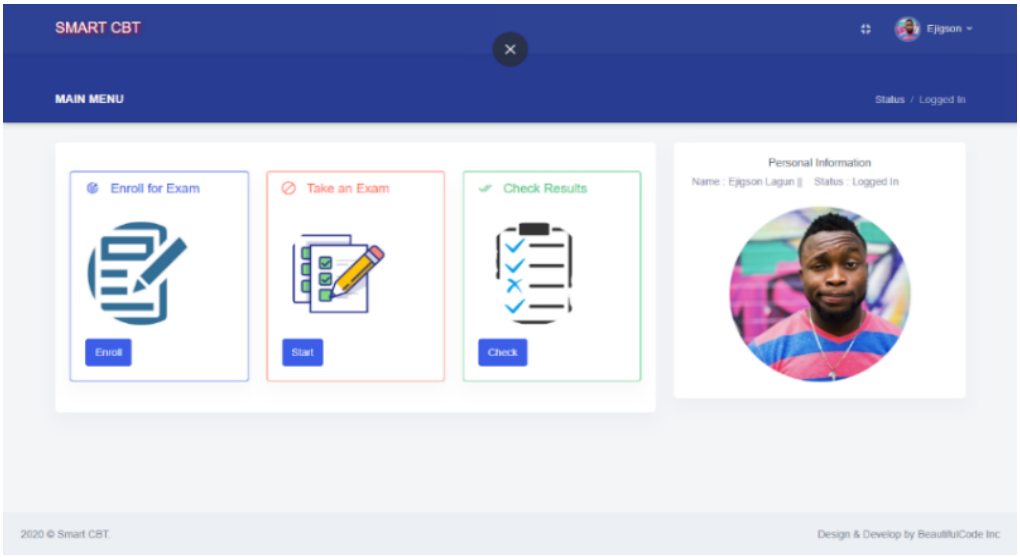

Fig. 5 Student's dashboard.

Fig. 6 shows the loading interface of the intelligent agent service for detecting impersonators. Fig. 7 shows the intelligent agent service assessment interface for testing each student before giving them access to the real examination environment if they pass the test. Fig. 8 shows the pop-up screen of an impersonation threat detected by the proposed system. Fig. 9 shows the student's examination interface for taking online examination in the proposed system.

Fig. 6 Intelligent agent service loading.

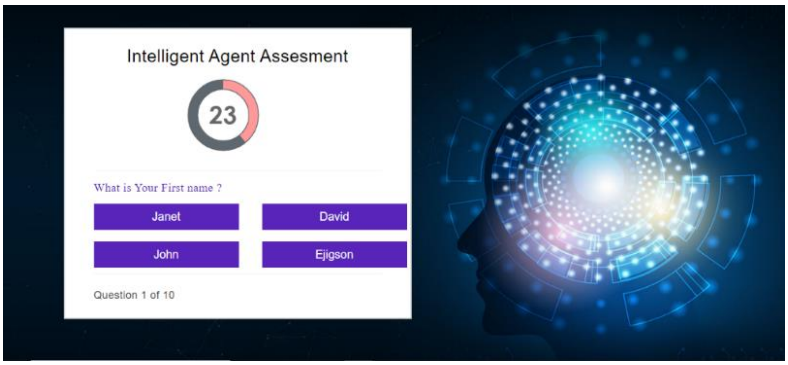

Fig. 7 Intelligent agent assessment interface.

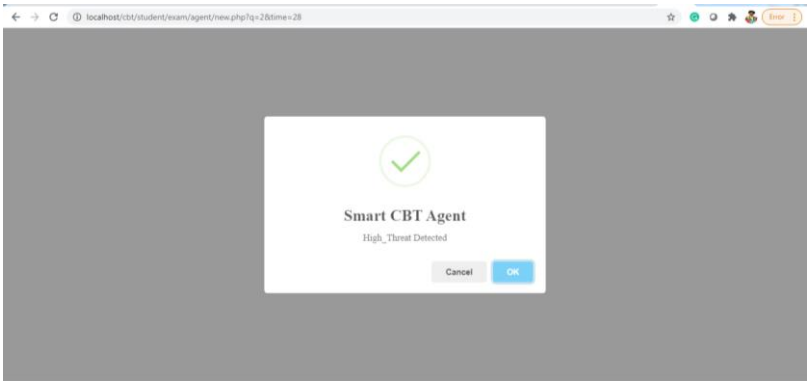

Fig. 8 Screen shot showing impersonator detected.

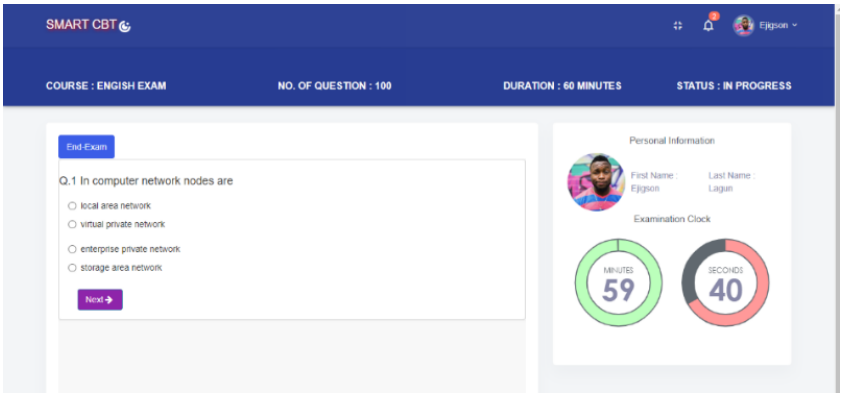

Fig. 9 Student's examination interface.

4.2. Performance Evaluation

The performance of the developed model was evaluated using the confusion matrix. The level of classification rate or accuracy rate, recall, precision and f-score were all considered in this evaluation with a total number of 617 test cases. Table 1 shows the predicted records with 124 test cases as low threat predicted, 77 test cases as no threat predicted and 416 test cases as high threat predicted.

Table 1 Predicted test cases.

N = 617	Predicted: low_threat	Predicted: no_threat	Predicted: threat
Actual: low_threat	124	0	0
Actual: no_threat	0	77	0
Actual: high_threat	0	0	416

Table 2 shows the confusion matrix breakdown of the 617 test cases. The breakdown further revealed that the 124 low threat test cases were True Positive (TP); the 77 no threat test cases were also TP, while the 416 high threat test cases were True Negative (TN).

Table 2 Confusion matrix false alarm breakdown.

N = 617	Predicted: low_threat	Predicted: no_threat	Predicted: Threat	
Actual: low_threat	TP = 124	FP = 0	FP = 0	124
Actual: no_threat	FN = 0	TP = 77	FN = 0	77
Actual: high_threat	FN = 0	FN = 0	TN = 416	416
	124	77	416	

Note: FP is False Positive, and FN is False Negative

*Calculation of Classification Rate /Accuracy***:** Classification Rate or Accuracy is given by the relation:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (2)$$

$$\text{Accuracy} = \frac{(TP + TP + TN)}{(TP + TP + TN + FP + FN + FN + FN + FN+ FN)}$$

$$\text{Accuracy} = \frac{(124+77+416)}{(124+77+416+0+0+0+0+0+0)}$$

$$\text{Accuracy} = 1.00$$

Calculation of Recall: Recall gives us an idea about when it is actually a yes, how often does it predict yes.

$$\text{Recall (No Threat)} = \frac{(TP)}{(TP + FN)} \quad (3)$$

$$\text{Recall (No Threat)} = \frac{(77)}{(77 + 0)}$$

$$\text{Recall (No Threat)} = 1.00$$

Calculation of Precision: it can be expressed as equation (4).

$$\text{Precision (No Threat)} = \frac{(TP)}{(TP + FP)} \quad (4)$$

$$\text{Precision (No Threat)} = \frac{(77)}{(77 + 0)}$$

$$\text{Precision (No Threat)} = 1.00$$

Calculation of F-measure: it can be expressed as equation (5).

$$\text{F1-score (F-Measure)} = \frac{(2 * \text{recall} * \text{precision})}{(\text{recall} + \text{precision})} \quad (5)$$

$$\text{F1-score (F-Measure)} = \frac{(2 * 1.0 * 1.0)}{(1.0 + 1.0)}$$

$$\text{F1-score (F-Measure)} = 1.00$$

The general evaluation results revealed that the K-NN algorithm used in building the intelligent agent services to detect impersonators during online examinations was 99.99% accurate.

5. Conclusion

This paper examined twenty published articles on online applications for writing examination downloaded from open journal access and also observed two industry-based software products on CBT platform (Moodle LMS and JAMB CBT software) based on the increasing trend of examination impersonation and curbing such cases that is quite worrisome due to the level of menace it has introduced into the online examination environment [6], [17]. The research work adopted the triple-A model standard (Assembling, Administering and Appraising) as proposed and recommended by [8], [9]. Special intelligent agent service was built into the existing CBT model using K-NN machine learning classification algorithm based on its strengths for active monitoring and detection of impersonators at a near real-time situation. The proposed model can detect possible cases of impersonation during online examination based on the intelligent internal control mechanism with a 99.99% detection accuracy when compared to other online examination systems examined [1]-[3], [5], [6], [11], [17]-[30], which relied exclusively on passive techniques such as CCTV and human invigilators as part of its internal control mechanism for detection of impersonators. Future research road map should consider hybridized techniques and deep neural networks for designing the intelligent control mechanism and also multi-level security measures integration to existing online examination platforms.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

ORCID

A. M. John-Otumu https://orcid.org/0000-0002-3138-4639

References

[1] M. Ajinaja, “The design and implementation of a computer-based testing system using component-based software engineering,” *International Journal of Computer Science and Technology*, vol. 8, no. 1, pp. 58 – 65, 2017.

[2] O. Adebayo, and S. M. Abdulhamid, “E-exams system for Nigerian universities with emphasis on security and results integrity,” *International Journal of the Computer, the Internet and Management*, vol. 18, no. 2, pp. 1 – 12, 2009.

[3] H. M. Ismail, and B. M. Soye, “Biometric enabled computer-based testing system with advanced encryption standard,” *Journal of Emerging Technologies and Innovative Research*, vol. 5, no. 8, pp. 1 – 8, 2018.

[4] M. Ismail, and P. A. Shami, “Computer-based vs paper-based examinations: perception of university teachers,” *The Turkish Online Journal of Educational Technology*, vol. 11, no. 4, pp. 371 – 381, 2012.

[5] T. M. Fagbola, A. A. Adigun, and A. O. Oke, “Computer based test (CBT) system for university academic enterprise examination,” *International Journal of Scientific and Technology Research*, vol. 2, no. 8, pp. 336 – 342, 2013.

[6] S. O. Kuyoro, G. U. Maminor, R. U. Kanu, and O. Akande, “The design and implementation of a computer-based testing system,” *Journal of Applied Computation*, vol. 1, no. 1, pp. 1 – 7, 2016.

[7] E. Delen, “Enhancing a computer-based testing environment with optimum item response time,” *Eurasia Journal of Mathematics, Science & Technology Education*, vol. 11, no. 6, pp. 1457-1472, 2015.

[8] N. E. Gronlund, and R. L. Linn, “Measurement and evaluation in teaching,” 6th Edition, MacMillan, New York, 1990.

[9] T. H. Wang, S. C. Huang, W. L. Wang, and L. Y. Kuo, “Development of a web-based assessment system: A case study,” Proceedings of the 6th Global Chinese Conference on Computers in Education/National Education Informalization Forum, Beijing Normal University, Beijing, China, pp. 385 – 388, 14 – 16 June 2002.

[10] A. R. Taylor, “A future in the process of arrival: Using computer technologies for the assessment of student learning,” Kelowna, British Columbia: Society for the Advancement of Excellence in Education, 2005

[11] A. Suleiman and N. Nachandiya, “Computer based testing (CBT) system for GST exams in Adamawa State University, Mubi,” *Asian Journal of Research in Computer Science*, vol. 2, no. 1, pp. 1 – 11, 2018

[12] H. Jeong, “A comparative study of scores on computer-based tests and paper-based tests,” *Behavior and Information Technology*, vol. 33, no. 4, pp. 410 – 422, 2014.

[13] Y. P. Chua and Z. M. Don, “Effects of computer-based educational achievement test on test performance and test takers’ motivation,” *Computers in Human Behavior*, vol. 29, no. 5, pp. 1889 – 1895, 2013.

[14] F. Kaya and E. Delen, "A computer-based peer nomination form to identify gifted and talented students," *The Australasian Journal of Gifted Education*, vol. 23, no. 2, pp. 29 – 36, 2014.
[15] P. Schatz and J. Browndyke, "Applications of computer-based neuropsychological assessment," *Journal of Head Trauma Rehabilitation*, vol. 17, no. 5, pp. 395 – 410, 2002.
[16] C. P. Newhouse, "Computer-based exams in schools: freedom from the limitations of paper," *Research and Practice in Technology Enhanced Learning*, vol. 8, no. 3, pp. 431 – 447, 2013.
[17] S. M. Al-Saleem and H. Ullah, "Security considerations and recommendations in computer-based testing," *The Scientific World Journal*, vol. 3, no. 4, pp. 1 – 7, 2014.
[18] T. O. Omotehinwa and D. S. Durojaye, "Computer based test: Security and results integrity," *International Journal of Computer and Information Technology*, vol. 2, no. 2, pp. 324 – 329, 2013.
[19] O. Oluwole, "Design and implementation of an open-source computer-based testing system with end user impact analysis in Africa," *International Journal of Modern Education and Computer Science*, vol. 8, no. 1, pp. 17 – 24, 2015.
[20] Y. Zhenming, Z. Liang, and Z. Guohua, "A novel web-based online examination system for computer science education," 33rd ASEE/IEEE Frontiers in Education Conference, S3F-7-S3F-10. 2003.
[21] Y. Yuan-Lung, H. Tsung-Chih, C. Li-Chaun, "Development of a web-based online examination system," *East Sea Management Comment*, vol. 7, no. 1, pp. 109 – 120, 2005.
[22] V. Indoria, P. Sharma and A. Soni, "Online examination," *International School of Informatics and Management*, 2012.
[23] M. Z. Rashad, M. S. Kandil, A. E. Hassan, and M. A. Zaher, "An Arabic web-based examination management system," *International Journal of Electrical and Computer Sciences*, vol. 10, no. 1, pp. 35 – 41, 2010.
[24] K. Henke, "Web-based test, examination and assessment system," *Advanced Technology for Learning*, vol. 4, no. 3, 2007; DOI: 10.2316/Journal.208.2007.3.208-0911
[25] C. K. Ayo, I. O. Akinyemi, A. A. Adebiyi, and U. O. Ekong, "The prospects of e-examination implementation in Nigeria," *Turkish Online Journal of Distance Education*, vol. 8, no. 4, pp. 125 – 135, 2007.
[26] O. Akinsanmi, O. T. R. Agbaji, and M. B. Soroyewun, "Development of an e-assessment platform for Nigerian universities," *Research Journal Applied Sciences, Engineering and Technology*, vol. 2, no. 2, pp. 170 - 175, 2010.
[27] T. Taşci, Z. Parlak, A. Kibar, N. Taşbaşi, and H. I. Cebeci, "A novel agent-supported academic online examination system," *Educational Technology & Society*, vol. 17, no. 1, pp. 154 – 168, 2014.
[28] Z. Qian-fang, and L. Yong-fei, "Research and development of online examination system," In proceedings of the 2012 2nd International Conference of Computer and Information Applications, 2012.
[29] O. A. Adewale, T. O. Ajadi, and J. O. Inegbedion, "Perception of learners on electronic examination in open and distance learning institutions: a case study of National Open University of Nigeria," *US-China Education Review*, vol. 5, no. 1, pp. 639 – 644, 2011.
[30] I. Y. Mohammed and S. H. Maysam, "Construction of an online examination system with resumption and randomization capabilities," *International Journal of Computing Academic Research*, vol. 4, no. 2, pp. 62 – 82, 2015.