University Of AL-Qadisiyah
Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Development of a machine learning-based online test system with a hierarchical structure for students

Ruqaia Jwad Kadhim

University of Information Technology and Communication, College of Business Informatics, Department of Informatics Systems Management (ISM), Baghdad ,Iraq. roqaia.jwad@uoitc.edu.iq

ARTICLE INFO

Article history:
Received: 25 /04/2026
Rrevised form: 22 /05/2026
Accepted : 24 /05/2026
Available online: 30 /06/2026

Keywords:

Online test

Fyzze logic

Language Processing

Extraction Machin learning

ABSTRACT

A significant data amount is being exchanged, turning internet to a contemporary Silk Road for information. Machine learning (ML) is a burgeoning discipline that is progressively employed for diverse purposes. Artificial intelligence (AI) is the area of study that grants machines the capacity to demonstrate intelligent behavior. In recent decades, there have been notable developments in the fields of ML and deep learning. Devices with low processing capacity are being equipped with advanced algorithms and technology. The machine learning-powered online student testing system frees students from the limitations of conventional paper-based testing and leads to the improvement of the effectiveness of testing procedure. At the same time, it results in maintaining fairness in evaluating students' performance while improving the efficiency of grading. The objective of the present work is creating an internet-based assessment system for students that utilizes machine learning to improve the evaluation of college courses. The main focus of research and design lies in functional modules, fundamental technologies, and implementation of the on-line testing system. The advanced educational software and on-line evaluation system can be helpful for the schools in developing a more systematic and rigorous administration. The conclusion highlights that the online examination system, which is an advanced and reliable educational software, can aid schools in implementing systematic and efficient management procedures

https://doi.org/10.29304/jqcsm.2026.18.22844

∗Corresponding author: ruqaia jwad kadhim

Email addresses: roqaia.jwad@uoitc.edu.iq

Communicated by ‘sub etitor’

1. Introduction

Important developments have been noticed within the E-learning systems in recent years. The reason behind this witness is the quick progress in the techniques of Artificial Intelligence (AI) and in machine learning as well. To add more, educational institutions have directed their focus towards employing smart systems concerning the processes of assessment and testing that are electronically conducted, i.e. online ones, aiming at improving the accuracy of measuring students' level and reducing the problems that may accompany traditional tests. A number of novel trends have appeared concerning this field. Chief among them are the electronic systems of exams that depend on dynamically adjusting questions, mainly their level of difficulty, during the exam. This is performed the student's performance, a matter that helps in fulfilling more accurate, fair, and flexible evaluation than those noticed during traditional fixed-level exams. The systems of Computerized Adaptive Testing (CAT) has been referred to by recent studies. These systems are among those that are the most efficient techniques used for improving the process of evaluation and minimizing the test let alone maintaining the results accuracy (1, 2).

What is noticed in most conventional systems of exam is that most of them depend on introducing the same questions to all students regardless of their different academic levels or their promptitude concerning the questions. This leads to weaken the ability of differentiation among the different-level students. To add more, the reliance of traditional tests on a fixed-questions pattern may reduce or the efficiency of evaluation due to the existence of very easy questions for some students or very difficult to some others. Accordingly, an urgent need for developing smart systems of examinations that are able to analyze student's performance and make adaptive decisions related to questions –level of difficulty during the exam depending on direct indicators like speed and accuracy or correctness of the answers (1).

Among the essential indicators employed for assessing the level of student during the test are the time and correctness of answer. Via the time and correctness of answer, student's ability to treat various levels of questions in a realistic way. A student who introduces correct answers during a short time may have a higher knowledge-level than their peers as compared to some other student who needs a long period of time or frequently gives incorrect answers. As a result, many recent studies have turned their focus on utilizing the time of response along with the results of the answers within adaptive examination systems in order to improve the next question accuracy and fulfill a real-time adjustment during the test (1).

Among the important techniques that are used for forming smart systems is the (Fuzzy Logic). This importance has emerged from its ability to treat the ambiguous cases, uncertainty, transform numerical values into flexible linguistic ones such as (fast/slow, correct/ incorrect). This makes it appropriate for making decisions within the adaptive learning systems. To add more, it can also be employed for analyzing the time of response and the answer-correctness in real time, and then identify the question-suitable level of difficulty for the student during the exam. Recent studies have shown that the use of this technique, i.e. (Fuzzy Logic), helps in improving the efficiency of intelligent systems of evaluation and the systems of adaptation due to its flexibility in processing inaccurate data and making decisions in a way that is similar to that followed by human reasoning (3)

The algorithm of Extreme Learning Machine (ELM) is considered one of the recent learning algorithms that are distinguished by rapid training and low complexity related to computation as compared to traditional neural networks. This makes this algorithm appropriate for the systems that require immediate response and data real-time treatment or processing. To add more, the combination of (Fuzzy Logic) and this algorithm helps in building an intelligent system that is capable of improving the accuracy of student-level prediction and making decisions that are more efficient and adaptive than before.

What has been shown by recent studies is that the models of (Fuzzy ELM) result in superior findings concerning classification and prediction compared to the individual employment of traditional models (4, 5).

The existence of various studies that have tackled adaptive examinations and smart systems reflects the fact that most of these studies have relied upon complex variables like analyzing the whole behavior of the student, or biometric data, or interaction monitoring within the system, a matter that leads to complicating the system and the computational cost required. Some other studies have employed the techniques of machine learning, but they have not relied on flexible mechanisms for decision making. Other studies have utilized the (Fuzzy Logic) in isolation, i.e. without being integrated with quick algorithms of learning that are capable of real-time prediction. Accordingly, the need for developing a smart adaptive system of examination is still there. This system depends on two crucial indicators only: quick response and the accuracy of answer, in addition to making use of the efficiency of (Fuzzy Logic) and the algorithm of (ELM) concerning adaptive-decision making in real time during the examination (6, 7).

Figure (1) represent designing a structure for the electronic examination with a multi-hierarchical levels that are dependent on a sequential integration of (Fuzzy Logic). The present study also suggests designing a smart adaptive system of examination that is based on (Fuzzy Logic) and the algorithm of (ELM) in analyzing the performance of the student during the exam in real time. Analyzing the speed of response and the accuracy of the answer are performed in the first level. This is done via utilizing (Fuzzy Logic) and identifying the conditions required for assessing the status of the exam. The second level is responsible for estimating the actual level of the student via the use of the algorithm of (ELM). The final level is responsible for making the suitable adaptive decision that is related to changing the questions-difficulty level or choosing new questions that dynamically suit the level of the student during the exam. The proposed system aims at improving the evaluation accuracy, reducing the test time, and fulfilling a more actual measurement for the academic level of the student.

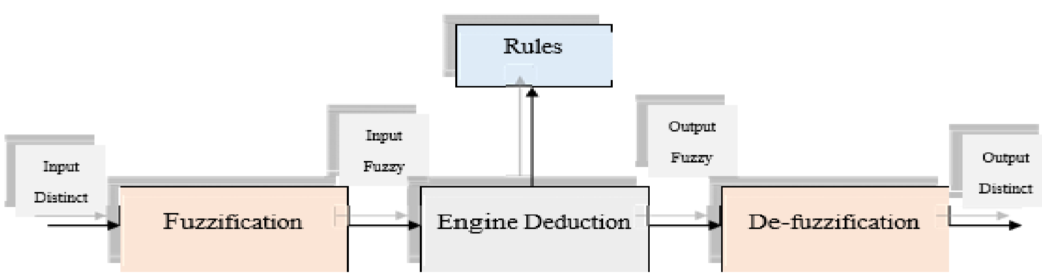

Fig. 1: Overview of the schematic representation of a fuzzy system.

2- Literature Review

This part tackles the effectiveness of the scoring model when evaluating and quantifying the assessments quality. This is performed via approaches such as computational linguistics, statistical analysis, and natural language processing (NLP) techniques [11,25]. An essential distinction among the various models of automated assessment is noticed within the divergence between the methodology of feature extraction and the actual mechanism of scoring. Reliance on surface linguistic features may be noticed the initial PEG model of scoring. It evaluated the texts of examination depending on certain metrics such as length as well as coherence. However, it failed in placing emphasis on the essential content of the tests. As a result, since this model, i.e. (PEG), showed a high accuracy of scoring, it clarified a robustness lack [12,26,27]. The examination scores of students show a fixed upward trend with longer test lengths, shorter structures of sentence, let alone simpler usage of vocabulary. To reinforce the identification of important attributes within the content of test, many specialists and researchers have integrated topic models into the feature extraction phase [13-15,28,29,30].

The main objective here is to fulfill a deep grasping of the organization and thematic composition that is related to the topics of examination. This is done via the techniques of topic modeling. The probabilistic latent semantic analysis represents the moist noticeable example of these models. Other examples are the processes of Dirichlet that shows that a core message of a paragraph is contingent on important lexical items within it. These pivotal concepts are identified by the topic model throughout computational analyses and utilizes the resulting vector in order to encapsulate certain topics. The advancements that are found in the technology of natural processing of language have fostered the relevant attributes extraction from the test texts of the students, including grammar facets and morphology, which strongly correlate to the content [16,31]. Regardless of this, the study of Hua emphasizes the importance of assessing the tests considering vocabulary as well as the structures of grammar employed. Refining the strategy of feature extraction requires directing the attention towards two crucial aspects: semantic coherence as well as consistency—traits that are particularly significant due to inherent limitations in semantic analysis. The findings support the relevance of these features and their contribution to reinforcing the effectiveness of model training [17,32].

An attribute of discourse analysis was introduced by Bingcai et al. The focus here was on the agreement degree with the argument found in the evaluation of automated testing. What is indicated by empirical research is that this feature increases the essay assessment accuracy. [18]. Fragluis has employed the fragmenting string kernel that functions with word embedding. His employment has been to generate the vectors of descriptive characteristics of textual data. The development of a scoring model, which is independent, is informed via these vectors. This model is for the assessment of the students, overcoming certain results that are fulfilled through the earlier methodologies of deep learning. Techniques of a multi-level feature extraction have been implemented by Xiaolei. This has been performed at text level and sentence level as well, and integrated within a framework of neural network in order to devise a notably precise automatic scoring model for student examinations [19].

A classification that is based on a hierarchical model of scoring was formulated by Lou. This classification classified texts depending on predefined characteristics and trains via the employment of different parameters and categories of paper [20,33]. Outputs were yielded from the results of this methodology. They were accompanied by an accuracy rate that reached 92%. To add more, in their study, Shim et al. used an unsupervised learning in order to assess the exam scores of the students. They emphasized the necessity of the segmentation of the training of end text into different types during the process of developing the model [21,34,35]. Reliance is made here on alignment, between test texts and the categorization objectives that are intended to show the effectiveness of this strategy, the facilitation of the weights-assignment to articles depending on their similarity to related documents. Remarkably, the unsupervised clustering technique performed a neighbor accuracy of 94% and an overall accuracy of 52% [22].

3- methodology

This part includes forming a system that has four structured integrated levels reflecting real-time assessment that is dependent on the speed and accuracy of the response. Figure (2) shows the flowchart of an electronic exam evaluation via the employment of machine learning. The subject is divided to multileveled to easily and analyses each one of them to achieve to the objects and as follows:-

Level 1: Preparatory Level: This level starts at the time of receiving the written answers. After that, the processing is performed as follows:

1-Language Processing: This stage includes:

First: Processing the texts:

These texts are processed as follows: in case that the question is essay-based, sentences are divided into words via the use of (Tokenization).

The role of the system is to standardize the words and show that the capital and small letters are the same as in (Database) and (database), both carry the same meaning via the use of the process called (normalization). Another crucial process is deleting (Function Words) since they supplement the words, though they have no meaning. This process is called (Stop Wording). What is produced via this process is a text that includes the principal concepts found in the student's answer.

Second: Extraction of Features:

For finding out the (Semantic Embedding), Bidirectional Encoder Representations From Transformers (BERT) technique has been utilized. It is extracted from the answers of students to be later transformed into numerical vectors that represent the contextual meaning of texts. Then, distances between words are then calculated via the employment of the equation known as (Euclidean Distance) between the vectors of student's answer and the ideal answers in order to measure the semantic similarity degree and determine how concepts are convergent (8, 3).

At the same time, student's behavior is monitored. It is represented by the time required for the answer of a question where creating a vector to represent the answer is required.

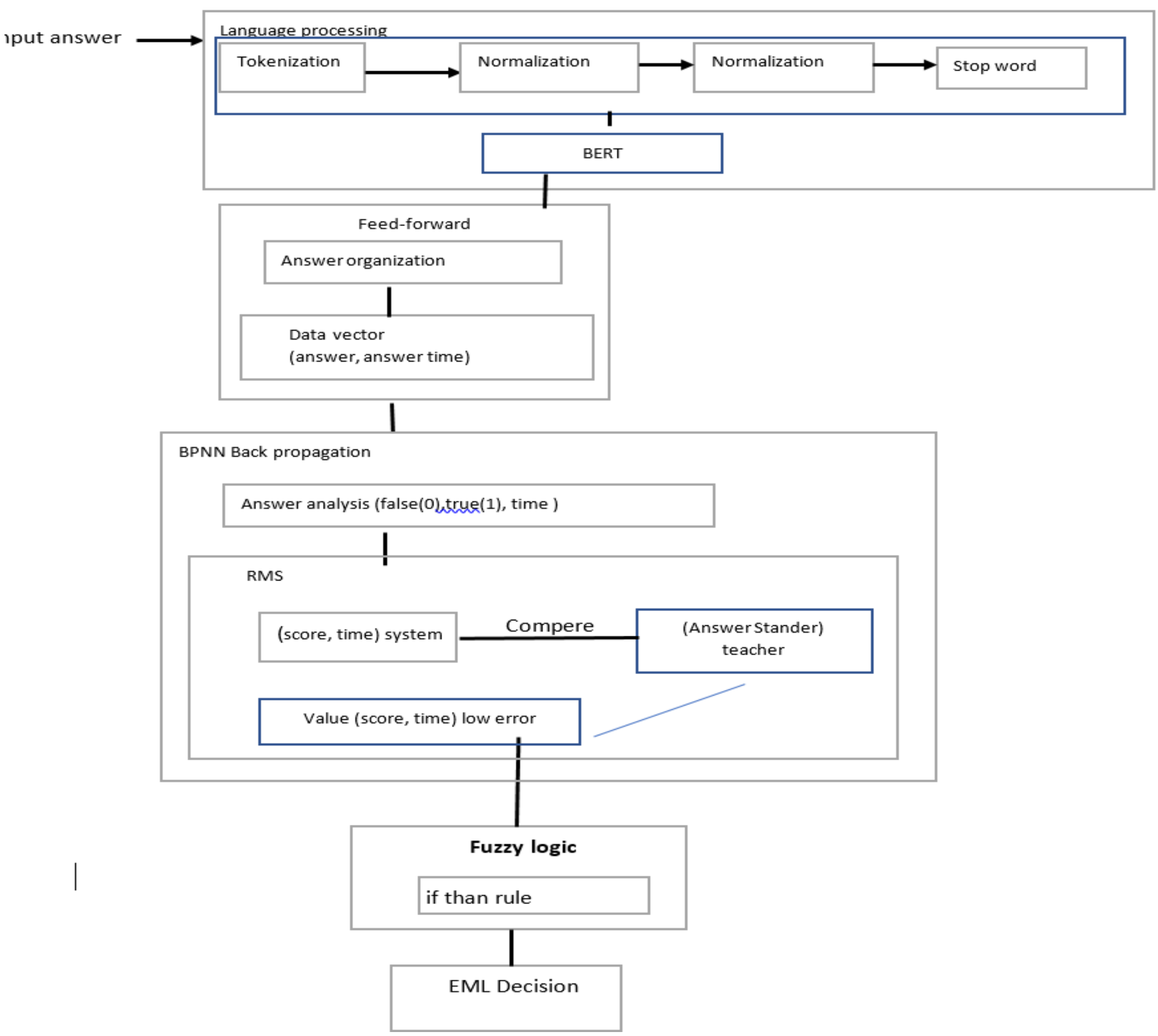

Fig.2 Flowchart ML TEST

Level Two: Feed-in Data Organization:

The role of this level is function as an intermediate stage for the standardization of the various data dimensions of the student. Here, the text vector is integrated with the data, time of response, and the index of question difficulty, let alone making sure of that all answers should lie within a numerical range (1, 0). The structure of the resulting matrix is as follows: [Answer Correctness, Response Time, Question Difficulty], preparing the way for equitable analysis in subsequent levels (9).

Level Three: Pattern Analysis and Prediction Accuracy Standards (BPNN) Back-propagation

Moving from data vector to understanding the behavior of the student is performed in this level. The (back-Propagation) network takes the numerical vectors which include [answer correctness, response time, question difficulty]. The basic role of the network is to analyze the interaction among the factors. Here appears the role of Root Mean Square Error (RMSE) to identify the accuracy of prediction via minimizing the difference between students' answers and the ideal answers. The result of this level is a numerical report that clarifies whether or not there is a deviant behavior for the student or it is a process of cheating or guessing (10).

Level Four: Fuzzy Logic: The Inference Engine - Implementing Instantaneous Interruption and Fuzzy Judgment

This level is the responsible for the transformation of numerical data into a logical educational decision since it does not tackle the uncertainty cases. The operation of this level is performed via a three-vector system, i.e. (answer correctness, time of response, and difficulty of the question), in addition to the accuracy of prediction. The inference engine works on incorporating the rules that govern this vector and make decisions related to keeping on the exam or stopping it depending on the inference rules: correct answer which are (standard time, very difficult question, and high RMSE). At this stage, a decision concerning the continuity or stoppage of the exam has to be taken concerning the specified student while keeping on the exam for the rest of students.

Level Five: EML Decision Making:

This level is a symbol for the engine of stability and adaptation. It provides the students with the following:

1-It ensures the continuity of the back-end monitoring of the students who continue performing the exam. It has no effect on the loading speed of questions or interface responsiveness, a matter that ensures smooth operation.

2-Dynamic standardization of the student: In this level, the student's linguistic voice in (BERT) is matched with his real-time performance in order to make sure that it is the same student who performs the answer.It also helps in reducing the cases of false alarms via addressing the cases of doubts. When the (Fuzzy Logic) level shows moderate suspicious cases that do not require warrant further investigation, the (EML) analyzes the answers and poses a question that ensures student's innocence.

The reason behind the employment of (Struts) is that related to its stability and reliability. The (Struts) is regarded as an operational stable framework, and its utilization in the structure of examination system is essential since it does not consume the system resources. To add more, there is a need for a high processing capability to treat the algorithms of intelligence, let alone it is merely a system container. It is to know that the real innovation is in the intelligent engine within the multiple level.

4- Results and Discussion

Table (1) and figure (3) shows that the matrix was constructed based on the structure of the proposed research model. The results were organized into five levels for the students, where all responses pass sequentially through the same stages.

These stages are as follows: the semantic linguistic level, the answer analysis and response time level, the correction level with comparison of responses to the model answer, the fuzzy analysis level responsible for determining the student's status and whether they may continue the examination, and finally the machine learning level, which is responsible for comparing suspicious responses with the (BERT) feature in the linguistic level in order to confirm the student's status and modify the state of the next question presented to the student.

Table1 student vector answer array

Student level	Vector array [BERT , true answer rate, answer time , RMSE]	Accuracy rater
Intelligent student	[0.96,093,60s,0.07]	97.32
Hardworking student	[0.88,0.82,67s,0.012]	95.27
Moderate student	[0.63,067,80s,0.026]	96.33
Less from moderate student	[0.63,0.67,180s,0.0033]	94.23
Weak student	[0.53,0.59,300s,0,041]	95.46
Average		95.72

The above matrix illustrates the classification of students into five main levels. The matrix indicates that an intelligent student provides answers with very few linguistic errors, accurate responses, and within a short period of time, while maintaining a very low error rate when compared with the model answers. The same principle applies to the remaining student levels, where the (BERT) score gradually decreases as the student's level declines. Response time is considered a highly important factor in determining the student's status when analyzed in conjunction with the other evaluation criteria. Based on these comparisons, the system determines whether the student is cheating during the examination and consequently terminates the exam, or alternatively sends signals to the other program levels to modify the type of questions presented in order to resolve issues related to electronic cheating.

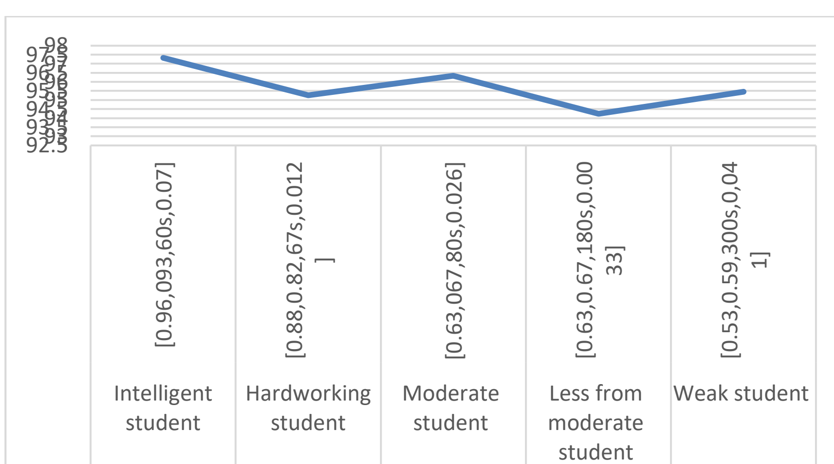

Fig. 3 Analyses Student level Vector array rater

Table (2) below represents the values of precision and recall, which were calculated according to Equations (1) and (2) for a total of (100) attempts.

$$P = \frac{TP}{TP + FP} \qquad (1)$$

$$R = \frac{TP}{TP + FN} \qquad (2)$$

Where:

TP (True Positive): Correct detection of cheating cases.

FP (False Positive): Cases suspected as cheating but were not actual cheating incidents.

FN (False Negative): Cheating cases that were not detected by the system.

TN (True Negative): Cases in which no cheating behavior occurred during the examination.

Table 2 evaluation result

MODLE	TP	FP	FN	TN	PRECISION	RECALL
BERT	70	14	6	10	0.83	0.87
BERT+ELM	77	12	3	15	0.86	0,96
BERT+ELM+FUZZY (Propos system)	84	7	2	7	0.92	0.97

The results presented in the table show a clear difference between the models used. The model that relied solely on linguistic fingerprints achieved moderate cheating detection rates, indicating that dependence on linguistic fingerprints alone is insufficient for accurately identifying student cheating.

When integrating the ELM model, an improvement in the results can be observed, with a higher rate of cheating detection. This demonstrates that the model contributed to enhancing the classification process, strengthening the relationship with linguistic fingerprints, and improving the understanding of students' answers.

The proposed model, which combined the aforementioned techniques with fuzzy logic, achieved superior results. It recorded the highest performance in detecting cheating cases, while also attaining high precision and recall rates compared with the other models. Furthermore, it reduced uncertainty in identifying student cheating by adapting the response level and improved the system's ability to handle uncertainty in students' answers.

References

[1] H-Ah Kang et al., "Location-Matching Adaptive Testing for Polytomous Technology-Enhanced Items", Applied Psychological Measurement, 2024. DOI: 10.1177/01466216241227548.

[2] H. Luo & Xiangdong Yang, "Efficiency of Computerized Adaptive Testing with a Cognitively Designed Item Bank", Frontiers in Psychology, 2024. DOI: 10.3389/fpsyg.2024.1353419.

[3] T. Glushkova et al., “Beyond Traditional Assessment: A Fuzzy Logic-Infused Hybrid Approach to Equitable Proficiency Evaluation via Online Practice Tests”, Mathematics, MDPI, 2024. DOI: 10.3390/math12030371

[4] Q. Gao et al., “Intuitionistic Fuzzy Extreme Learning Machine with the Truncated Pinball Loss”, Neural Processing Letters, Springer, 2024. DOI: 10.1007/s11063-024-11492-5.

[5] Qi. Chen et al., “A Medical Disease Assisted Diagnosis Method Based on Lightweight Fuzzy SZGWO-ELM Neural Network Model”, Scientific Reports, Nature, 2024. DOI: 10.1038/s41598-024-79426-8

[6] T. Glushkova et al., “Beyond Traditional Assessment: A Fuzzy Logic-Infused Hybrid Approach to Equitable Proficiency Evaluation via Online Practice Tests”, Mathematics, MDPI, 2024.

[7] Y.Yang, , H. Lou, Z.Wang, , et al. “Pinball-Huber boosted extreme learning machine regression: a multiobjective approach to accurate power load forecasting”. Appl Intell 54, 8745–8760 (2024). https://doi.org/10.1007/s10489-024-05651-3

[8]W.Tianhao.” An ELM-based approach to promoting reading of library books”. Int. J. Information and Communication Technology, Vol. 26, No. 2, pp.82–95. 2025

[9] M.Almsallti, , A. B .Alzubi,., & Adegboye, O. R.”Hybrid Metaheuristic Optimized Extreme Learning Machine for Sustainability Focused CO_2 Emission Prediction Using Globalization-Driven Indicators”. SUSTAINABILITY, 17(15), 6783. (2025) . https://doi.org/10.3390/su17156783

[10] T.Tara P A, L.F Simões, K.Hou Yip, Nikolaos Nikolaou, João M Mendonça, Ingo P Waldmann,”Extreme learning machines for exoplanet simulations: a faster, lightweight alternative to deep learning”, RAS techniques and instruments, Volume 4, 2025, rzaf050, https://doi.org/10.1093/rasti/rzaf050

[11] R.jwad kadhem, "semantic information retrieval for hadith based on query expansion of ontological knowledge" phd thesis . Universiti Sains Islam Malaysia, Malaysia, 2016

[12] M .Uto,”A review of deep-neural automated essay scoring models”. *Behaviormetrika* **48**, 459–484 .(2021). https://doi.org/10.1007/s41237-021-00142-y

[13] N. Khayi, V .Rus. “Automated Essay Scoring Using Discourse External Knowledge”. Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence.2024. https://doi.org/10.24963/ijcai.2024/791

[14] X.Wu,., T.Nguyen. & Luu, A.T. A survey on neural topic models: methods, applications, and challenges. *Artif Intell Rev* **57**, 18 (2024). https://doi.org/10.1007/s10462-023-10661-7

[15] D. Blei, A. Ng, and M. Jordan, "Latent Dirichlet Allocation," J. Mach. Learn. Res., vol. 3, pp. 993–1022, Jan. 2003.

[16] K. Taghipour and H. T. Ng, "A neural approach to automated essay scoring," in Proc. EMNLP, 2016, pp. 1882–1891.

[17] Z. Hua, "Semantic analysis for essay evaluation: A coherence-based approach," Journal of Language Technology, vol. 7, no. 3, pp. 45–52, 2019.

[18] Y. Bingcai et al., "Discourse analysis for automated essay scoring systems," in Proc. ACL, 2020, pp. 201–210.

[19] X. Xiaolei, "Deep learning for automatic essay scoring using hierarchical features," Neural Comput. Appl., vol. 32, pp. 14223–14233, 2020.

[20] Y. Lou, "Text classification-based essay scoring system using hierarchical techniques," Int. J. Artif. Intell., vol. 12, no. 2, pp. 88–96, 2021.

[21] Faseeh, M., Jaleel, A., Iqbal, N., Ghani, A., Abdusalomov, A., Mehmood, A., & Cho, Y.-I. “ Hybrid Approach to Automated Essay Scoring: Integrating Deep Learning Embeddings with Handcrafted Linguistic Features for Improved Accuracy”. *Mathematics*, *12*(21), 3416. 2024.https://doi.org/10.3390/math12213416

[22] A. Fragulis et al., "A string kernel-based model for automatic student evaluation," Expert Systems with Applications, vol. 169, p. 114341, Apr. 2021.

[23] N.Gardazi,, A.Daud, , Malik, M.K. et al. “BERT applications in natural language processing: a review”. Artif Intell Rev 58, 166 (2025). https://doi.org/10.1007/s10462-025-11162-5

[24] Li, C., Xie, Z., & Wang, H. “Short Text Classification Based on Enhanced Word Embedding and Hybrid Neural Networks”. *Applied Sciences*, *15*(9), 5102. (2025). https://doi.org/10.3390/app15095102

[25] O. Sythar..P,Mishra.A,Singhal.S.”exploring the landscape of natural language processing for text analytics:a comprehensive review”. Sixth International Conference on Futuristic Trends in Networks and Computing Technologies (FTNCT06) . 2025 held, India

[26] H.Misgna, , On, BW., Lee, I. et al. “A survey on deep learning-based automated essay scoring and feedback generation”. Artif Intell Rev 58, 36 .2025. https://doi.org/10.1007/s10462-024-11017-5

[27]Y.Huang, and Wilson, Joshua,. “Evaluating LLM-Based Automated Essay Scoring: Accuracy, Fairness, and Validity”. Conference: Artificial Intelligence in Measurement and Education Conference (AIME-Con) At: Pittsburgh, Pennsylvania, United States. 2025.

[28] J.Pecuchova, Ľ .Benko,. & Drlik, M. “Automated Grading of Open-Ended Questions in Higher Education Using GenAI Models”. Int J Artif Intell Educ 35, 3813–3846 (2025). https://doi.org/10.1007/s40593-025-00517-2

[29]C. Dennis , Gabon, Albert A. Vinluan, Jennifer T. Carpio.”Automated Grading of Essay Using Natural Language Processing: A Comparative Analysis with Human Raters Across Multiple Essay Types”. Journal of Information Systems Engineering and Management.(2025). https://doi.org/10.52783/jisem.v10i6s.700

[30] B.Quah, , L Zheng,., Sng, T.J.H. *et al.* Reliability of ChatGPT in automated essay scoring for dental undergraduate examinations. *BMC Med Educ* **24**, 962 (2024). https://doi.org/10.1186/s12909-024-05881-6

[31]M.Bexte, A.Horbach, & Zesch, T. “Strengths and weaknesses of automated scoring of free-text student answers”. Informatik Spektrum 47, 78–86 (2024). https://doi.org/10.1007/s00287-024-01573-z

[32] B. Miao, C.Xu, . “Aspect-level multimodal sentiment analysis model based on multi-scale feature extraction”. Sci Rep 15, 31591 (2025). https://doi.org/10.1038/s41598-025-16051-z

[33] H.Allam, L.Makubvure, ,B. Gyamfi, , Graham, K. N., & Akinwolere, K. “Text Classification: How Machine Learning Is Revolutionizing Text Categorization”. *Information*, *16*(2), 130. (2025).https://doi.org/10.3390/info16020130

[34] Mustafa, S., Hama Saeed, M. Empowering text classification with NLP and explainable AI for enhanced interpretability. Journal of Electrical Systems and Inf Technol 12, 81 (2025). https://doi.org/10.1186/s43067-025-00273-2

[35] Q.Liu, , K.Xiao, & Qian, Z. A hybrid re-fusion model for text classification. Sci Rep 15, 9333 (2025). https://doi.org/10.1038/s41598-025-90864-w