

ISSN 1840-4855
e-ISSN 2233-0046

Original scientific article
<http://dx.doi.org/10.70102/afts.2025.1833.271>

AUTOMATED ESSAY EVALUATION SYSTEMS FOR SCIENTIFIC WRITING IN ENGINEERING EDUCATION

Izzatbek Rejapov¹, Abdurahim Mannonov², Gavharxon Saydaliyeva³,
Gaybullaev Otabek⁴, Sherzod Djabbarov⁵, Dilnoza Ziyoyeva⁶, Ulugbek Eshkuvvatov⁷,
Nargiza Mannapova⁸

¹Deputy Dean, Faculty of Philology & Dean, Faculty for Foreign Students, Mamun university, Uzbekistan. e-mail: rejapov_izzatbek1@mamunedu.uz,
orcid: <https://orcid.org/0009-0000-0590-8710>

²Professor, Iran-Afghan Philology High School, Tashkent State University of Oriental Studies, Uzbekistan. e-mail: abdurahim.mannonov@mail.ru,
orcid: <https://orcid.org/0009-0009-3570-3455>

³Department of Foreign Languages, Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan. e-mail: gavharwlu@mail.ru,
orcid: <https://orcid.org/0000-0002-1078-7818>

⁴Professor, Vice-Rector for Youth Affairs, Samarkand State Institute of Foreign Languages, Samarkand, Uzbekistan. e-mail: ogaybullaev1979@gmail.com,
orcid: <https://orcid.org/0000-0003-3161-9212>

⁵Associate professor, Head, Department of English Teaching Methodology, Jizzakh state pedagogical university, Jizzakh, Uzbekistan. e-mail: sjabborov1980@gmail.com,
orcid: <https://orcid.org/0000-0002-1374-6843>

⁶Head, Department of English, Bukhara State Medical Institute named after Abu Ali ibn Sina, Bukhara, Uzbekistan. e-mail: ziyoyeva.dilnoza@bsmi.uz,
orcid: <https://orcid.org/0009-0004-4141-4805>

⁷Head, Department of Road Engineering, Termez State University of Engineering and Agrotechnology, Termez, Uzbekistan. e-mail: eshkuvvatov.ulugbek@mail.ru,
orcid: <https://orcid.org/0000-0002-6129-2583>

⁸Associate Professor, Uzbek National Pedagogical University named after Nizami Tashkent, Uzbekistan. e-mail: mannapova70@internet.ru,
orcid: <https://orcid.org/0009-0005-1878-0791>

Received: May 24, 2025; Revised: August 08, 2025; Accepted: September 11, 2025; Published: October 30, 2025

SUMMARY

Despite the importance of scientific writing for engineering students, the technical discourse poses a significant challenge for learners. The promise of Automated Essay Evaluation (AEE) Systems to provide timely and reliable feedback has enabled greater development of writing instruction through expedited and automated support. This paper examines the development of AEE systems designed specifically for scientific writing in the context of engineering education. We review existing research that applies Natural Language Processing (NLP), machine learning, and even rule-based linguistics to evaluate structural coherence, domain vocabulary, and argumentative discourse. Additionally, we evaluate the pedagogical advantages and drawbacks of AEE concerning fostering self-regulated learning, instructor support, and assessment rigor. Results from pilot case studies and experiments illustrate increased student-writing performance and engagement with the task. This research emphasizes the need for

domain-specific adaptation of AEE systems and algorithmic transparency, automated ethics, and machine-logic debates, challenging the predisposed notions of the use of AEE in engineering writing instruction.

Key words: *automated essay evaluation (aee), scientific writing, engineering education, natural language processing (nlp), writing assessment, machine learning, technical communication, feedback systems, academic writing, educational technology.*

INTRODUCTION

Automated Essay Evaluation Systems (AEE) are computer applications that focus on Natural Language Processing (NLP), Machine Learning (ML), and linguistic algorithms to assess essays and provide feedback [18]. The AEE systems aim to provide unbiased, prompt, and scalable evaluations of the essays which consider a wide range of features like grammar, vocabulary relevance, coherence, organization, and information content relevance [1]. To aid in educational settings, AEE systems attempt to emulate the scoring mechanisms of humans to enhance instructional assessment.

In the realm of engineering education, scientific writing is of utmost significance. In the case of broad engineering disciplines, students must proficiently elaborate on sophisticated topics within technical reports, research papers, documentation projects, and other professional communications. Acquisition of skills related to scientific writing not only assists the students in academics but also prepares them for career roles in industry or academia. However, their technical writing poses challenges due to the inadequately structured ideas concerning appropriate nomenclature which is terse, conventionalized and thus adheres to various regimens of technical workmanship [13]. Conventional approaches to teaching and evaluating writing that rely on extensive feedback from instructors tend to have limitations related to time and objectivity, especially in larger class settings.

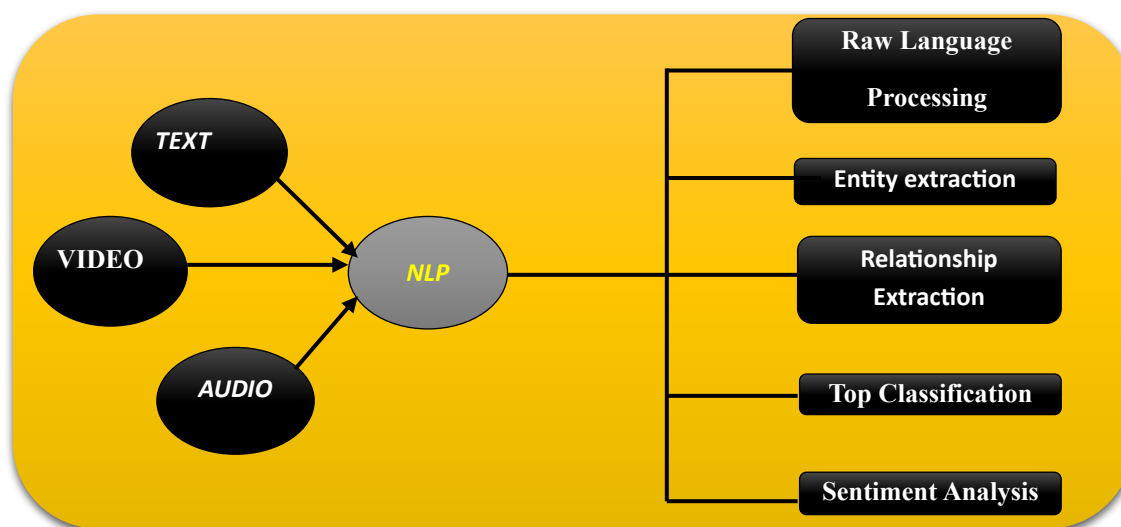


Figure 1. Natural language processing

Figure 1 describes the concept of Natural Language Processing (NLP) which is a branch of artificial intelligence (AI) and computational linguistics dealing with the manner in which computers interact with natural languages. Its main aim is to enable machines read, understand, interpret and generate human language in a meaningful and contextually correct way. NLP covers several tasks like text classification, sentiment analysis, machine translation, named entity recognition, syntactic parsing and semantic understanding among others. In educational environments more so in Automated Essay Evaluation (AEE) systems, NLP has been used in analysing student writing beyond grammatical errors and spelling. For instance nowadays modern NLP techniques use machine learning as well as deep learning algorithms to evaluate essay coherence, relevance, argument structure and content quality [2]. Through tokenization, lemmatization, part-of-speech tagging and using vector based language models such as word2vec, Glove, BERT, GPT etc, NLP allows systems to identify semantic relationships

between words occurring within students' writings. This enhances accurate scoring plus feedback generation that makes machine evaluations closer to human judgment. Therefore, NLEPs are instrumental in assessing technical reports project documentation as well as scientific explanations amidst the difficulty of domain specific language and highly structured writing. The changes in NLP technologies are result oriented in terms of personalized learning, formative assessment and the improvement of writing skills in various educational sectors.

Emerging technologies in AEE systems appear to address these issues effectively. Investigations have looked into the application of AEE technologies towards helping engineering learners enhance their writing skills by providing automated formative feedback and summative assessment [3],[14]. While AEE systems that focus on general tasks, such as e-rater and IntelliMetric, have received considerable attention, there is an increasing demand for specific tools that handle scientific and technical writing in a more context-sensitive manner. Tailoring the AEE systems to specific language and rhetorical patterns of a discipline has been shown to improve system accuracy and the learning outcomes of students significantly [5]. Even with all these advancements, the use of AEE tools in engineering education is still an emerging area of research that seeks to refine the quality, clarity, usefulness, and educational alignment of the provided feedback.

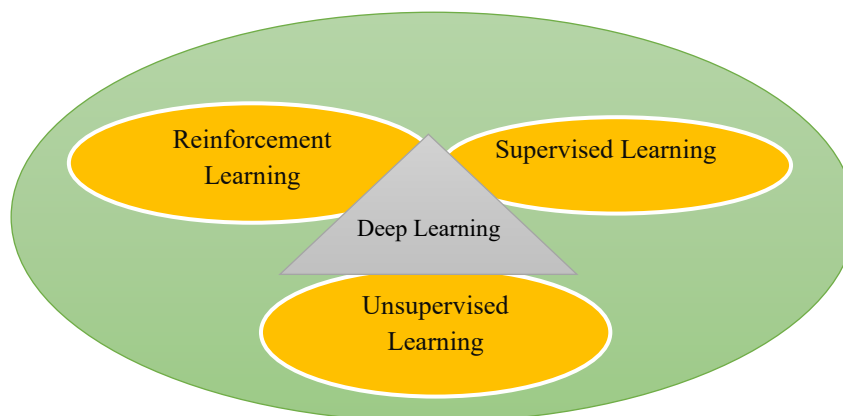


Figure 2. Machine learning

In Figure 2, it can be noted that machine learning (ML) is a subset of artificial intelligence which allows for a computer system to learn from data, improving systematically over time on a given task without the need of being programmed [6]. Essentially, ML algorithms detect patterns and make predictions or decisions based on input data. There is three main types of machine learning: supervised learning, where an algorithm is trained on data that is already labelled; unsupervised learning, which deals with unlabelled data and seeks to discover hidden patterns; and reinforcement learning, where learners discover the best possible actions to take through trial and error within an environment [10]. As far as the technologies we are considering here, and particularly in the case of Automated Essay Evaluation (AEE), ML models are trained using large corpora of human-scored essays to learn how to accurately assess new essays. These models apply numerous methods, including scoring essays using algorithms that analyse grammar, vocabulary, coherence, and argumentation, to produce scores comparable to human evaluators. More advanced ML methods, particularly those involving deep learning and natural language processing (NLP), significantly improved AEE system's abilities to capture or understand the semantic meaning, the context, and the quality of writing.

Thus, AEE systems that are powered by machine learning are being integrated into engineering education for automated assistance in marking examinations and supporting students with timely feedback to help improve their writing and communication skills in technical subjects.

The objective of this study is to describe AEE systems in the context of scientific writing for engineering students. It analyses the technology foundations of these systems, evaluates their impact in education, discusses the obstacles and prospects in engineering education curriculum design, and integrates these systems into teaching practices.

THEORETICAL FRAMEWORK

The foundations of Automated Essay Evaluation (AEE) systems are rooted in research on computational linguistics, natural language processing (NLP), machine learning (ML), and educational measurement [19]. In AEE systems, scoring is simulated by automating algorithms that extract and process linguistic, syntactic, and semantic features of text. These systems are largely based on two approaches: statistical machine learning models like the e-rater developed by ETS which relies on scoring patterns within extensive annotated datasets, and rule-based models which operate on pre-set linguistic and rhetorical rules [8], [15]. Such systems take into account theories about writing, genre theory and discourse analysis, as well as constructivist learning theory, all of which examine the role of structure, aim, and context in writing [7].

AEE systems in scientific writing context () offer both formative and summative feedback, assisting learners at surface level (grammar, spelling, and syntax) and higher level issues (coherence, argumentation, and domain-specific lexicon) [20]. By providing instantaneous feedback, these systems foster self-regulated learning, enabling learners to revise their work multiple times, which corresponds to Vygotsky's Zone of Proximal Development (ZPD) where optimal learning occurs with scaffolding support [16]. With regards to engineering students, who tend to have difficulties articulating complex technical concepts, AEE systems aid with structured feedback on clarity, logical progression, and scientific vocabulary [9].

Implementing AEE systems within engineering education has notable pedagogical advantages [4]. For example, writing practices are graded less, which frees up instructor time, while students have better access to objective and timely feedback [17]. Moreover, AEE tools may help in the development of uniform assessment criteria and provide insight into students' performance data over time. Despite the benefits, AEE systems encounter significant difficulties. One major challenge is the very limited ability to assess creatively, nuanced argumentation, and the contextual relevance of scientific content [21]. Additionally, automated feedback devoid of human intervention can result in superficial changes rather than meaningful learning. Deep learning may not occur when receiving automated feedback. Meeting the engineering discipline's specific criteria for AEE tools can present further challenges for general AEE systems [12].

In summary, AEE systems can revolutionize the teaching of scientific writing in engineering education, but such transformation hinges on adapting pedagogical frameworks, refining evaluative systems, and strategically balancing automated and manual feedback mechanisms [11].

METHODOLOGY

This investigation utilized the SLR methodology to assess the treatment and pedagogical impact of Automated Essay Evaluation (AEE) systems in relation to engineering education. The selection of an SLR was aimed at offering a summary that synthesized actual research which was in existence as AEE systems encompasses a wide range of technologies used in assigning essays and auto grading students with different techniques based on their education systems and settings. The design of the study was intended to comprehensively retrieve, analyse and evaluate published literature on the application of AEE technology in engineering education with regard to its impact on the level of educational assessment, teaching and student writing performance. In order to guarantee that the review could be repeated and was unbiased, the procedure complied with PRISMA guidelines for systematic reviews that outline unusual practices for harsh evaluation of each aspect of the investigation. The time span of literature was confined to focus on publications from 2013 and 2024 as there was increased development in the fields of Artificial Intelligence (AI), Natural Language Processing (NLP), and educational technology AEE systems were increasingly emerging during this period as useful instruments for writing assessment in STEM fields.

The selection criteria was undertaken in a step-by-step method that included identification, screening, eligibility evaluation, and final inclusion. Studies were obtained through thorough searches of various intermingling academic databases, such as IEEE Xplore, Scopus, Web of Science, ERIC, SpringerLink,

and Google Scholar. A series of keywords were combined with Boolean operators and more search phrases to ensure relevance and precision, including, “automated essay evaluation,” “engineering education,” AI based writing assessment,” and “machine learning for essay grading,” in addition to “automated scoring systems in technical writing.” The criteria for inclusion were tight to retain a predetermined level of academic rigor and relevance. The focus was limited to addressing the application of AEE systems for teaching automation in undergraduate and graduate engineering courses. Also, there was a requirement to document empirical data irrespective of the methodology used, provided it pertains to the system, instruction, or student evaluation outcomes, whether described in terms of system performance, instruction, or student evaluation. Furthermore, the studies had to be conducted in English and published in reliable peer reviewed journals, conference proceedings or institutional repositories. The exclusion criteria eliminated articles from outside the field of engineering, lacking methodological rigor, or discussing broadly automated grading without focusing on essay type writing tasks.

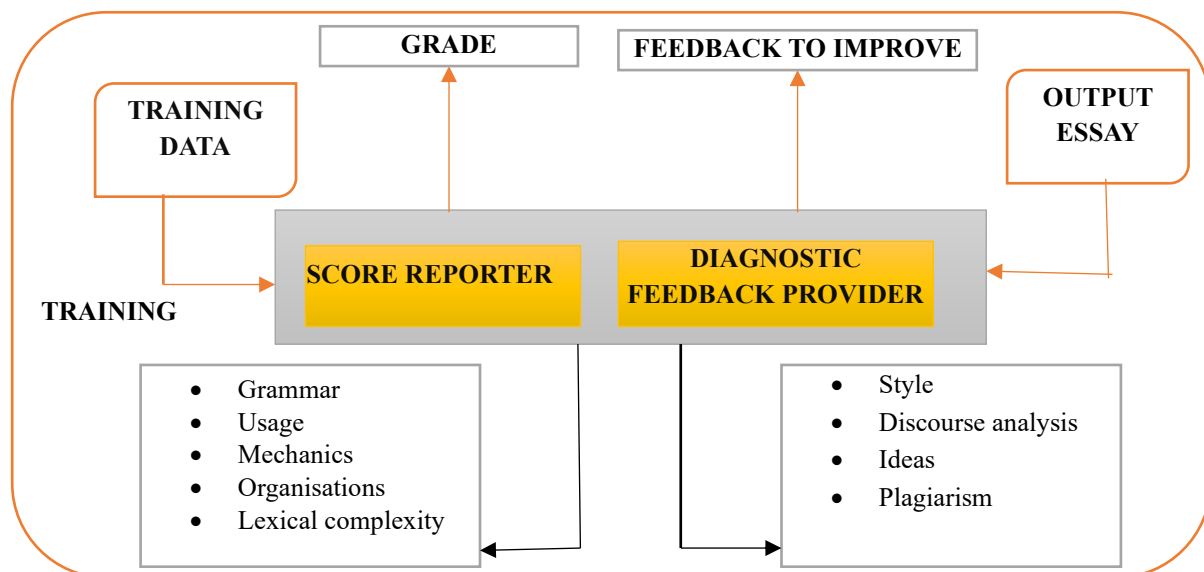


Figure 3. Common framework for AEE system

Automated Essays Evaluation (AEE) systems are examples of feedback-focused essay evaluation solutions spanning an entire processing pipeline from a modular, holistic perspective. Systems of this type are generally organized around four core activities: preprocessing, feature extraction, scoring model application, and feedback generation. The subprocess within the pre-processing phase entails cleaning the input essay into a more usable format through tasks such as tokenization, stop-word removal, lemmatization, and part-of-speech tagging. This task is referred to as feature extraction, and it entails finding relevant linguistic, syntactic, and semantic descriptors associated with the student's essay. More specifically, this can comprise surface level features (word count, sentence length), lexical diversity, and grammar, coherence and cohesion, domain-specific vocabulary (as shown in figure 3), and even essay specific parameters. Some more sophisticated systems perform feature extraction with the aid of NLP models such as BERT or GPT, which provide deeper contextual understanding via vectorized representations. During the scoring phase, the retrieved features are input into a trained ML model (regression model, SVM, or neural network) which has been trained on a dataset of human-scored essays. This process often results in aligned categorical or numerical scores to the essay's quality in accordance with a rubric. The final phase is feedback generation, during which formative comments concerning a writing sample's strengths and weaknesses are provided.

Certain systems aim at automating the process of suggesting corrections for grammar, coherence, and content issues. This approach makes it possible for AEE systems automated essay evaluation systems to be scalable, uniform, and to work with both summative assessment and formative instruction in teaching – in engineering education, for example, where clarity and precision in technical writing is crucial.

Once a final selection of relevant studies was found, a twofold form of analysis was performed in order to fully integrate the findings. Quantitative data was obtained to assess the technical aspects of system

performance such as scoring, accuracy, alignment with human raters, inter-rater reliability, precision and recall, and kappa coefficients. These metrics were evaluated in order to approximate the level of human mimicry performed by AEE systems in complex engineering writing assignments, which often involve domain knowledge and logic, along with the intricate reasoning and coherence expected at the graduate level. Concurrently, a qualitative content analysis was performed focusing on user experience and instructional elements, usability of the system, and their educational significance. Reports on student perceptions regarding feedback quality, system responsiveness to different writing styles, and instructor perceptions regarding the role of AEE in formative assessment were coded and thematically analysed. Special emphasis was placed on the difficulties dealing with the interpretation of technical language, counterargument assessment, and providing arguments which foster conceptual understanding while maintaining linguistic precision.

The merging of quantitative and qualitative data provided a deeper insight into the strengths and limitations of Automated Essay Evaluation AEE systems in engineering education. This analysis not only described the tangible efficacy and productivity of such systems, but also revealed the more subtle, sociological sides of assessment that writing AEE tools are not yet capable of addressing. This approach helped in formulating well-grounded, critically informed conclusions on the actual usefulness, incapacities, and emerging possibilities of AEE technologies in aiding engineering learners' academic writing and educators' assessment systems.

FINDINGS

The systematic review of the literature revealed several key findings regarding the use and effectiveness of Automated Essay Evaluation (AEE) systems in the engineering field education. In general, the studies suggested that AEE systems have received more attention as aids for the evaluation of technical writing, formative assessment, and essay enhancement in pedagogical settings. One of the key findings was that AEE systems, when properly implemented, are effective in alleviating instructors' grading burden without major discrepancies in scores compared to human raters. Moreover, several studies highlighted the usefulness of AEE tools in large class scenarios where individualized feedback is otherwise not possible within the time constraints. Yet, some studies observed that AEE systems often performed poorly with feedback, and domain-specific terminology as well as complex logical constructs and non-standard engineering report formats posed problems.

Differences in the performance, flexibility, and even teaching efficiency of various AEE systems used in engineering education was striking after conducting a comparative analysis. As an instance, rule-based systems like PEG (Project Essay Grade) are capable of evaluating a student's essay based on superficial grammar checks. However, they are unable to assess if the student's answer flows logically or whether there are reasonable arguments given from a technical standpoint. On the other hand, systems based on statistics and machine learning, such as e-rater® and Intelli Metric, proved to be more accurate in their compliance with human evaluators along with offering detailed comments. The latest AI models, particularly those employing deep learning and natural language processing (NLP), performed exceedingly well at understanding the context, assessing precision, and tailoring to different technical writing styles. The greatest drawback stemmed from the training, which had limited diversity for engineering specific content. Moreover, some systems were able to provide suggestions for interactive feedback and revisions which did foster higher levels of student engagement as well as enhancing longitudinal writing outcomes, while others remained unchanged during the entire feedback process.

The results are useful in the context of engineering education focused on scientific and technical writing. An example of an AEE system value is how it allows for self-paced independent learning and revision skills by providing immediate feedback. Instructors are now able to spend their time mentoring learners on other more sophisticated issues instead of routine error correction that deals with clarity and innovation on higher-level functions dealing with self-verbalization. The review does emphasize, however, certain limitations, like the need for human assessment, since machine systems lack the ability to interpret imagination, reasoning that justifies design, and the synthesis of advanced engineering design concepts. To that end, it is possible to assert that formative assessment with AEE can be efficiently administered, but careful integration is needed to ensure students know the bounds of feedback. Control

over grade-related decisions enables educators. It is observed that the use of AEE in engineering education relies on the use of technology concerning AEE accuracy, which is only one part of the training framework that is pedagogically structured as guidance for developing scientific writing skills central to the engineering profession.

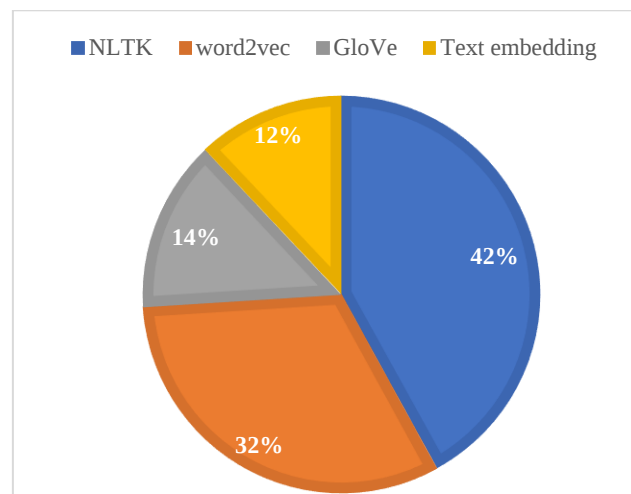


Figure 4. Systematic literature view of an automated essay evaluation system

The Automated Essay Evaluation systems (AEE) in engineering education context were organized and analysed systematically. The process is represented in the figure that follows the PRISMA methodology assessing AEE Essay Assessors by breaking them down into tools, criteria, and thematic areas of focus for identification, screening, eligibility, and inclusion phases of systematized literature review (SLR) engineering education review.

Focusing on the engineering education field, the figure illustrates the steps involved in performing a Systematic Literature Review (SLR) on AEE systems, capturing the distinct stages as aligned with the PRISMA framework.

CONCLUSION

This review has noted the increasing relevance and use of Automated Essay Evaluation (AEE) systems in engineering education, specifically pertaining to the evaluation of scientific and technical writing. The central conclusions of the review noted that AEE systems have merits with regard to their efficiency and effectiveness, reliability, and the speed with which they provide feedback, especially in large classrooms or online settings. These systems may facilitate the development of students' writing skills through multiple revisions while decreasing the grading workload for instructors. The findings also, however, highlight the fact that the efficacy of the AEE tools is dependent on the system's technology, its content knowledge processing ability, and its integration into the teaching framework. While systems utilizing machine learning and NLP have demonstrated greater flexibility and accuracy than rule-based systems, challenges remain with evaluating the engineering writing on the conceptual level, creativity, and technically sound reasoning.

Based on these findings, there are suggestions to address other areas of research. Primarily, there is a need in engineering AEE systems design and their subsequent assessments to consider the specialized language and unique features of engineering discourse. Later research should focus on how these instruments can adapt to elaborate formats like lab reports, design rationales, and project files. This study also needs to analyse the consequences of AEE-assisted instruction on enduring outcomes of writing skill, student attention, and independent thinking within engineering disciplines. In addition, some actionable recommendations may arise from comparative research assessing the use of AEE tools in blended versus project-based learning environments. Cross-cultural research may also be useful in looking at the relative effectiveness of AEE systems in different educational contexts and with different groups of students.

To summarize, although AEE systems are not ready to replace human evaluators, they offer remarkable possibilities as adjuncts in the instructional processes relating to scientific writing at the engineering education level. If well integrated, these systems can enhance learner independence, reduce the workload involved in evaluation, and serve as a single mechanism for sharpened attention to writing at different levels of the engineering curriculum. With continuing technological development, its educational use will hinge on the partnerships of teachers, builders, and scholars to make certain that AEE instruments are educationally sound and technically solid.

REFERENCE

- [1] Shermis MD, Burstein J. Handbook of automated essay evaluation. NY: Routledge. 2013.
- [2] Monir NI, Akter FY, Sayed SR. Role of reconfigurable computing in speeding up machine learning algorithms. SCCTS Transactions on Reconfigurable Computing, 2025;2(2):8–14. <https://doi.org/10.31838/RCC/02.02.2025;2>.
- [3] Attali Y, Burstein J. Automated essay scoring with e-rater® V. 2. The Journal of Technology, Learning and Assessment. 2006 Feb 1;4(3).
- [4] Acar S, Cevik E, Fesli E, Bozkurt RN, Kaufman JC. Testing the domain specificity of creativity with Kaufman domains of creativity scale: A meta-analytic confirmatory factor analysis. The Journal of Creative Behavior. 2024;58(1):171-189. <https://doi.org/10.1002/jocb.641>
- [5] McCorkindale W, Ghahramani R. Machine learning in chemical engineering for future trends and recent applications. Innovative Reviews in Engineering and Science. 2025;3(2):1-2. <https://doi.org/10.31838/INES/03.02.01>.
- [6] Hyland K. Disciplinary discourses: Social interactions. Ann Arbor: University of Michigan Press. 2004.
- [7] Siti A, Putri B. Enhancing performance of IoT sensor network on machine learning algorithms. Journal of Wireless Sensor Networks and IoT. 2025;2(1):13-9.
- [8] Warschauer M, Grimes D. Automated writing assessment in the classroom. Pedagogies: An International Journal. 2008 Jan 9;3(1):22-36. <https://doi.org/10.1080/15544800701771580>
- [9] Barhoumi EM, Charabi Y, Farhani S. Detailed guide to machine learning techniques in signal processing. Progress in Electronics and Communication Engineering. 2024;2(1):39-47.
- [10] Ramesh D, Sanampudi SK. An automated essay scoring systems: a systematic literature review. Artificial Intelligence Review. 2022 Mar;55(3):2495-527.
- [11] Zhang Z, Wang Y, Wang W. Machine Learning in Gel-Based Additive Manufacturing: From Material Design to Process Optimization. Gels. 2025 Aug;11(8):582. <https://doi.org/10.3390/gels11080582>
- [12] Connor U. Intercultural rhetoric research: Beyond texts. Journal of English for academic purposes. 2004 Oct 1;3(4):291-304. <https://doi.org/10.1016/j.jeap.2004.07.003>
- [13] Khan I, Siddiqui S. Machine Design a Systematic Approach to Designing Mechanical Systems. Association Journal of Interdisciplinary Technics in Engineering Mechanics. 2024 Sep 30;2(3):6-11.
- [14] Dikli S. An overview of automated scoring of essays. The Journal of Technology, Learning and Assessment. 2006 Aug 16;5(1).
- [15] Shermis MD, Burstein J. Handbook of automated essay evaluation. NY: Routledge. 2013.
- [16] Vygotsky LS. Mind in society: The development of higher psychological processes. Harvard university press; 1978.
- [17] Chinnasamy. (2024). A Blockchain and Machine Learning Integrated Hybrid System for Drug Supply Chain Management for the Smart Pharmaceutical Industry. Clinical Journal for Medicine, Health and Pharmacy, 2(2), 29-40.
- [18] Roscoe, R. D., Snow, E. L., & McNamara, D. S. (2014). Feedback and Revision: Using Natural Language Processing to Improve Writing. *International Journal of Artificial Intelligence in Education*, 24(3), 293–312.
- [19] Michael P, Jackson K. Advancing scientific discovery: A high performance computing architecture for AI and machine learning. Journal of Integrated VLSI, Embedded and Computing Technologies. 2025;2(2):18-26.