UDC 004.8:378.146

225. MODELS AND TOOLS FOR AUTOMATION OF QUALITY CONTROL IN EDUCATIONAL ASSIGNMENT ASSESSMENT

Pakutnik D.V., Master's degree student, gr. 526441

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Pristupa N.N. – PhD in Philology, Associate Professor

Annotation. This paper analyzes models and tools for automating quality control in educational assessment. It reviews rule-based systems, machine learning, and natural language processing approaches, and examines LMS-integrated solutions and their effectiveness. Key challenges include algorithmic bias, assessing creative work, and ensuring transparency. The study shows that hybrid approaches combining automated assessment with human oversight provide the most effective and reliable results for modern higher education systems.

Keywords. automated assessment, quality control, educational technology, machine learning, natural language processing.

The rapid expansion of higher education globally has created unprecedented challenges for educational institutions seeking to maintain assessment quality while managing increasing student numbers. Traditional manual grading approaches, while offering pedagogical benefits through personalized feedback, face significant scalability constraints that threaten the integrity of educational evaluation processes. According to recent research, automated grading systems can reduce assessment time by up to 80 % while maintaining accuracy levels comparable to human evaluators [1]. This transformation has prompted educators and administrators to explore automated solutions that preserve assessment quality while addressing operational efficiency concerns.

Quality control in educational assessment encompasses multiple dimensions including reliability, validity, fairness, and transparency. These dimensions ensure that evaluation outcomes accurately reflect student learning and maintain stakeholder confidence in educational credentials. Automated assessment systems must address each dimension while providing the scalability benefits that motivate their adoption. The integration of artificial intelligence and machine learning technologies has significantly expanded the capabilities of these systems, enabling assessment of increasingly complex assignment types beyond traditional multiple-choice questions.

This paper presents a comprehensive analysis of models and tools for automating quality control in educational assignment assessment. We examine the theoretical foundations of automated assessment, review current technological implementations, and identify emerging trends that will shape future developments in this field. The research contributes to ongoing discussions about educational technology adoption by providing evidence-based recommendations for institutions considering automated assessment solutions.

The development of automated assessment systems traces back to the 1960s with the introduction of optical mark recognition for standardized testing. These early systems provided efficient processing of multiple-choice assessments but offered limited pedagogical value. The subsequent decades witnessed gradual improvements in automated testing capabilities, with computer-assisted assessment gaining traction in the 1990s through Learning Management Systems (LMS) that enabled online quiz administration and automated scoring.

A comprehensive framework for evaluating the deployability of automated scoring systems is presented in [2], emphasizing the importance of validity evidence and fairness considerations. This work established evaluation criteria that remain relevant for contemporary AI-powered assessment tools. The implementation of automated assessment feedback and quality assurance systems for ICT courses, as documented in earlier research [3], demonstrated that automated approaches can provide more effective feedback mechanisms than manual alternatives under certain conditions.

The application of machine learning to educational assessment has transformed the capabilities of automated evaluation systems. Natural Language Processing (NLP) techniques enable automated essay scoring systems to evaluate written responses with accuracy approaching human rater performance. According to IEEE research, automated essay scoring systems have demonstrated the ability to provide consistent and objective assessments comparable to human evaluators [4]. These systems analyze textual features including grammar, vocabulary usage, argument structure, and semantic coherence to generate holistic evaluations.

Recent advances in transformer-based architectures have further improved automated essay scoring performance. Research published in arXiv demonstrated that efficient transformer-based language models achieve state-of-the-art results on standardized essay evaluation benchmarks [5]. The comprehensive review by Springer publications examining AI-powered grading and tailored feedback systems analyzed 77 core studies from 2018 to 2025, revealing significant improvements in assessment accuracy and feedback quality through machine learning integration [6].

Rule-based assessment models represent the foundational approach to automated evaluation, employing explicit criteria and decision trees to evaluate student submissions. These models excel in domains with well-defined correct answers, such as mathematical problem-solving, programming assignments, and objective knowledge assessments. The Automated Grading System with Student Performance Analytics developed by researchers demonstrates how rule-based approaches can effectively address challenges in traditional grading systems while providing detailed performance analytics [7].

The primary advantages of rule-based models include interpretability, consistency, and ease of maintenance. Educators can directly examine and modify assessment rules to align with pedagogical objectives. However, these models face limitations in evaluating open-ended responses and creative work that requires subjective judgment. Research indicates that purely rule-based approaches may inadequately capture the nuanced dimensions of higher-order thinking skills assessment [8].

Machine learning-based assessment models leverage data-driven approaches to learn evaluation patterns from training datasets. These models encompass various architectures including neural networks, support vector machines, random forests, and ensemble methods. The AI-based Automated Essay Grading System using NLP achieves grading times up to 80 % faster than human graders while maintaining comparable accuracy [9]. Such systems demonstrate the potential for significant efficiency gains without sacrificing assessment quality.

Deep learning approaches, particularly transformer-based architectures, have fundamentally transformed automated text assessment capabilities in recent years. Unlike earlier rule-based or statistical methods, these models are capable of capturing highly complex linguistic patterns, contextual dependencies, and deep semantic relationships within text. This allows them to evaluate not only surface-level features such as grammar and vocabulary, but also coherence, argumentation quality, and overall meaning, significantly improving the accuracy and flexibility of automated assessment systems.

However, despite these advantages, the inherently opaque, “black-box” nature of deep learning models introduces serious challenges. In educational contexts, transparency and interpretability are essential, as both instructors and students must clearly understand the rationale behind assigned grades or feedback. Without this clarity, trust in automated systems may be undermined, and their adoption in high-stakes assessment scenarios becomes problematic.

As a result, the development of explainable artificial intelligence techniques has emerged as a critical and rapidly evolving research direction. These approaches aim to make model decisions more interpretable by providing insights into which features, phrases, or structures influenced the evaluation. Advancing such methods is key to ensuring that automated assessment systems are not only accurate.

Hybrid assessment models combine automated preliminary evaluation with human oversight to leverage the strengths of both approaches. These systems typically employ automated tools for initial assessment and identification of responses requiring human attention, optimizing the allocation of instructor time. Research on human-in-the-loop Intelligent Visual Assessment systems demonstrates how hybrid approaches can enhance higher education quality assessment while maintaining pedagogical validity [10].

The implementation considerations for automated AI grading platforms in K-12 settings, explored through co-design pilots with teachers, reveal that successful hybrid systems require careful attention to educator workflows and pedagogical context [11]. These findings emphasize that technological capability alone does not ensure effective implementation; alignment with educational practices.

Learning Management Systems serve as the primary platform for deploying automated assessment tools in educational institutions. Research examining LMS utilization in higher education reveals that despite being standard components, these systems remain underutilized in universities [12]. The integration of automated assessment capabilities within LMS platforms offers significant advantages including centralized grade management, student progress tracking, and seamless workflow integration.

Quality assurance frameworks embedded within LMS platforms enable continuous monitoring of assessment effectiveness. Studies investigating automated assessment of Learning Management Systems in higher education institutions examine system features, compliance with regulatory standards, quality of service, and technology acceptance as key evaluation dimensions [13]. These frameworks support institutional efforts to maintain assessment quality while scaling educational delivery.

Natural Language Processing tools form the technological foundation for automated text assessment. These tools analyze linguistic features including syntax, semantics, discourse structure, and stylistic elements to evaluate written work. The comprehensive survey of automated essay scoring published in Neurocomputing provides detailed analysis of recent advances, datasets, and evaluation methodologies in this field [14]. Such resources inform the development of more sophisticated assessment tools.

The emergence of large language models has expanded the possibilities for automated assessment, enabling evaluation of increasingly sophisticated response types. Research exploring ChatGPT's potential for grading programming tasks demonstrates the capability of contemporary AI systems to understand and assess code quality [15]. However, these technologies also raise concerns about academic integrity and the authenticity of student work that institutions must address through policy and procedural frameworks.

Despite significant technological advances, automated assessment systems face substantial challenges that limit their applicability and effectiveness. Algorithmic bias represents a critical concern, as machine learning models trained on historical data may perpetuate existing inequalities in educational outcomes. Research on trustworthy AI in educational assessment emphasizes the importance of fairness, accountability, and transparency in developing assessment technologies [16]. Institutions must implement ongoing monitoring and adjustment processes to address bias in automated systems.

The assessment of creative and innovative work poses particular challenges for automated systems. Unlike objective assessments with clear correct answers, creative submissions require evaluation of originality, artistic merit, and subjective quality dimensions that resist algorithmic quantification. The promises and limitations of automated grading systems in computer science education, as examined through qualitative research with expert practitioners, reveal ongoing debates about appropriate scope and application of these technologies [17].

Pedagogical transparency constitutes another significant challenge. Students and educators need to understand assessment criteria and feedback mechanisms to support learning processes. Black-box AI systems that cannot explain their evaluation rationale undermine the formative potential of assessment. The development of explainable AI approaches specifically designed for educational contexts represents an important research direction addressing this limitation.

This paper has examined models and tools for automating quality control in educational assignment assessment, analyzing their theoretical foundations, technological implementations, and practical limitations. The research reveals that automated assessment systems offer significant potential for improving evaluation efficiency and consistency, particularly for large-enrollment courses and standardized assessment formats. Machine learning approaches, especially those incorporating natural language processing, have expanded automated assessment capabilities to include written responses and complex problem-solving tasks.

However, the findings also underscore the importance of addressing challenges related to algorithmic bias, assessment of creative work, and pedagogical transparency. Hybrid approaches combining automated preliminary assessment with human oversight emerge as the most promising direction for institutional implementation, offering efficiency gains while preserving pedagogical integrity. Future research should focus on developing explainable AI techniques for educational assessment, establishing robust frameworks for bias detection and mitigation, and designing hybrid assessment workflows that optimize the allocation of human and automated evaluation resources.

Educational institutions considering automated assessment adoption should prioritize alignment with pedagogical objectives, stakeholder engagement, and ongoing quality monitoring. The technology should serve educational goals rather than driving implementation decisions. As the field continues to evolve, collaboration between educational researchers, technologists, and practitioners will be essential for developing assessment systems that enhance learning outcomes while maintaining the integrity of educational evaluation.

References:

1. Automated grading system with student performance analytics [website]. – URL: https://www.researchgate.net/publication/392357301 (date of access: 17.03.2026).

2. Evaluations of Automated Scoring Systems in Practice [website]. – URL: https://onlinelibrary.wiley.com/doi/full/10.1002/ets2.12293 (date of access: 17.03.2026).

3. The Implementation of an Automated Assessment Feedback and Quality Assurance System for ICT Courses [website]. – URL: https://aisel.aisnet.org/jise/vol18/iss4/12 (date of access: 17.03.2026).

4. Automated Essay Scoring Using Natural Language [website]. – URL: https://ieeexplore.ieee.org/document/9310845 (date of access: 17.03.2026).

5. Automated essay scoring using efficient transformer-based language models [website]. – URL: https://arxiv.org/abs/2102.13136 (date of access: 17.03.2026).

6. A comprehensive review of AI-powered grading and tailored feedback systems [website]. – URL: https://link.springer.com/article/10.1007/s44163-025-00517-0 (date of access: 17.03.2026).

7. Analysis of an automatic grading system within first year Computer Science programming modules [website]. – URL: https://dl.acm.org/doi/10.1145/3437914.3437973 (date of access: 17.03.2026).

8. AI-based Automated Essay Grading System using NLP [website]. – URL: https://www.researchgate.net/publication/371429878 (date of access: 17.03.2026).

9. AI-assisted assessment of higher education quality: A visual analytical approach [website]. – URL: https://www.sciencedirect.com/science/article/pii/S2468502X26000021 (date of access: 17.03.2026).

10. Implementation Considerations for Automated AI Grading [website]. – URL: https://arxiv.org/html/2506.07955v1 (date of access: 17.03.2026).

11. An Assessment of Learning Management System Use in Higher Education: Perspectives from a Comprehensive Sample of Teachers and Students [website]. – URL: https://www.researchgate.net/publication/380633032 (date of access: 17.03.2026).

12. Towards Automated Assessment of Learning Management Systems in Higher Education Institutions in Zambia [website]. – URL: https://www.scirp.org/pdf/ojapps2024145_92312502.pdf (date of access: 17.03.2026).

13. A survey of automated essay scoring: Challenges, methods, and future directions [website]. – URL: https://www.sciencedirect.com/science/article/abs/pii/S0925231225015887 (date of access: 17.03.2026).

14. The future of grading programming assignments in education: The role of ChatGPT in automating the assessment and feedback process [website]. – URL: https://www.sciencedirect.com/science/article/pii/S1871187124000609 (date of access: 17.03.2026).

15. Trustworthy AI Meets Educational Assessment [website]. – URL: https://ojs.aaai.org/index.php/AAAI/article/download/35089/37244 (date of access: 17.03.2026).

16. Promises and breakages of automated grading systems: a qualitative study in computer science education [website]. – URL: https://www.tandfonline.com/doi/full/10.1080/20004508.2025.2464996 (date of access: 17.03.2026).

17. AI-Powered Learning: Revolutionizing Education and Assessment [website]. – URL: https://www.mdpi.com/2078-2489/16/11/1015 (date of access: 17.03.2026).