World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

A survey on automated student evaluation and analysis using machine learning

Chitoor Venkat Ajay Kumar, Adithya Krishna Eemani *, Gouri Charan Kalluri and Gunadeep Rudra

Department of CSE (Artificial Intelligence & Machine Learning), ACE Engineering College, Hyderabad, Telangana, India

World Journal of Advanced Research and Reviews, 2024, 21(02), 2547–2554

Publication history: Received on 19 February 2024; revised on 28 March 2024; accepted on 31 March 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.21.3.0901

Abstract

The problem of automated student evaluation and analysis is a widespread issue in the education sector, where the manual evaluation of students' performance and progress can be time-consuming, prone to errors, and lacking in efficiency. The current evaluation methods do not provide real-time insights and feedback to students, leading to a lack of engagement and motivation. Many Methodologies have been developed for estimating and analyzing the students' performance. One such methodology developed using: Machine learning, educational data mining, Predicting achievement. This methodology has no concern for the aspects like non-academic activities, personality traits of students and other technical and non-technical skills. The goal of automated student evaluation and analysis is to develop a system that can accurately and efficiently evaluate and analyze students' performance and provide meaningful feedback in real-time. The system should be able to analyze large amounts of data from various sources, exams, and provide detailed insights into the strengths and weaknesses of individual students. The system should also provide teachers with the tools to track student progress over time, identify areas for improvement, and provide personalized feedback to each student. The challenge of creating such a system lies in accurately analyzing and interpreting large amounts of data, identifying patterns and trends, and providing meaningful insights and feedback to students and teachers in real-time. It requires the integration of advanced artificial intelligence and machine learning techniques, as well as user-friendly interfaces and data visualization tools.

Keywords: Automated Student Evaluation; Performance Analysis; Real-Time Feedback; Machine Learning; Educational Data Mining; Predictive Achievement; Non-Academic Activities; Personality Traits; Technical Skills.

1. Introduction

Automated student evaluation and analysis pose a prevalent challenge in the education sector. Manual assessment of students' performance is time- consuming, error-prone, and inefficient. The existing evaluation methods fail to offer real-time insights and feedback, resulting in decreased student engagement and motivation. In the realm of education, the traditional methods of manually assessing and analyzing student performance have proven to be inefficient and inadequate. The emergence of advanced technologies, particularly in the fields of machine learning and educational data mining, has paved the way for the development of more sophisticated and dynamic systems. These systems are envisioned to revolutionize student evaluation, providing timely and accurate insights to both educators and students. The integration of artificial intelligence (AI) promises to address the shortcomings of conventional evaluation methods, offering a comprehensive approach that goes beyond academic achievements. A number of methodologies have been devised to estimate and analyze students' performance, and one notable approach utilizes machine learning and educational data mining to predict achievement.

* Corresponding author: Adithya krishna Eemani

This particular system focuses primarily on predicting scores and does not take into account other aspects of students' performance. It is important to note that this system may not encompass the full range of factors that contribute to a student's overall performance. It doesn't provide a proper system between mentor and mentee. The aim of automated student evaluation and analysis is to create a system that can effectively assess and analyze students' performance while delivering timely and valuable feedback. The System is capable of analyzing large amounts of data from various sources, exams, and provide detailed insights into the strengths and weaknesses of individual students. The system should also provide teachers with the tools to track student progress over time, identify areas for improvement, and provide personalized feedback to each student.

In expanding the vision of an automated student evaluation and analysis system, several pivotal considerations come to the forefront. Beyond the fundamental dimensions of academic assessment, the system must adopt a holistic approach, encompassing a spectrum of non-academic factors. These include extracurricular achievements, interpersonal skills, and emotional intelligence. Recognizing and evaluating these diverse aspects contributes to a more nuanced understanding of a student's overall development, fostering a comprehensive view that goes beyond traditional metrics. Ethical considerations stand as a cornerstone in the development of such systems. The integration of algorithms must be accompanied by meticulous efforts to mitigate biases and ensure fairness. Transparent and accountable AI models are imperative, not only for the sake of accurate evaluations but also to prevent inadvertent discrimination. Upholding ethical standards is crucial to maintaining the integrity of the automated evaluation process.

The success of the system is intrinsically tied to its accessibility and usability. Designing user-friendly interfaces tailored for both educators and students is essential. Intuitive platforms encourage regular interaction, promoting engagement and ensuring that the system becomes an integral part of the educational workflow. Seamlessly integrating the automated evaluation system with existing Learning Management Systems (LMS) further streamlines data flow, enhances interoperability, and contributes to a cohesive educational technology ecosystem. Emphasizing adaptability is vital for the sustained relevance of the system. The educational landscape is dynamic, and the system must evolve in tandem with changes in methodologies and curricula. Regular updates, feedback mechanisms, and a commitment to a continuous improvement cycle are essential components for ensuring the enduring effectiveness of the automated evaluation system.

Privacy and data security are paramount concerns in handling sensitive student information. Robust measures should be implemented to safeguard privacy and uphold data integrity. Adherence to data protection regulations not only ensures the confidentiality of student data but also fosters trust among both educators and students, a foundational element for the successful implementation of the system.

Involving stakeholders in the development process is indispensable. Collaboration with educators, students, and relevant stakeholders provides valuable insights and perspectives. Establishing feedback loops and embracing iterative design processes ensures that the system aligns with the practical needs and expectations of the education community. By incorporating these additional considerations, the automated student evaluation and analysis system can transcend its role as a mere assessment tool, becoming a comprehensive, adaptive solution that enriches the educational experience for all involved parties.

2. Literature review

"In recent years, the challenges associated with manual student evaluation and analysis within the education sector have spurred a growing interest in the development and implementation of automated systems. The time-consuming nature, error proneness, and inefficiency of traditional evaluation methods have prompted researchers and educators to explore innovative approaches. This literature review aims to explore the current landscape of automated student evaluation methodologies, focusing on advancements in machine learning, educational data mining, and predictive achievement models. By examining existing literature, we seek to identify gaps and opportunities for improvement, particularly in addressing non-academic factors such as personality traits and technical skills.

Hamzah Setiawan; Chastine Fatichah; Ahmad Saikhu et al.[1]

Evaluating student feedback is crucial for educational institutions to enhance their services. The primary objective is to refine the services provided to students through an information system that monitors their interests. The involvement of guardians in providing feedback on infrastructure and learning services is essential for continuous improvement. Furthermore, the evaluation process aims to address student queries promptly by eliciting responses from relevant departments, ensuring efficient delivery of student services. To expedite response times and enhance service quality, an automated system is employed for the classification of student feedback. Immediate categorization into relevant

departments is crucial, necessitating a system that prioritizes rapid development. Given that a single feedback item may pertain to multiple units, the challenge involves multi-label classification.This study focuses on the multi-label classification of student feedback data, utilizing Bidirectional Encoder Representation from Transformers (BERT) to extract word vectors. Various machine learning methods, including Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Random Forests (RF), and Decision Trees (DT), are employed for the classification of multi-label student feedback. The dataset comprises assessments from 3323 students regarding the guardianship information system. The experimental setup involves an 80% training data and 20% testing data split. Among the methods, the SVM approach with a linear kernel demonstrates superior performance, achieving an accuracy of 82% and an F1 value of 90%.

Ganga Sanuvala; Syeda Sameen Fatima et al.[2]

The written examination serves as a universally adopted method for assessing student performance in the education domain. It acts as a tool to ensure the consistency of the assessment process for both instructors and educational institutions. However, the manual effort required for assessment is significant and influenced by factors such as the teacher's knowledge, application-level understanding, marking criteria, and time constraints. Traditional evaluation methods are resource-intensive, demanding considerable time for the entire evaluation, verification, and result publication process. This study proposes the development and implementation of a Handwritten Answer Evaluation System (HAES) for assessing student exam papers. HAES is an automated response assessment system designed to recognize text in answer sheets and determine the grade for each answer based on the model's prior knowledge. The approach involves utilizing an Optical Character Recognition (OCR) tool to extract text from scanned handwritten answer scripts, while machine learning and natural language processing (NLP) techniques are applied for grading. The scoring system is based on cosine set similarity measures, assigning corresponding marks to each sentence in the evaluated answer paper. The developed model can be employed to evaluate scores for unscored descriptive answers

T M Sivanesan; Smitha J A; U. Arul; K. Balasubramanian et al.[3]

This research endeavor involves the development and implementation of an integrated system that utilizes IoT technologies and a machine learning algorithm for the measurement of Body Mass Index (BMI) and the conduct of health analytics within hospital settings. In order to effectively manage the widespread challenges of obesity, overweight, and chronic diseases, healthcare practitioners must obtain accurate and timely BMI readings in order to monitor weight status and address health concerns. The suggested solution uses Internet of Things (IoT) sensors to automate BMI calculations and collect real-time health data in order to overcome these difficulties. The hospital information systems and wireless IoT devices interact with ease. Data about vital signs, physical activity, and BMI are analyzed using the Support Vector Machine (SVM) method. Based on the gathered data, this program can efficiently identify and forecast BMI values. Moreover, SVM can forecast BMI, which allows for the creation of customized weight loss suggestions. By streamlining BMI readings in hospitals, the suggested solution seeks to reduce manual intervention and ease the workload for medical personnel. Early identification and tailored treatment of obesity-related health issues are made possible by real-time monitoring and analysis made possible by IoT devices and the SVM algorithm. In the end, data-driven decision-making is supported by the integration of IoT devices and the SVM algorithm, improving patient outcomes and overall healthcare efficiency.

Samuel Cunningham-Nelson; Mahsa Baktashmotlagh; Wageeh Boles et al.[4]

This study presents an automated approach for visualizing students' free-text comments obtained from course satisfaction surveys, with a specific focus on sentiment analysis. The goal of these visualizations is to identify areas of a course's teaching and learning components that are doing effectively or that may need improvement. The methodology gives teachers an easy-to-use and methodical way to keep an eye on their classes and decide on pedagogical tactics. Free-text comments from students are frequently included in course satisfaction surveys, which offer instructors insightful information. However, the abundance of data in these comments poses challenges for actionable insights. Existing visualization methods lack suitability for this context, necessitating additional capabilities. The research investigates the following queries: How can a lot of data about student happiness be compiled and shown graphically? In what ways can these visualizations offer meaningful insights into courses? What recurring themes emerge across different semesters? The study visualizes student satisfaction remarks using a variety of text analysis and machine learning techniques. Student opinions on a course are identified by the statistical technique known as latent Dirichlet allocation (LDA), while the sentiments conveyed in student comments are categorized through sentiment analysis. A case study with examples of these visualizations is then used to graphically present the findings to educators. The visualization techniques under investigation provide teachers with a summary of the course elements and the emotions

that go along with them. Simple comparisons between courses or teaching intervals within the same course are made easier by the condensed visuals.

Amit Rokade; Bhushan Patil; Sana Rajani; Surabhi Revandkar; Rajashree Shedge et al .[5]

Many articles discussing automated grading emphasize the significance of keyword matching in assessing answers. While this is crucial, humans often forget uncommon terms and substitute them with words of similar meaning. This paper proposes a solution for grading theory-based subject papers through Natural Language Processing (NLP) and adopts machine learning techniques, particularly Semantic Analysis. Recognizing that a single answer can be expressed in various ways by different students, relying solely on keyword matching is deemed inefficient. Therefore, the paper advocates the use of ontology to extract words and their synonyms relevant to the domain,

enabling a more comprehensive evaluation process. This method takes into account the coverage of concepts, the presence of keywords, synonyms, and appropriate word combinations. The suggested methods, which make use of Ontology, will be evaluated using standard input data that includes technical solutions. A comprehensive analysis of the data will be conducted with the goal of developing an automated grading system for theory-based subjects that is impartial, extremely accurate, and has a low error rate—one that is similar to the differential error rate amongst humans. The program was developed using teacher survey replies about what criteria they used to manually edit papers.

Binita Prasain; Simi Kamini Bajaj et al.[6]

The automated grading of student assignments and exam papers has been a subject of study for over a decade, offering a time-efficient method for evaluating academic assessments. While some educational institutions have successfully implemented this technology, others are still in search of optimal systems for accurate and efficient evaluation. This research aims to explore the functionality of automated marking systems and their potential to enhance the teaching and learning process. The paper provides a concise overview of current research in natural language processing and machine learning algorithms applied to develop automated grading systems. It introduces a sample framework utilizing data from diverse schools, universities, and datasets to assess student answers against a model answer. The system determines similarity levels between student and model answers, ultimately grading the student responses based on the degree of resemblance. This paper offers readers insights into recent developments in education, showcasing the integration of automated marking systems to evaluate student performance. Such technology not only aids teachers in saving time and effort but also ensures students receive timely feedback, contributing to their continuous improvement.

Marlo Haering; Muneera Bano; Didar Zowghi; Matthew Kearney; Walid Maalej et al.[7]

As the number of learning apps continues to grow, selecting a suitable app for courses has become a challenging task for teachers. Numerous evaluation frameworks have been proposed to aid teachers in this process. The iPAC framework, which emphasizes personalization, authenticity, and collaboration in the learners' experience, stands out as a well-established mobile learning framework. This article introduces an automated approach for identifying and comparing apps relevant to iPAC. Utilizing data from publicly available app descriptions and evaluations, the approach makes use of machine learning and natural language processing techniques The keyword base of the iPAC framework is empirically validated through the language used by app users in reviews. The results of the automated identification of iPAC relevant apps show promising outcomes (F1 score $\sim$ 72%), and their evaluation aligns closely with assessments by domain experts (Spearman's rank correlation 0.54). The discussion delves into the potential utility of these findings for teachers, students, and app vendors.

Mohammad Idhom; I Gusti Putu Asto Buditjahjanto; Munoto; Muchlas Samani et al.[8]

Traditionally, the assessment of essays has relied on manual scoring methods, presenting challenges in the learning evaluation process. To address this, Automated Essay Scoring (AES) systems have emerged as an alternative for the development of automated scoring applications. This research paper aims to assess the performance of an online AES system in evaluating vocational education competencies within the IT Multimedia field at the Community Academies in Blitar and Pacitan, East Java. The methodology involves text pre-processing, similarity calculation utilizing cosine similarity, and correlation analysis to gauge the AES system's performance against human raters. The study compares the assessment outcomes between the AES online system and human raters. Results indicate a correlation value of 0.903569186, signifying an excellent level of agreement, and a relatively small Mean Absolute Error (MAE) value of 0.133547009.

Arepelly Shylesh; Abdul Raafeh; Shaik Mathin; Vallakati Bhanu Prakash; Hariharan Shanmugasundaram et al.[9]

Recent advancements in deep learning have started to replace traditional systems in various aspects of our lives. Deep learning offers a cost-effective and consistent alternative to human marking, particularly in the context of script evaluation. The introduction of automated script analyzers addresses the challenges associated with manual assessment. The heightened competition in the education sector has placed a significant burden on evaluators, leading to time-consuming processes and variations in paper evaluation. This variability is not a reliable or efficient approach to assessment. To alleviate these issues and enhance the efficiency of evaluations, we propose the implementation of automated scripting for the assessment of scripts. This approach aims to reduce manual workload, expedite lengthy evaluation processes, and ensure a more evenly distributed assessment based on the quality of students' writing and the standard of their answers. The system takes into account various factors, such as grammatical errors, synonyms, and stop words, for comprehensive analysis and scoring. Many global examinations rely on descriptive written responses, which, while advantageous, come with inherent drawbacks. The meticulous evaluation of answer scripts demands careful consideration of each line, presenting a time-consuming challenge for evaluators. This not only delays the grading process but also leaves students awaiting their scores for an extended period. The proposed automated script evaluation system seeks to address these issues by saving time, reducing costs, and providing a more efficient means of evaluation.

Angad Sethi; Kavinder Singh et al.[10]

To attain reasonable performance, contemporary automated scoring models frequently rely on classical recurrent neural networks. However, their limited ability to codify long-term dependence limits their efficacy. In this study, a novel architecture including state-of-the-art language models from the natural language processing field is presented. By combining pre-trained language models with adapter modules with a bottleneck design, we are able to reduce trainable parameters without sacrificing performance. Furthermore, we present a strategy for identifying adversarial essays by reusing the bidirectional attention flow model. On the majority of the essay prompts in the Automated Student Assessment Prize dataset, our model outperforms earlier approaches used for this task, exhibiting state-of-the-art performance. The growing demand for automated essay scoring in large-scale standardized testing and model evaluation has driven the development of robust deep-learning algorithms. Chodorow & Burstein [1] addressed this need with their attempt to solve automated essay scoring, introducing the e-rater99 algorithm, which initially computed scores based solely on essay length. Subsequent advancements, such as the e-rater01 system, decreased dependence on essay length by taking closeness to the assigned topic into account. A significant breakthrough in automated essay scoring came with Klebanov et al.'s [2] introduction of argumentation as a crucial metric, minimizing dependence on hand-engineered features like grammatical errors and ambiguities.

Table 1 Comparison Between Automated Student Evaluation Techniques, Pros and Cons.

Paper	Year	Technique/Methodology	Pros	Cons
[1]	2023	Multilabel Classification of Student Feedback Data Using BERT and ML Methods	Enhances student services through automated feedback classification	Challenges with multi-label classification and the need for advanced ML methods
[2]	2021	A Study of Automated Evaluation of Student Examination Paper using ML Techniques	Automated grading, providing efficient assessment of exam papers.	Dependency on technology, potential challenges in implementing OCR accurately
[3]	2023	Automated Personalized Health Analytics using IoT and ML Algorithms	Integrates IoT and SVM algorithm for automated BMI measurement	Requires robust infrastructure, potential privacy concerns with health data.
[4]	2019	Visualizing Student Opinion with Text Analysis	Automated visualization of student feedback, aiding educators in course	Limited to sentiment analysis, may not capture nuanced feedback.

			improvement through sentiment analysis	
[5]	2018	Automated Grading System Using NLP	Utilizes NLP and Semantic Analysis for automated grading, considering synonyms	Relies on survey responses for algorithm design, potential bias in survey data.
[6]	2020	Analysis of Algorithm in Automated Marking in Education Hybrid Algorithm	Explores automated grading frameworks, providing insights into recent developments	Dependency on model answers, potential challenges in handling diverse answer patterns.
[7]	2021	Automating the Evaluation of Education Apps With App Store Data and ML	Automates app evaluation for iPAC framework using NLP and machine learning	Relies on publicly available data, potential limitations in language-based analysis
[8]	2022	Automated Essay Scoring Online System for Competency Assessment	Evaluates AES online system for vocational education, achieving a high correlation with human raters	Limited to online assessments, potential challenges in handling diverse essay styles..
[9]	2023	Automated Answer Script Evaluation Using Deep Learning(DL)	Proposes automated scripting for efficient script evaluation, considering grammatical errors	Challenges in handling descriptive answers, potential biases in scoring.
[10]	2022	NLP based Automated Essay Scoring with Parameter-Efficient Transformer Approach	Introduces an innovative architecture incorporating pre-trained language models	Challenges with traditional recurrent neural networks,

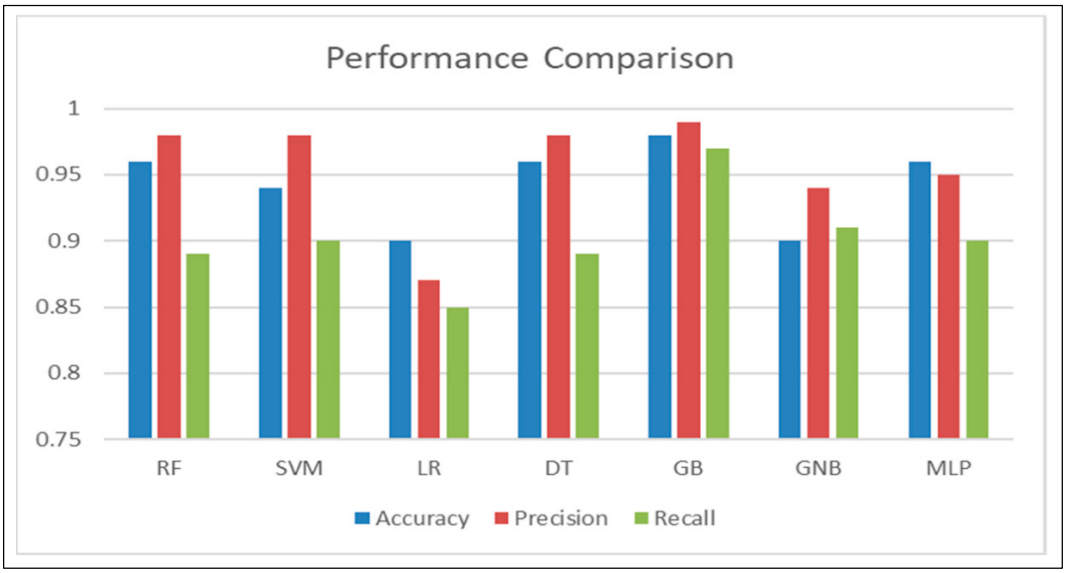

Figure 1 Comparison of Evaluating Student Knowledge Assessment Using Machine Learning Techniques

Figure 2 Error measure comparison of the classifiers

3. Conclusion

Creating an effective automated student evaluation and analysis system poses the challenge of accurately processing large volumes of data, recognizing patterns, and delivering meaningful insights and real-time feedback to both students and teachers. This necessitates the incorporation of advanced artificial intelligence and machine learning techniques, along with user-friendly interfaces and data visualization tools. In summary, the proposed system addresses critical drawbacks in current manual evaluation methods within the education sector. The existing approach, marked by its time-intensive nature, susceptibility to errors, and inefficiency, lacks the ability to offer real-time insights and comprehensive feedback to students. Furthermore, the limitations are underscored by the exclusive focus on predicting academic scores, neglecting crucial elements like non-academic activities, personality traits, and various technical and non-technical skills. The software requirements underscore the necessity for advanced algorithms, data mining tools, real-time feedback mechanisms, user-friendly interfaces, and structured communication platforms. On the hardware front, robust server infrastructure, reliable network components, accessible computing devices, and stringent security measures are deemed essential for successful implementation. In essence, the proposed system signifies a significant advancement in automating student evaluation and analysis, presenting a holistic solution that transcends traditional academic metrics. By leveraging advanced technologies and adopting a more inclusive evaluation approach, the system has the potential to enhance students' educational experience, engagement, and motivation, contributing to their comprehensive development.

Compliance with ethical standards

Acknowledgments

We would like to thank our guide Mr. C V AJAY KUMAR, Assistant Professor, Department of CSE (Artificial Intelligence & Machine Learning), Ace Engineering College. Also, we are extremely grateful to Dr. MRS. SOPPARI. KAVITHA, Head and Professor of Department of Computer Science (Artificial Intelligence and Machine Learning), Ace Engineering College for her support and invaluable time.

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Hamzah Setiawan; Chastine Fatichah; Ahmad Saikhu, Multilabel Classification of Student Feedback Data Using BERT and ML Methods

[2] Ganga Sanuvala; Syeda Sameen Fatima, A Study of Automated Evaluation of Student Examination Paper using ML Techniques

[3] T M Sivanesan; Smitha J A; U. Arul; K. Balasubramanian, Automated Personalized Health Analytics using IoT and ML Algorithms

[4] Samuel Cunningham-Nelson; Mahsa Baktashmotlagh; Wageeh Boles, Visualizing Student Opinion with Text Analysis

[5] Amit Rokade; Bhushan Patil; Sana Rajani; Surabhi Revandkar; Rajashree Shedge, Automated Grading System Using NLP

[6] Binita Prasain; Simi Kamini Bajaj, Analysis of Algorithm in Automated Marking in Education Hybrid Algorithm

[7] Marlo Haering; Muneera Bano; Didar Zowghi; Matthew Kearney; Walid Maalej, Automating the Evaluation of Education Apps With App Store Data and ML

[8] Mohammad Idhom; I Gusti Putu Asto Buditjahjanto; Munoto; Muchlas Samani, Automated Essay Scoring Online System for Competency Assessment

[9] Arepelly Shylesh; Abdul Raafeh; Shaik Mathin; Vallakati Bhanu Prakash; Hariharan Shanmugasundaram, Automated Answer Script Evaluation Using Deep Learning(DL)

[10] Angad Sethi; Kavinder Singh, NLP based Automated Essay Scoring with Parameter-Efficient Transformer Approach