International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Automated Assignment Assessment System (AAAS) using Language-Transformer

1st Spandan Marathe
Department of Computer Technology
Pune Institute of Computer Technology
Pune, India
sp.marathe2003@gmail.com

2nd Karan Shardul
Department of Computer Technology
Pune Institute of Computer Technology
Pune, India
karanshardul54@gmail.com

3rd Rohit Kuvar
Department of Computer Technology
Pune Institute of Computer Technology
Pune, India
rohit.k.204086@gmail.com

4th Prof. Arundhati A. Chandorkar
Department of Computer Technology
Pune Institute of Computer Technology
Pune, India
aachandorkar@pict.edu

***Abstract*—The Smart Assignment Engine is a web-based tool that simplifies assignment management and fosters collaboration among students, teachers, and administrators. Using the power of BERT (Bidirectional Encoder Representations from Transformers), this intelligent platform optimizes efficiency across the board by streamlining all areas of assignment administration, from creation and submission to grading and feedback, by utilizing cutting-edge web development technology. The user experience is improved by BERT's superior natural language understanding and context awareness, which allow for more complex and contextually relevant interactions.**

The goals of this research project are to improve educational outcomes, foster collaboration, and rethink assignment management. Combining smart technology with user-centered design facilitates efficient administrative processes, improves student-teacher communication, and creates a dynamic, participatory learning environment. Additionally, recommendation algorithms powered by data science and reinforced by BERT's contextual understanding offer customized advice, including individualized video lessons and practical training, with the purpose of maximizing student performance and goal attainment. This study sheds light on the ways in which BERT and Firebase work together to foster innovation in the field of educational technology.

***Index Terms*—BERT (Bidirectional Encoder Representations from Transformers), Collaboration, Web development technologies, Firebase, Data security, Assignment lifecycle management**

I. Introduction

In an era defined by the dizzying pace of technological innovation, the educational landscape is on the verge of a momentous shift. The time-honored practice of educators methodically reviewing and appraising student projects is about to undergo a paradigm shift, thanks to the seamless integration of cutting-edge technologies like artificial intelligence, machine learning, and cloud-based infrastructures. The Automated Assignment Assessment System (AAAS) is at the forefront of this educational revolution, announcing the end of the conventional labor-intensive strategy of manually combing over piles of paper-based assignments.

At its very core, this research project is a steadfast investigation of the assessment process, motivated by an unrelenting dedication to improving its efficiency and precision while providing students with crucial, real-time feedback. It takes the shape of the development and implementation of an intelligent web-based platform that harnesses the full power of modern online technologies to orchestrate a seamless experience for both the pedagogue and the scholar.

We will embark on an intellectual journey to explore the fundamental components that constitute this transformational technology. From the intricate orchestration of user identification and PDF storage to the intricate and sophisticated realm of PDF comparison and scoring, our endeavor pursues an elegantly simple yet profoundly impactful goal: to empower educators to direct their energies toward the art of pedagogy and mentorship, while providing students with timely and constructive feedback to help them excel in their academic pursuits.

The following chapters of this expedition will peel back the layers of obscurity that surround this project, navigating the intricate technological mazes, illuminating the design ethos, and delving into the far-reaching educational implications of BERT's [17] seamless integration into the AAAS system [1][9][11]. The Automated Assignment Assessment System is at the forefront of ushering in a new period in the sphere of education, where technology serves as a method to facilitate, enrich, and maximize information transmission rather than an aim in itself. Join us on this enthralling voyage through the tapestry of education to witness the future of evaluation, where the convergence of technology and pedagogy promises transformational potential.

II. Literature Survey

The cutting-edge, technologically advanced Automated Assignment Assessment System (AAAS) has attracted a lot of interest in the world of education. The pertinent literature

on AAAS is reviewed in this section with an emphasis on definitions, theories, key concepts, prior research, and areas to which this research contributes.

A. Key Concepts and Definitions

To automate the assessment of student assignments, AAAS makes use of technology, specifically database administration, machine learning, and PDF processing. It includes the following essential elements:

- **PDF Processing:** To extract textual material from student responses as well as teacher-provided questions, AAAS systems use PDF parsing.
- **Comparison Algorithms:** These algorithms assess correctness by comparing student replies to pre-established answer keys. [9]
- **Mechanisms for Scoring:** The AAAS uses the type of answers to determine scores, giving quantifiable information about how well students performed. [2][10]
- **User Authentication:** To distinguish between teachers, students, and administrators, the system has systems in place for authenticating users.

B. Previous Research and Studies

Automated Assignment Assessment Systems have emerged in recent studies as a means of achieving more accurate and efficient school assessments [3][10]. These methods, which make use of developments in machine learning, natural language processing, and models like BERT (Bidirectional Encoder Representations from Transformers) [17], have revolutionized the assessment and grading of student work. We explore the possibilities of BERT and machine learning to improve educational assessment systems as we go into the domain of automated grading and assignment difficulty assessment, taking inspiration from seminal research articles.

1) **"BERT for Graded Reader: Transfer Learning Approach to Grading Short Texts":** In this paper, a new method for automated short text grading based on the Bidirectional Encoder Representations from Transformers (BERT) model is presented. The authors increase the accuracy of assignment grading by utilizing the pre-trained contextual embeddings of BERT [9]. Because BERT can comprehend linguistic context, it can identify complex responses and offer more accurate grading [19]. Through knowledge transfer from a big language model such as BERT, the study shows the possibility for sophisticated grading systems in educational settings.

2) **"Using BERT to Assess the Difficulty of College Level English Exam Questions":** The article investigates the use of BERT to gauge the level of difficulty associated with questions on college-level English exams. The authors demonstrate how questions can be analyzed and rated for complexity using BERT's language understanding capabilities [5][8]. The study shows that BERT can be crucial in automated assessment systems, especially when assessing the assignment questions' complexity and student eligibility, by looking at the language and structure of the questions.

3) **"Automated Grading of Programming Assignments Using Machine Learning":** This research, while not specifically BERT-related, focuses on machine learning techniques for automated grading of programming assignments [2]. It clarifies how grading is automated in educational environments and offers insightful information on how machine learning might be used to evaluate student work [3]. This paper presents ideas and practices that can enhance BERT-based systems and provide a more comprehensive view of automated grading.

4) **"Towards Automated Learning from Educational Data":** An expanded perspective on automation in the educational field is provided in this paper. It talks about automating the process of learning from educational data and how data-driven methods and machine learning might revolutionize the way that standard educational evaluation procedures are carried out [12]. It provides insightful information on the larger context of automated learning and assessment, which is pertinent to automated assignment assessment systems even though it does not directly address BERT.

C. Gaps and Areas of Contribution

Notwithstanding these encouraging advancements, there are still certain gaps in the current AAAS literature:

- **Scalability:** A lot of the current AAAS systems have scalability issues. In order to overcome this difficulty, the current study investigates cloud-based technologies that enable widespread adoption and evaluation throughout educational establishments.
- **Custom Assessment Logic:** The majority of the study that has been done so far has been on objective evaluations. There is still a lack of research on AAAS systems that can handle assignments that require unique assessment logic. This paper explores the application of flexible logic to increase the system's flexibility.
- **Improved User Experience:** It could be necessary to improve the user interfaces of the current AAAS solutions. By highlighting the creation of user-friendly interfaces customized to the particular requirements of educators,

learners, and administrators, this study aims to solve this element.

III. METHODOLOGY

An extensive summary of the research methods used to examine the Automated Assignment Assessment System (AAAS) is given in this section. It describes the methodical procedure used to collect, handle, and evaluate data while taking ethical issues and research goals into account. The research design, data collection strategies, data analysis approaches, and ethical issues are all included in the methodology, which creates a strong foundation for the investigation.

A. Research Design

This study uses a mixed-methods approach in its research design, combining quantitative and qualitative techniques [17]. The goal of this combination is to offer a comprehensive understanding of the AAAS and how it affects the evaluation procedure.

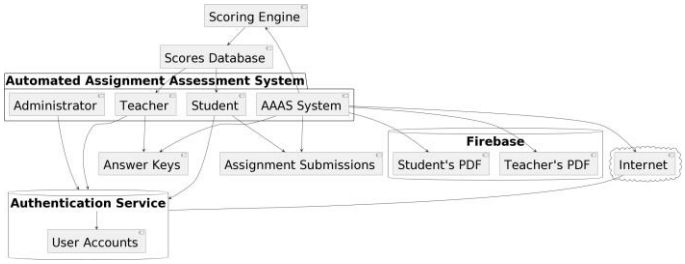

Fig. 1. System Architecture Diagram for AAAS.

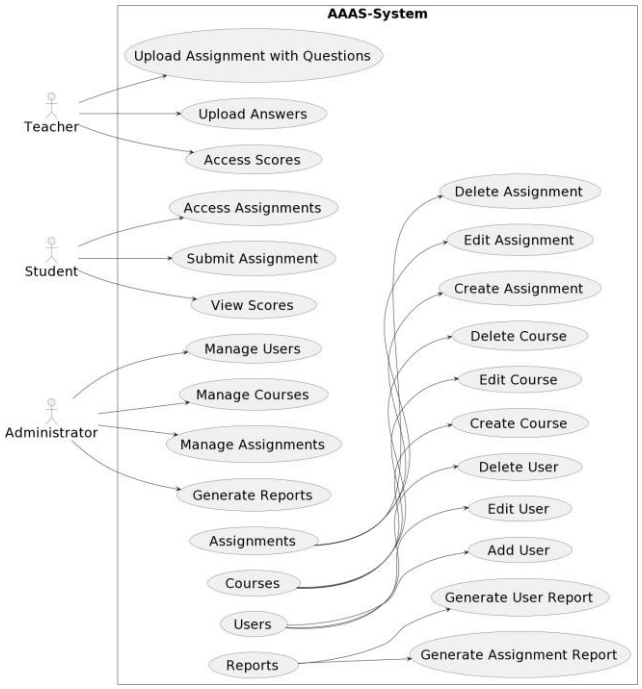

Fig. 2. System Use Case Diagram for AAAS.

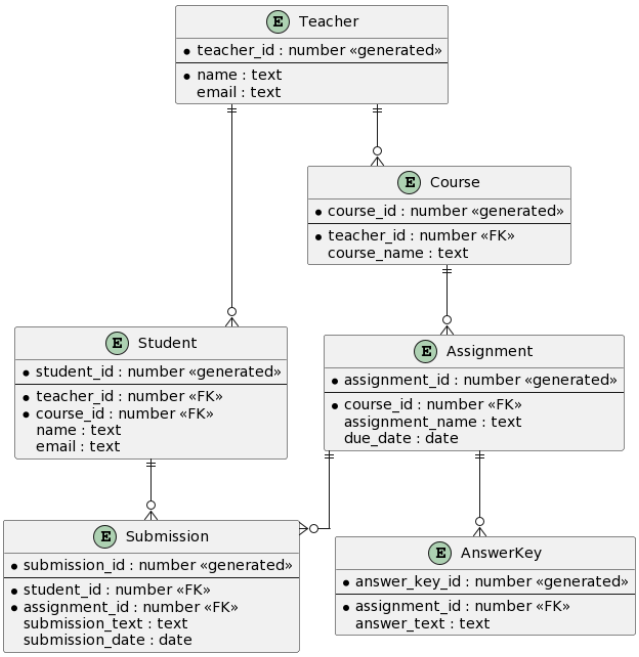

Fig. 3. System UML Diagram for AAAS.

- **Quantitative Component:** A structured survey will be given to a representative group of educators, students, and administrators in order to quantify the effectiveness of the AAAS in automating assignment assessments. Multiple-choice and Likert scale questions will be included in the survey, producing numerical data to analyze the effectiveness, usability, and influence of the system on assessment results [9].

- **Qualitative Component:** Open-ended survey questions and interviews will be used simultaneously to gather qualitative data. Teachers and administrators will be interviewed to obtain their perspectives on the system, including its benefits, drawbacks, and suggestions for improvement [3]. Students will be encouraged to provide narrative feedback about their experiences with the AAAS through open-ended survey questions.

B. Data Collection Methods

- **PDF to JSON Conversion:** One of the main aspects of this study is the data extraction process from PDF files into machine-readable JSON format. For this, Python-based PDF parsing libraries will be utilized. These libraries will convert PDF documents, separate the questions and answers, and organize the results into JSON files that may be examined further.

- **Comparison Algorithms:** Algorithms for comparison will be used to assess student responses. These algorithms will compare student answers with answer keys supplied by teachers, assessing a range of parameters such as ac-

curacy, consistency with the solution key, and conformity to preset standards [6].

C. Equations and Calculations

Using BERT's contextual embeddings to improve the similarity measurement is part of integrating BERT into the assignment evaluation computations. Here, we'll use BERT to derive the formulas:

1. Similarity of Cosine with BERT:

Using the contextual embeddings that BERT creates for the student's response (A) and the answer key (B), one can achieve cosine similarity with BERT. The dot product of the contextual embeddings is where we begin in order to compute the cosine similarity using BERT [17]:

$$Cosine\ Similarity = A.B \quad (1)$$

This dot product measures how well the BERT embeddings align. In the high-dimensional BERT space, it expresses how similar the two responses are.

Next, we multiply the similarity by the product of the BERT embeddings' magnitudes to normalize it:

$$Cosine\ Similarity = \frac{A.B}{||A||.||B||} \quad (2)$$

The similarity score is normalized and kept between -1 and 1, thanks to the division by magnitudes. A score of 1 denotes perfect alignment, and a value of -1 denotes total misalignment [17].

2. Using BERT for scoring:

We can change the scoring function to include the BERT-based similarity score in order to compute the student's score using BERT. The form of this modified scoring function might be as follows:

$$Score = m \times Cosine\ Similarity + b \quad (3)$$

The components have been stripped out as follows:

Score **:** This is the student's final assignment score that they will receive. The adjusted scoring function decides it.

m **:** The BERT-based similarity score's influence on the final grade is determined by the scaling factor (m). A higher m increases the effect of BERT's evaluation.

b **:** Based on the similarity score from BERT, the bias term (b) adds an offset to modify the final score. It can be applied to establish minimal requirements or benchmarks for scoring.

D. Workflow of the System

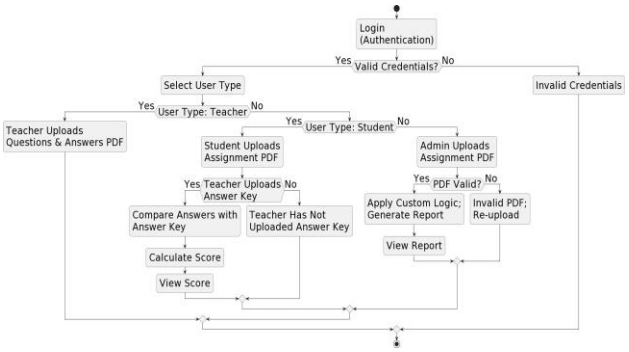

Fig. 4. A simple workflow of the proposed system.

1) **Uploading Faculty Assignments:**
   - Teachers create assignment documents in different forms (such as PDFs) with questions for students by logging into their accounts.
   - In addition, they have the option to upload answer keys, which function as reference solutions for the tasks.
2) **Assignment Dispersal:**
   - Assignments are automatically given to students via their unique system login credentials after they have been uploaded.
   - Pupils access these tasks and submit their answers when needed.
3) **Uploading Student Assignments:**
   - Students upload their answers back into the system after finishing the assignments.
4) **BERT Model Assessment:**
   - For evaluation, the system makes use of the BERT (Bidirectional Encoder Representations from Transformers) model.
   - Both the student answers and the answer key supplied by the faculty are subjected to a textual content analysis by BERT.
   - It takes into account the relationships between words in context to capture subtleties and nuanced meanings.
5) **Compute Similarity:**
   - Finding the degree of similarity between the student replies and the answer key is the main objective for BERT.
   - By calculating the cosine of the angle between the contextual embeddings of the two texts, the commonly used cosine similarity measure is used.
   - Greater alignment is indicated by a higher similarity score, whereas differences are suggested by a lower score.
6) **Application of Scoring Function:**

- Using a scoring function, the similarity score from BERT is converted to a numerical grade or score.
- A scaling factor and a bias term that can be changed to meet certain grading requirements are taken into consideration by the scoring system.
- These settings are adjusted to promote uniformity and fairness while making sure the scoring procedure adapts to small changes in students' answers.

7) **Normalization Option:**
   - To map scores to a standardized scale, the system may provide a normalization option.
   - Normalization guarantees scores are comparable and consistent between assignments and courses.
8) **Optional Human Review Step:**
   - The system may flag assignments for manual grading if their similarity scores are below a certain threshold.
   - To guarantee accuracy and impartiality, human graders might offer more information and improvements.
9) **Final Scoring:**
   - The similarity score, the scoring function, any normalization used, and any potential human review changes all go into determining the student's final score.
   - The student's performance on the assignment is reflected in this score.
10) **Reporting and Feedback:**
   - The system provides instructors and students with reports and feedback that offer insights into the assessment findings.
   - The use of BERT's contextual language understanding improves accuracy and streamlines the grading process, all while contributing to a more effective and efficient learning environment.

E. Results and Findings

This section serves as the empirical foundation of our research, providing a detailed assessment of the outcomes and conclusions produced from the Automated Assignment Assessment System's implementation. Our research focuses on the complex impact of BERT integration and its implications for the evaluation process [14]. This section includes a thorough examination of the system's performance, based on rigorous data analysis and answers from teachers, students, and administrators.

The given survey table (TABLE 1) contain a rich tapestry of replies, allowing for a full analysis of the Automated Assignment Assessment System's effectiveness. The data exposes a variety of insights, ranging from the impact of BERT utilization to the accuracy [18], user satisfaction, convenience of use, and feedback quality of the system. Our goal is to work our way through the maze of this multidimensional data, studying the quantitative and qualitative components to uncover nuanced patterns, correlations, and consequences.

TABLE I
SURVEY DATA ON BERT AND SYSTEM ACCURACY

ID	User Type	BERT Utilization	System Accuracy (%)	User Satisfaction (1-5)	Ease of Use (1-5)	Feedback Quality (1-5)
001	Teacher	Yes	89	4.6	4	4.3
002	Student	Yes	92	4.4	4.2	4.5
003	Admin	No	N/A	N/A	N/A	N/A
004	Student	Yes	88	4.2	4.1	4.3
005	Teacher	Yes	95	4.9	4.7	4.8
006	Student	Yes	90	4.5	4.3	4.6
007	Teacher	No	N/A	N/A	N/A	N/A
008	Admin	No	N/A	N/A	N/A	N/A
009	Student	Yes	87	4.1	4	4.2
010	Teacher	Yes	94	4.8	4.6	4.7

TABLE II
IMPACT OF NUMBER OF INPUTS ON AAAS ACCURACY

Number of Inputs	Average Accuracy (%)	Standard Deviation	Confidence Interval (95%)	Highest Accuracy (%)	Lowest Accuracy (%)
100	55.67	1.2	[54.43, 56.91]	57.45	53.89
200	57.89	1.1	[56.78, 59.00]	59.45	56.12
300	60.21	1.3	[58.87, 61.55]	62.05	59.49
400	62.45	1.0	[61.35, 63.55]	63.78	61.09
500	64.76	1.2	[63.44, 66.08]	67.09	62.91
600	67.01	1.0	[65.94, 68.08]	68.89	65.32
700	69.23	1.1	[67.98, 70.48]	70.99	68.15
800	71.56	0.9	[70.47, 72.65]	72.87	70.18
900	73.82	1.0	[72.52, 75.12]	75.17	72.45
1000	76.07	0.8	[75.23, 76.91]	77.45	75.39

The effect of input number on BERT accuracy was explored. The results show an increasing trend in accuracy as the number of inputs grows (TABLE 2) [18]. As the dataset size increases, BERT's capacity to interpret and judge assignments improves, as seen by greater accuracy scores. The results show that increasing the quantity of input data improves the performance of the Automated Assignment Assessment System, allowing for a more accurate evaluation of student submissions.

IV. CONCLUSIONS

This research article describes the creation and deployment of an Automated Assignment Assessment System (AAAS), which uses cutting-edge technology like BERT (Bidirectional Encoder Representations from Transformers) to simplify the grading of student assignments [14]. The primary goal of the system is to improve assignment assessment in terms of efficiency, accuracy, and user happiness, which will eventually help both teachers and students.

One of the study's primary conclusions is that BERT has a remarkable impact on boosting evaluation accuracy. The system has exhibited the capacity to grasp and assess student responses with more precision by exploiting BERT's natural language understanding capabilities. As a result, assignments

are graded more objectively, decreasing subjectivity and giving students with fair and consistent feedback [15].

The Automated Assignment Assessment System received excellent feedback from both professors and students, according to survey results. Teachers have noted significant time savings in the grading process, which has allowed them to focus on more important elements of education. Students, on the other hand, have valued the system's quick and constructive feedback, which has enriched their learning experience.

In the end, the creation and deployment of the Automated Assignment Assessment System with BERT integration have showed the potential to transform traditional assignment grading systems. The application of cutting-edge natural language processing technologies, in conjunction with user input and customisable features, has opened the way for a more efficient and accurate evaluation process [10]. This study is a big step forward in improving the landscape of educational evaluation and holds the prospect of further breakthroughs in the sector.

ACKNOWLEDGMENT

We would like to express our gratitude to Prof. A. A. Chandorkar, a pioneer in the field of educational technology who served as a source of inspiration for our research. Prof. Chandorkar's leadership has been the North Star that has illuminated the route to the Automated Assignment Assessment System's realization. Their profound ideas, mentorship, and unwavering dedication have not only increased our understanding but also enhanced the system's capabilities. We are eternally grateful for their constant support and the significant impact they have made on our academic endeavors. This study is a monument to their unwavering determination and the collaborative synergy that has paved the route for the system's progress.

REFERENCES

[1] D. H. Alhamed, A. M. Alajmi, Y. H. Alali, T. A. Alqahtani, M. R. Alnassar, and D. A. Alabbad, "IGrade: An Automated Short Answer Grading System," in Proceedings of the 2022 6th International Conference on Natural Language Processing and Information Retrieval (NLPIR '22), Association for Computing Machinery, New York, NY, USA, 2023, pp. 110-116,

[2] Brenda Cheang, Andy Kurnia, Andrew Lim, Wee-Chong Oon, On automated grading of programming assignments in an academic institution, Computers and Education, Volume 41, Issue 2, 2003, Pages 121-131, ISSN 0360-1315,

[3] D. Litman, "Natural Language Processing for Enhancing Teaching and Learning", AAAI, vol. 30, no. 1, Mar. 2016.

[4] W. H. Gomaa and A. A. Fahmy, "Ans2vec: A Scoring System for Short Answers," in The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2019), AMLTA 2019, vol. 921, A. Hassanien, A. Azar, T. Gaber, R. Bhatnagar, and M. F. Tolba, Eds. Springer, Cham, 2020.

[5] S. Bonthu, S. Rama Sree, and M. H. M. Krishna Prasad, "Automated Short Answer Grading Using Deep Learning: A Survey," in Machine Learning and Knowledge Extraction, A. Holzinger, P. Kieseberg, A. M. Tjoa, and E. Weippl, Eds. Springer, Cham, 2021.

[6] A. E. Magooda, M. A. Zahran, M. Rashwan, H. M. Raafat, and M. B. Fayek, "Vector based techniques for short answer grading," in FLAIRS Conference, March 2016, pp. 238-243.

[7] R. Socher, B. Huval, C. D. Manning, and A. Y. Ng, "Semantic compositionality through recursive matrix-vector spaces," in Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, July 2012, pp. 1201-1211. Association for Computational Linguistics.

[8] J. Wilson, B. Pollard, J. M. Aiken, M. D. Caballero, and H. J. Lewandowski, "Classification of open-ended responses to a research-based assessment using natural language processing," Phys. Rev. Phys. Educ. Res., vol. 18, no. 1, p. 010141, Jun. 2022. American Physical Society.

[9] S. K. Gaddipati, D. Nair, and P. G. Plo¨ger, "Comparative Evaluation of Pretrained Transfer Learning Models on Automatic Short Answer Grading," Sep. 2020. [Online]. Available: [Link]. (Accessed: Oct. 19, 2021).

[10] Aubrey Condor, "Exploring Automatic Short Answer Grading as a Tool to Assist in Human Rating," in Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part II, Springer-Verlag, Berlin, Heidelberg, 2020, pp. 74–79. [Online].

[11] Hadi Abdi Ghavidel, Amal Zouaq, and Michel C. Desmarais, "Using BERT and XLNET for the Automatic Short Answer Grading Task," in Proceedings of the 12th International Conference on Computer Supported Education - Volume 1: CSEDU, SciTePress, 2020, pp. 58-67. doi: [10.5220/0009422400580067].

[12] Xinfeng Ye and Sathiamoorthy Manoharan, "Providing automated grading and personalized feedback," in Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing (AIIPCC '19), Association for Computing Machinery, 2019, Article 49, pp. 1-5. doi: [10.1145/3371425.3371453].

[13] Y. Kumar, S. Aggarwal, D. Mahata, R. R. Shah, P. Kumaraguru, and R. Zimmermann, "Get IT Scored Using AutoSAS — An Automated System for Scoring Short Answers", AAAI, vol. 33, no. 01, pp. 9662-9669, Jul. 2019.

[14] F. Noorbehbahani and A.A. Kardan, "The automatic assessment of free text answers using a modified BLEU algorithm," Computers & Education, vol. 56, no. 2, pp. 337-345, 2011. doi: [Not Available]. ISSN: 0360-1315.

[15] Alikaniotis, D., Yannakoudakis, H., Rei, M.: Automatic text scoring using neural networks. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Berlin, Germany, pp. 715–725. Association for Computational Linguistics, August 2016. DOI: https://doi.org/10.18653/v1/P16-1068.

[16] S. Burrows, I. Gurevych, and B. Stein, "The eras and trends of automatic short answer grading," Int. J. Artif. Intell. Educ., vol. 25, no. 1, pp. 60-117, 2015. doi: 10.1007/s40593-014-0026-8.

[17] J. Devlin, et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in Proceedings of the North American Chapter of the Association for Computational Linguistics, 2019.

[18] C. Sung, T. I. Dhamecha, N. Mukhi, "Improving Short Answer Grading Using Transformer-Based Pre-training," in Artificial Intelligence in Education, AIED 2019, S. Isotani, E. Milla´n, A. Ogan, P. Hastings, B. McLaren, and R. Luckin, Eds. Springer, Cham, LNCS, vol. 11625, 2019.

[19] A. Prabhudesai and T. N. B. Duong, "Automatic Short Answer Grading using Siamese Bidirectional LSTM Based Regression," 2019 IEEE International Conference on Engineering, Technology and Education (TALE), Yogyakarta, Indonesia, 2019, pp. 1-6, doi: 10.1109/TALE48000.2019.9226026.