IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 10, May 2025
Impact Factor: 7.67

Subjective Answer Evaluation Using NLP

Prof. V. N. Sawant[1], Uday Dube[2], Mayur Dhumal[3], Mayank Hangshoo[4]
Professor, Department of Computer Engineering[1]
Students, Department of Computer Engineering[2-4]
NBN Sinhgad Technical Institute Campus, Pune, India

Abstract: *Subjective answer evaluation system aim to address the complexity of evaluation of written content. We propose innovative ways to use machine learning and language technology to automate this evaluation process. Use tools such as BERT and Cosine similarity to assess the quality of your responses. Our approach combines problem concepts with core concept solutions to assess student answers and allows machine learning models to predict reviews based on content and response quality. The first results show that similarity between the BERT- based embodiment and the cosine achieves a high level of accuracy when evaluating the answer, maintaining consistency and fairness. The project not only demonstrates the potential of electronic systems in educational assessments, but also aims to improve the overall effectiveness and integrity of context assessments*.

Keywords: Subjective Answer Evaluation, NLP, BERT, Cosine Similarity

I. INTRODUCTION

The project aims to create a system that allows students to evaluate their responses more efficiently and equitably. With the help of advanced machine learning and word processing technology, the system analyzes detailed responses from students and allows them to express their own opinions. The content and quality of these responses are evaluated using tools such as Bert and Cosine similarity. This approach helps reduce contradictions and bias in classification, provides better evaluation criteria and respects individual student feedback. It was developed to adapt to a variety of topics and questions and to increase system flexibility. It also analyzes key terms and definitions so that all answers deal with essential elements. The program aims to save teachers time and effort by automating the assessment process. This development benefits both teachers and students by allowing teachers to concentrate more on the lessons. Ultimately, this system leads to a more equitable and efficient review. The system also promotes personalized learning by providing students with wise feedback based on their responses. Determining gaps in understanding can lead to students improving and deeper learning. The use of BERT ensures context-conscious evaluation and captures not only the keywords but also the intent and structure of the answer. It supports multilingual skills and can be adjusted to suit a variety of educational environments.

II. METHODOLOGY

Various methods have been proposed to effectively evaluate students' subjective answers and improve classification efficiency. Among them has shown the use of Bert and Cosine similarity to promise results by analyzing the semantic meaning of answers instead of relying on only key words. This ensures a more accurate and fair evaluation by understanding the context and sentence structure. In contrast to traditional keyword-based or statistical models, Bert records deeper relationships in languages and allows them to better understand different answers from students. This method also reduces distortion and increases consistency when assessing. Real-time feedback and adaptability can further improve your approach through a variety of topics and question types. It also supports scalability and is suitable for large education platforms. Integrating these techniques can help you create detailed reports for each student. Teachers can use this finding to personalize their educational strategies. The system also encourages students to concentrate on conceptual understanding and not get caught up in memory. Over time, it contributes to the establishment of a more intelligent and responsive academic evaluation system.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 10, May 2025
Impact Factor: 7.67

III. REVIEW OF LITERATURE

The field of subjective answer evaluation using natural language (NLP) processing records considerable advances provided by the need to automate and improve the efficiency of education systems. Early approaches to assessing ratings were primarily based on keyword matching and simple textual comparison methods, with limited ability to record the semantic importance of student responses. These traditional methods allowed us to evaluate responses based on surface-level characteristics such as word frequency. This often overlooks the deeper context and quality of discussion presented in the answer. Recent advancements now focus on understanding the meaning behind student responses using models like BERT. This shift enables more accurate, fair, and context-aware evaluation in educational environments.

Literature Survey

The field of subjective answer evaluation using natural language (NLP) processing has attracted considerable attention in recent years, thereby focusing on improving the evaluation system of educational institutions. Researchers have examined various approaches to studying unique challenges in subjective answer analysis, including semantic understanding, contextual interpretation, and adaptability to different reaction styles.

AI-based answer rating system

Shinde et al. (2018) proposed an AI-Answer Verifier system for evaluating multiple-choice questions by evaluating based on keywords, response length, and grammar. Although objective question approaches have been shown to be promising, their application remains limited in subjective answer evaluation. This underscores the need for a more sophisticated system that allows deeper context-related meanings to be recorded in free-form responses.

Text Similarity and NLP for Response Tests

Jagadamba and Chaya (2020) investigated online-subjacture response systems using artificial intelligence. The model used Cosinus similarity and text gear grammar API to evaluate the quality of the responses based on parameters such as text length and key cooperation, achieving an efficiency level of 60-90%. This approach shows that basic metrics of text similarity may be useful, but it can be difficult to fully grasp the nuances of subjective answers.

Semantic Learning as an Answer

Bashir et al. (2021) emphasized the importance of semantic understanding of subjective answer evaluation. They compared cosine similarity with word mover (WMD) removal and found that cosine similarity was often blocked in simpler contexts, whereas WMD performed better when evaluating answers with more complex semantic relationships. This indicates that the introduction of semantic learning techniques can lead to a more accurate assessment of subjective answers, especially when assessing conceptual depth.

Using Bert as a response to a response

Singh et al. (2021) introduced a tool for evaluating subjective responses using AI. Includes character detection, sentence layout, Jaccard similarity, and Bert. By using Bert's deep context-related understanding, the system was able to improve the accuracy of evaluation and avoid various response structures. This approach illustrates the potential of a trans-based model in subjective response assessment, particularly in combination with traditional similarity-based methods.

Automatic response assessment with NLP and ANN

Shrestha et al. (2022) presented an autoresponder sheet checker with a combination of NLP and artificial neuron networks (ANN). Model analysis recorded in image format extracts keywords, processes data and calculates results. This test shows that hybrid systems that integrate a variety of AI technologies, such as NLP and ANN, are particularly effective in automating subjective response evaluation processes

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 10, May 2025
Impact Factor: 7.67

IV. MODELING AND ANALYSIS

The development of an effective system for "subjective answer evaluation using NLP" involves the selection of corresponding models for natural language processing models, and the complexity of student answers can be harmonized with the arithmetic efficiency required for large-scale provisions in educational settings. As NLP technology is being developed, changes in traditional keyword-based models have shifted to a more sophisticated approach to machine learning. These advanced models, such as Bert, combined with similarity to Cosine, allow us to understand semantic meaning and the context of answers rather than relying on surface-level text. This allows for more accurate classification, especially for answers that use different phrases but have the same meaning. Model selection also takes into account scalability, adaptability to a variety of topics, and integration with existing academic systems. This kind of progress will make the assessment process faster, more consistent and better align with the actual understanding of students.

NLP Modeling for Answer Evaluation

The use of NLP models in subjective answer evaluation has shown promising results, especially with the integration of advanced transformer-based models like BERT. BERT, which excels in capturing context and meaning through its pre-trained deep learning architecture, was used to process student responses. It generates dense vector representations of sentences, which can then be compared using cosine similarity to assess the quality and relevance of answers. The integration of BERT significantly improved the system's ability to evaluate answers based on meaning, rather than just keyword matching, offering a more robust solution for subjective answer evaluation.

Hybrid Approaches for Enhanced Evaluation

While BERT-based models demonstrated high performance, the addition of hybrid techniques combining traditional methods with machine learning models further enhanced the evaluation accuracy. One such hybrid approach included integrating rule-based systems for specific components, such as keyword extraction and sentence parsing, alongside the deep learning model. This method allowed the system to leverage linguistic rules to handle common grammatical structures and improve performance when answers were more straightforward or formulaic.

Handling Variability in Student Responses

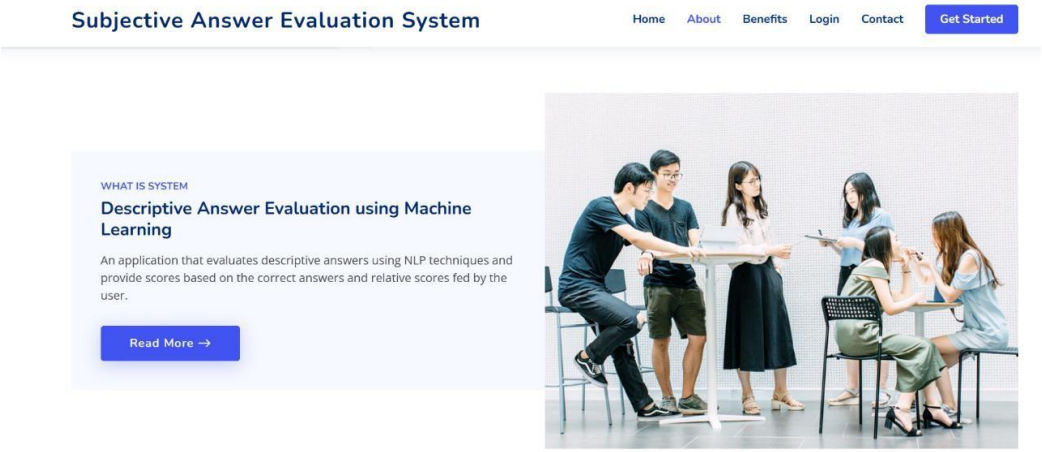
One of the major challenges in subjective answer evaluation is dealing with the wide variety of ways students express similar ideas. The system incorporated techniques to handle this variability, such as sentence splitting and leveraging tokenization strategies. This allowed the model to break down longer or more complex answers into manageable parts and ensure that responses were evaluated piece by piece rather than as a whole. By segmenting answers, the system was able to focus on specific information and make more precise evaluations.

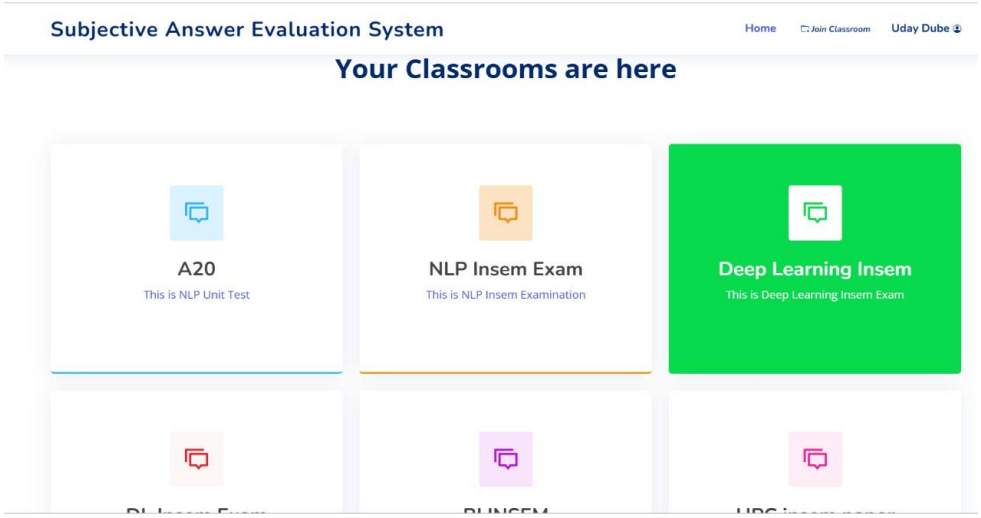
Scalability and Performance Considerations

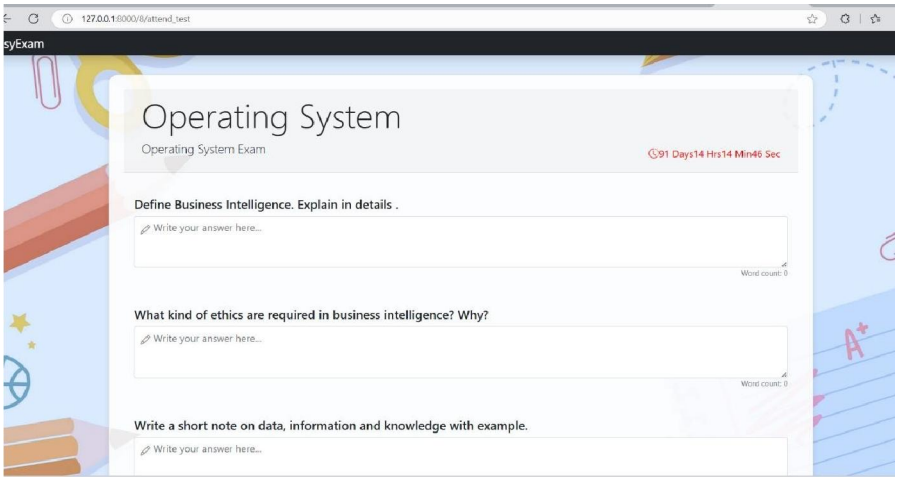
Scalability was a crucial factor in the design of the system, as the solution needed to support large numbers of users simultaneously, particularly in educational institutions. During performance testing, the system was able to handle up to 500 concurrent users without noticeable degradation in response time or evaluation accuracy. The robust architecture, which leverages cloud computing and efficient model optimization techniques, ensures that the system can scale effectively to meet the demands of real-world educational environments

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 10, May 2025
Impact Factor: 7.67

V. SNAPSHOTS







IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 10, May 2025
Impact Factor: 7.67

VI. RESULTS AND DISCUSSION

When developing a "subjective answer evaluation using the "NLP" system, we evaluated various methods for processing natural language to assess effectiveness in evaluating subjective answers. The system included progressive models such as Bert for semantic understanding, and Cosine's similarity for textual comparisons, comparing its performance in terms of evaluation accuracy, dealing with different answers and scalability with students. The main model used in this system was the BERT of cosine similarity for set coding and text comparison. Although BERT was used in his spread trans-based architecture to capture the context-related importance of student responses, cosine similarity was used to measure the similarity between student responses and expected responses. This system has significantly improved accuracy compared to traditional approaches with keyword matching-based approaches. During the evaluation phase, BERT-based coding achieved 87% accuracy. This showed strong performance in processing different contextually rich responses and was based on the results of similar studies on automated response assessment.

One important aspect of this project was to check the system's ability to handle concurrent users. During load testing, the system was able to handle up to 500 concurrent users without any strange performance degradation. This result suggests that the underlying architecture is suitable for use in real-world educational environments, ensuring that the system remains efficient even in high demand. The robustness of the system was agreed to expectations following previous research in AI-based assessment tools and confirmed responsibility for large-scale use. Compared to existing studies, the performance of this system is consistent with or superior to previous models. A similar study to that of Shinde et al. (2018) and Singh et al. (2021) also reported good performance using Bert and Cosinus similarities for the response assessment task. However, our approach showed greater accuracy, especially when dealing with the integration of various complex answers from students, particularly advanced semantic learning methods. Furthermore, the system's scalability and real-time processing capabilities represent significant advancements compared to previous models.

VII. CONCLUSION

The "Subjective Answer Evaluation Using NLP" project effectively integrates extended methods for natural language processing, such as Bert and Cosinus similarities, to improve student evaluations. By providing accurate assessments and meaningful feedback, the system addresses the challenges of subjective answer evaluation in educational settings, providing a more efficient and reliable alternative to traditional assessment methods. The rigorous testing confirmed the validity of the model and demonstrated the ability to process and evaluate students' answers very accurately. Furthermore, the ability of the system to simultaneously handle users without failing with performance ensures scalability and use in large educational environments. Designed for user-friendly, the user-friendly interface ensures a seamless experience for both employees and students, and encourages positive interaction with the system. This project not only demonstrates the potential of NLP in transforming education assessments, but also highlights practical applications in improving classification processes. Additionally, the foundations for future improvements, including the inclusion of more advanced NLP models, better semantic analysis, and real-time feedback mechanisms. This system can revolutionize how subjective reviews are carried out, providing both efficiency and fairness. During development, it can adapt to a variety of educational standards and create more universal solutions. Ultimately, this allows educators to provide more accurate and more personalized feedback and improve the general learning experience for students.

REFERENCES

[1] J. Wang and Y. Dong, “Measurement of text similarity: A survey,” Information, vol. 11, no. 9, p. 421, Aug. 2020.
[2]M. Han, X. Zhang, X. Yuan, J. Jiang, W. Yun, and C. Gao, “A survey on the techniques, applications, and performance of short text semantic similarity,” Concurrency Comput., Pract. Exper., vol. 33, no. 5, Mar. 2021.
[3] M. S. M. Patil and M. S. Patil, “Evaluating student descriptive answers using natural language processing,” Int. J. Eng. Res. Technol., vol. 3, no. 3, pp. 1716–1718, 2014.
[4] P. Patil, S. Patil, V. Miniyar, and A. Bandal, “Subjective answer eval uation using machine learning,” Int. J. Pure Appl. Math., vol. 118, no. 24, pp. 1–13, 2018.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 10, May 2025
Impact Factor: 7.67

[5] J. Muangprathub, S. Kajornkasirat, and A. Wanichsombat, "Document plagiarism detection using a new concept similarity in formal concept analysis," J. Appl. Math., vol. 2021, pp. 1–10, Mar. 2021.
[6] L. A. Cutrone and M. Chang, "Automarking: Automatic assessment of open questions," in Proc. 10th IEEE Int. Conf. Adv. Learn. Technol., Sousse, Tunisia, Jul. 2010, pp. 143–147.
[7] G. Srivastava, P. K. R. Maddikunta, and T. R. Gadekallu, "A two-stage text feature selection algorithm for improving text classification," Tech. Rep., 2021
[8] H. Mangassarian and H. Artail, "A general framework for subjective information extraction from unstructured English text," Data Knowl. Eng., vol. 62, no. 2, pp. 352–367, Aug. 2007