ZKG INTERNATIONAL ISSN: 2366-1313

AI BASED AUTOMATIC SUBJECTIVE ANSWER EVALUATION SYSTEM

[1]Mrs.K.PRATHYUSHA, [2]K.SAI KUMAR, [3]K.RUKMINI, [4]M.UDAY KIRAN REDDY

[1](Assistant Professor) ,**CSE.** Teegala Krishna Reddy Engineering College Hyderabad

[234]B,tech scholar ,**CSE.** Teegala Krishna Reddy Engineering College Hyderabad

ABSTRACT

Currently, we've seen that the descriptive answering systems are replaced by objective type of examinations. But when the universities, and institutions want to check the factual and deep knowledge of a pupil the descriptive answering system plays an important part. Checking of the descriptive answers are a excited job for the preceptors hence we bear some automatic system for Checking of the descriptive answers to reduce this excited Work. Descriptive paper evaluation is a tricky and tiresome task to do by homemade labor. The AI grounded automatic digital answer evaluation system for descriptive answering is an automatic system which checks the descriptive type of answers on its Grounded on given input Keywords. The colorful machine literacy, natural language processing ways, and tools similar as Wordnet, word2vec, Word Mover's Distance(WMD), Multinomial Naive Bayes(MNB) are used to estimate descriptive answers automatically.

1.INTRODUCTION

Private questions and answers can assess the performance and capability of a pupil in an open-concluded manner. The answers, naturally, aren't bound to any constraint, and scholars are free to write them according to their mindset and understanding of the conception. With that said, several other vital differences separate private answers from their objective counterpart. For one, they are much longer than the objective questions. Secondly, they take further time to write . also, they carry much further environment and take a lot of attention neutrality from the school teacher assessing them. Evaluation of similar questions using computers is a tricky task, substantially because natural language is nebulous. Several preprocessing way must be performed, similar as drawing the data and

ISSN: 2366-1313

tokenization before working on it. Also the textual data can be compared using colourful ways similar as document similarity, idle semantic structures, conception graphs, ontologies. The final score can be estimated grounded on similarity, Keywords presence, structure, language. Several attempts have been made in the history to break this problem, but there is still room for advancements, some of which is bandied in this . Private examinations are considered more complex and scary by both scholars and preceptors due to their one point,, environment.

A private answer demands the checker check every word of the answer for scoring laboriously, and the checker's internal health, fatigue, and neutrality play a massive part in the overall result. Thus, it's much further time and resource-effective to let a system handle this tedious and critical task of assessing private answers.assessing object answers with machines is veritably easy and doable. A program can be feed with question and one word answers that can snappily collude scholars responses. Nonetheless, private answers are much more grueling to attack. They are varied in length and contain a vast quantum of vocabulary. Likewise, people tend to use antonyms and accessible bowdlerization, which make the process that important tricky.

2. LITERATURE SURVEY

Automatic Subjective Answer Evaluation (ASAE) systems have gained significant attention due to their potential to streamline the grading process in educational institutions. These systems utilize Artificial Intelligence (AI) techniques to assess subjective answers, providing a faster and often more consistent evaluation compared to manual grading. This literature survey aims to explore various approaches, methodologies, and advancements in AI-based ASAE systems.

1. Traditional Techniques in ASAE:

o Early ASAE systems often relied on rule-based or heuristic approaches. These systems used predefined rules to evaluate answers based on grammar, syntax, and keywords.

o However, these approaches lacked flexibility and struggled to handle diverse answers or nuanced language.

2. Machine Learning-Based Approaches:

o With advancements in machine learning, ASAE systems began utilizing techniques

ISSN: 2366-1313

such as Natural Language Processing (NLP) and supervised learning.

o Supervised learning models, particularly deep learning architectures like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have shown promise in automatically evaluating subjective answers.

o These models can learn from labeled datasets, mapping input answers to corresponding grades.

3. Semantic Analysis and Contextual Understanding:

o Recent research focuses on enhancing ASAE systems' understanding of semantics and context in subjective answers.

o Techniques such as word embeddings, contextual embeddings (e.g., BERT), and attention mechanisms help capture the semantic meaning of answers.

o By considering the context of the question and the answer, these systems can provide more accurate evaluations.

4. Ensemble and Hybrid Approaches:

o Ensemble methods, combining multiple models or features, have demonstrated improved performance in ASAE.

o Hybrid approaches, integrating rule-based heuristics with machine learning models, aim to leverage the strengths of both approaches for better evaluation accuracy.

5. Evaluation Metrics and Benchmarking:

o Developing appropriate evaluation metrics is crucial for assessing the performance of ASAE systems.

o Common metrics include accuracy, precision, recall, F1-score, and Cohen's kappa coefficient.

o Benchmarking datasets, such as ASAG (Automatic Short Answer Grading), provide standardized datasets for evaluating ASAE systems' performance.

6. Challenges and Future Directions:

o Despite advancements, challenges remain in developing robust ASAE systems.

o Handling linguistic variations, understanding context, and dealing with subjective grading criteria are ongoing challenges.

o Future research directions include exploring Explainable AI (XAI) techniques to provide transparent grading criteria and addressing biases in ASAE systems.

3.SYSTEM DESIGN

ISSN: 2366-1313

The automated subjective paper evaluation system is structured around a modular architecture designed to efficiently handle the evaluation of descriptive answers using advanced machine learning and natural language processing techniques. At its core, the system consists of several interdependent components, each fulfilling specific roles and responsibilities to ensure seamless operation and accurate evaluation.

The User Interface component serves as the primary point of interaction between the system and its users, primarily students. It provides an intuitive and user-friendly interface through which students can submit their descriptive answers for evaluation and view their corresponding grades. This component is typically implemented using web-based or desktop applications, offering accessibility across different devices and platforms. Complementing the User Interface is the Admin Module, which empowers administrators, typically educators or instructors, with comprehensive control over the system. This module facilitates various administrative tasks, including the provision of labeled training data for the machine learning model, configuration of grading criteria and parameters, and monitoring of system performance. Administrators can adjust settings and parameters as needed to tailor the evaluation process to specific requirements or preferences.

At the heart of the system lies the Machine Learning and Natural Language Processing Component, which is responsible for performing the actual evaluation of descriptive answers. This component employs a combination of sophisticated machine learning algorithms and natural language processing techniques to extract features from submitted answers, measure similarity to solution statements or keywords, and classify answers based on predefined grading criteria. Techniques such as Word2Vec, TF-IDF, cosine similarity, and multinomial naive Bayes are commonly utilized within this component to achieve accurate and reliable evaluation results. Additionally, the component is trained on labeled training data provided by 7 administrators, allowing it to continuously improve and adapt to evolving patterns and nuances in student responses. To facilitate efficient data management and retrieval,

the system incorporates a Database Component, which serves as the central repository for all relevant data. This includes not only the training data provided by administrators but also system

ISSN: 2366-1313

configurations and parameters, as well as user data such as submitted answers and corresponding grades. By storing data in a structured and organized manner, the Database Component enables quick and reliable access for analysis, reporting, and system maintenance. Finally, the Feedback and Reporting Component plays a crucial role in facilitating ongoing improvement and optimization of the system. This component collects feedback from both users (students) and administrators, allowing for insights into the system's performance, accuracy, and usability. Feedback can be used to identify areas for improvement, refine evaluation criteria, and enhance overall system effectiveness. Additionally, the component generates reports on system performance, providing administrators with valuable insights and metrics to inform decision-making and strategic planning.

3.1 System Architecture

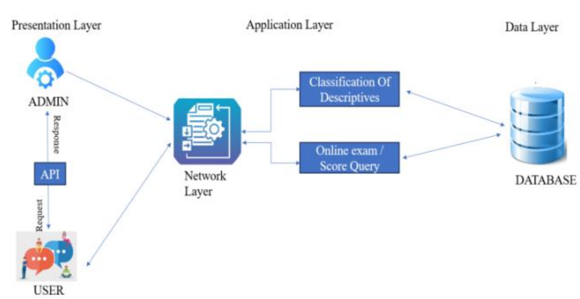

Fig 3.1 System Architecture

3.2 Activity Diagram

An Activity Diagram is a type of Unified Modeling Language (UML) diagram that illustrates the flow of control or the flow of activities within a system. It depicts the actions, decisions, and transitions between different states or activities in a system or business process. Activity Diagrams are widely used in software development and business process modeling to visualize and analyze complex workflows, business processes, and system behaviors. They provide stakeholders with a clear and intuitive representation of the sequence of activities and decision points within a system or process, helping to identify bottlenecks, optimize workflows, and improve system efficiency.

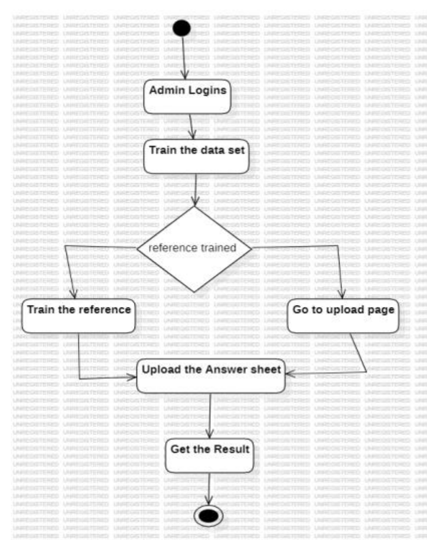

Fig 3.2 Activity Diagram

ISSN: 2366-1313

7.OUTPUT SCREENS

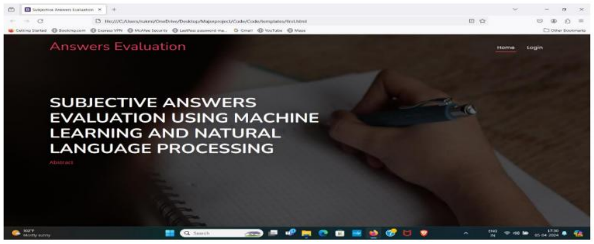

Fig 4.1 Interface

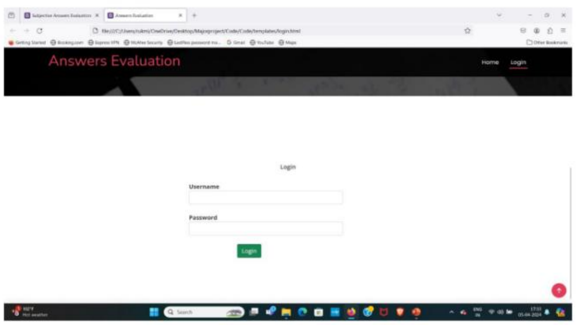

Fig 4.2 Admin logins

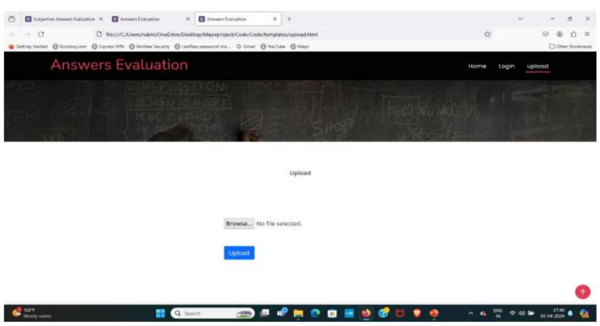

Fig 4.3 Upload The Dataset

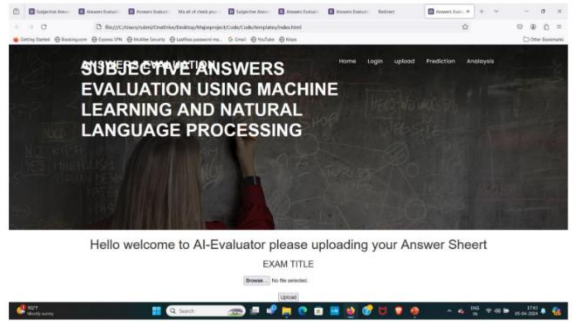

Fig 4.4 Upload the Answer Script

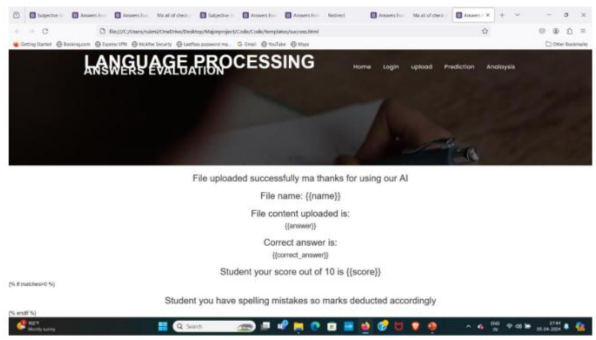

Fig 4.5 Displays the result

5.CONCLUSION

This paper proposed a new approach to private answers evaluation grounded on machine literacy and natural language processing ways. Two score vaticination algorithms are proposed, which produce up to 88 accurate scores. colorful similarity and diversity thresholds are studied, and colorful other measures similar as the keyword's presence and chance mapping of rulings are employed to overcome the abnormal cases of semantically loose answers. The trial results show that on average word2vec approach performs better than traditional

ISSN: 2366-1313

word embedding ways as it keeps the semantics complete. likewise, Word Mover's Distance performs better than Cosine Similarity in utmost cases and helps train the machine literacy model briskly. With enough training, the model can stand on its own and prognosticate scores without the need for any semantics checking. In terms of unborn advancements, the word2vec model can be trained especially for private answers evaluation of a particular sphere, and with large data sets, the number of classes or grades in the model can be significantly increased. private answers evaluation remains an intriguing problem to attack, and in the future, we hope to find more effective ways to break this problem.

6.FUTURE ENHANCEMENT

Future enhancements for AI-based automatic subjective answer evaluation systems can focus on several key areas to further improve their accuracy, efficiency, and applicability:

1. Contextual Understanding: Enhance the system's ability to understand and analyze the context of both questions and answers. Incorporating contextual embeddings and attention mechanisms can help capture the subtle nuances in language, leading to more accurate evaluations.

2. Explainability: Develop techniques to make the evaluation process more transparent and interpretable. Explainable AI (XAI) methods can help provide insights into how the system arrives at its decisions, enabling instructors and students to understand the grading criteria better.

3. Adaptability to Diverse Subjects: Design the system to be adaptable to a wide range of subjects and domains. Customizable models and training pipelines can allow the system to learn specific grading criteria for different subjects, ensuring accurate evaluations across diverse disciplines.

4. Feedback Mechanisms: Implement feedback mechanisms that allow instructors to provide feedback on the system's evaluations. This feedback loop can be used to continuously improve the system's performance over time and address any discrepancies or errors in grading.

5. Bias Mitigation: Develop techniques to mitigate biases inherent in the training data or the evaluation process. Fairness-aware learning algorithms and bias detection mechanisms can help ensure that the system's evaluations are fair and unbiased across different demographic groups.

ISSN: 2366-1313

6. Multimodal Integration: Explore the integration of multimodal inputs, such as text, images, and audio, to evaluate answers comprehensively. This can be particularly useful for subjects that involve visual or auditory components, such as art or music.

7. Real-time Feedback and Adaptive Learning: Enable real-time feedback and adaptive learning capabilities that allow the system to adapt and improve based on user interactions and performance data. This can enhance the system's effectiveness in providing personalized feedback and support to students.

8. Integration with Learning Management Systems (LMS): Integrate the ASAE system seamlessly with existing learning management systems used in educational institutions. This integration can streamline the grading process for instructors and provide students with immediate feedback on their performance.

9. Scalability and Efficiency: Develop techniques to ensure scalability and efficiency, allowing the system to handle large volumes of answers and perform evaluations in a timely manner, especially during peak periods such as exams.

10. Ethical Considerations: Address ethical considerations related to privacy, data security, and the responsible use of AI in education. Ensure that the system adheres to ethical guidelines and regulations to protect the rights and interests of students and instructors.

7.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 evaluation using machine learning,’’ Int. J. Pure Appl. Math., vol. 118, no. 24, pp. 1–13, 2018.

[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] X. Hu and H. Xia, ''Automated assessment system for subjective questions based on LSI,'' in Proc. 3rd Int. Symp. Intell. Inf. Technol. Secur. Informat., Apr. 2010, pp. 250–254.

[7] M. Kusner, Y. Sun, N. Kolkin, and K. Weinberger, ''From word embeddings to document distances,'' in Proc. Int. Conf. Mach. Learn., 2015, pp. 957–966.

[8] C. Xia, T. He, W. Li, Z. Qin, and Z. Zou, ''Similarity analysis of law documents based on Word2vec,'' in Proc. IEEE 19th Int. Conf. Softw. Qual., Rel. Secur. Companion (QRS-C), Jul. 2019, pp. 354–357.

[9] H. Mittal and M. S. Devi, ''Subjective evaluation: A comparison of several statistical techniques,'' Appl. Artif. Intell., vol. 32, no. 1, pp. 85–95, Jan. 2018.

[10] 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.