Seminar Report

On

**Automatic Answer Checker Using NLP**

By

**Swanand Mahajan**

PA32 | 1032210868

Under the Guidance of -

**Prof. Sagar Apune**

**Dr. Vishwanath Karad MIT-World Peace University**

**Department of Computer Engineering & Technology**

**School of Computer Science & Engineering**

**\* 2023-2024 \***

**Dr. Vishwanath Karad MIT-World Peace University (MIT-WPU)**

**School of Computer Science & Engineering**

**CERTIFICATE**

This is to certify that Mr. /Ms. Swanand Mahajan of B.Tech CSE (AIDS), Semester-VI, PRN. No. 1032210868, has successfully completed a seminar on Automatic Answer Checker using NLP

This seminar is satisfactorily submitted & delivered during the academic year 2023-2024 towards the partial fulfilment of degree of Bachelor of Computer Science Engineering under Dr. Vishwanath Karad MIT- World Peace University, Pune.

Prof. Sagar Apune Dr. Balaji Patil

Seminar Guide Name & Sign Program Director

Department of CET, SoCSE

Date & Place -:

**List of Figures:-**

| **Sr. No** | **Figure** |
| --- | --- |
| **1** | Basic Working of Model |
| **2** | Working of Answer Checker |

**Abbreviations:-**

| **Sr.No.** | **Abbreviations** | **Definition** |
| --- | --- | --- |
| 1 | NLP | Natural Language Processing |
| 2 | AI | Artificial Intelligence |
| 3 | OCR | Optical Character Recognition |
| 4 | PoS | Part-of-Speech |
| 5 | BoW | Bag of Words |

**INDEX**

| **Sr. No.** | **Topic Name** | **Page no.** |
| --- | --- | --- |
| 1 | Abstract & Keywords | 6 |
| 2 | Introduction | 7 - 8 |
| 3 | Literature Survey | 9 - 10 |
| 4 | Details of Technology | 11 - 12 |
| 5 | Details of Design | 13 - 15 |
| 6 | Conclusion | 16 - 19 |
| 7 | References | 20 |
| 8 | Appendix | 21 |

**Acknowledgement**

I, Swanand Mahajan (PRN: 1032210868), would like to extend my sincere gratitude to Prof. Sagar Apune, my esteemed seminar guide, for his invaluable support, guidance, and encouragement throughout the preparation and presentation of this seminar. His expertise has been instrumental in shaping the content and direction of this Seminar. I am deeply grateful for his unwavering commitment and dedication, which have greatly enriched my understanding and perspective on this vital topic.

Thank you Sir, for your mentorship and inspiration.

**Abstract**

The need for automation in answer assessment systems is covered in the study, especially as it relates to theory answers. The current system makes it more difficult for evaluators because it takes more time and effort to evaluate theory responses. The study introduces a computerised technology that streamlines the paper-correcting procedure. Although handwritten evaluations have the potential to be biased, this technique is essential for assessing student performance. This report presents an automatic answer script evaluation method based on natural language processing. Our experiment involves extracting text from the answer script, comparing the summarised extracted text to the previously saved correct answers, and then giving each computed parameter a weighted value to rate the answer script. Using keyword-based summarization approaches, we have generated summaries from the retrieved text. Here, the final mark is generated using four similarity measures (Cosine, Jaccard, Bigram, and Synonym). The study also examines the use of natural language processing (NLP) techniques to assess exam papers, including looking for grammatical faults, performing syntactic analysis, semantic similarity, and database storage. The research comes to the opinion that assessment and evaluation processes can be made more efficient and reliable through automation.

Keywords—Natural Language Processing, Similarity Score, Answer Checker.

**Introduction**

The integration of technology into education has transformed the learning landscape, offering new possibilities for assessment and evaluation. Automated assessment systems, powered by advancements in natural language processing (NLP) and artificial intelligence (AI), have emerged as effective tools for streamlining the grading process and providing timely feedback to students. These systems are particularly beneficial for evaluating descriptive answers, which require a deeper understanding of language and context.

Manual evaluation of descriptive answers is a time-consuming and labour-intensive task, often prone to subjectivity and inconsistency. Automating this process not only saves time and effort but also ensures more objective and consistent grading across all submissions. By leveraging NLP techniques, such as text extraction, summarization, and similarity calculation, automated assessment systems can analyse and evaluate answer scripts with a high degree of accuracy.

This report presents a study on automating the assessment of answer scripts for descriptive questions using NLP. The objective is to develop a system that can extract text from images, summarise the text, preprocess it for analysis, calculate similarity measures, and assign grades based on predefined parameters. The system aims to demonstrate the efficiency and accuracy of automated grading compared to manual evaluation.

The report is structured as follows: it begins with a literature review on automated assessment systems and NLP in education, highlighting the benefits and challenges of integrating technology into the assessment process. The methodology section describes the approach used to develop the automated grading system, including text extraction, summarization, preprocessing, similarity calculation, and grading. The experimental setup details the testing environment and dataset used for evaluation

.

Results from the experiment show the effectiveness of the automated grading system in accurately and efficiently evaluating descriptive answers. The discussion section interprets the results in the context of the research objectives and explores the implications for automated assessment systems in education. Finally, the conclusion summarises the key findings of the study and provides recommendations for future research and implementation.

**Literature Survey**

Examining student's answer scripts is an important part of determining their performance. Teachers use various techniques to determine the level of student knowledge, including short-answer questions, descriptive answers, and multiple-choice questions. Evaluation of multiple-choice and short-answer questions is quite simple and takes less time, while evaluation of descriptive replies can take more time. As a result, numerous methods for automating the evaluation of response scripts have been developed; some of these methods are covered in the sections that follow.

| Title | Author | Year | Work Done |
| --- | --- | --- | --- |
| Automatic Answer Sheet Checker | Ronika Shrestha,  Raj Gupta and Priya Kumari | April 5, 2022 | The literature review explores an AI-driven automatic answer checker system employing CNN and image processing for human-like grading, aiming to save time and resources. The proposed model emphasises keyword identification, specifically addressing the lack of handwriting scanning in previous systems, proposing a solution to evaluate handwritten answer sheets in PDFs/documents through keyword-based grading. |
| Automated Answer-Checker | Vasu Bansal,  M.L. Sharma, Krishna Chandra Tripathi | 2020 | The paper "Automated Answer-Checker" presents an algorithm using AI and text similarity to automatically evaluate subjective answers in exams. The system simplifies the assessment process, considering keywords, grammar, and similarity. It offers a user-friendly interface for teachers and students, aiming to streamline the evaluation of subjective answers in online exams. |
| Online Subjective Answer Verifying System Using Artificial Intelligence | Prof. Priyadarshani Doke, Priyanka Gangane, Kesia S Babu, Pratiksha Lagad, Neha Vaidya | 2021 | The literature review categorises techniques into Statistical, Information Extraction, and Full Natural Language Processing. Statistical methods face limitations like handling synonyms. Recent studies stress the significance of weighted averages for optimised subjective answer evaluation. Systems like TESA and ASSESS aim to enhance efficiency. Specific 2021 studies explore graphical comparisons and machine learning models, emphasising the evolving landscape. The review highlights the need for a comprehensive system to address challenges in subjective assessment during online exams. |
| Automatic generation of short-answer questions in reading comprehension using NLP and KNN | Riza, L.S., Firdaus, Y., Sukamto | 2023 | The literature review covers methodologies like DeconStructure algorithm, POS tagging, and semantic pattern recognition for question generation. This study uses NLP and KNN, emphasising quality maintenance, automated generation, and POS tagging versatility. |

**Details of Technology**

The automated answer script evaluation system leverages several key technologies and libraries to achieve its functionality. These technologies enable the system to extract text, summarise the text, preprocess it for analysis, calculate similarity measures, and assign grades based on predefined parameters. The following are the main technologies used in the system:

1. **Python:**

* Python is the primary programming language used for developing the automated grading system.
* It offers readability and extensive library support, making it suitable for natural language processing (NLP) tasks.
* Python's versatility allows for the implementation of various NLP functionalities required for text processing and analysis.

1. **NLTK (Natural Language Toolkit):**

* NLTK is a powerful platform for building Python programs that work with human language data.
* It provides tools for text processing and analysis, including tokenization, stemming, tagging, parsing, and semantic reasoning.
* NLTK is used for text preprocessing in the system, such as tokenization, stopword removal, lemmatization, and part-of-speech tagging.

1. **Pytesseract:**

* Pytesseract is a Python wrapper for Google's Tesseract-OCR Engine.
* It is used for optical character recognition (OCR), allowing the system to extract text from images of answer scripts.
* Pytesseract enables the conversion of handwritten or printed text into a machine-readable format for further analysis.

1. **Text Summarization Techniques:**

* Text summarization techniques are used to condense extracted text into a concise form, focusing on key information.
* The system employs keyword-based summarization methods to identify and extract key terms or phrases representing the main ideas.
* This helps reduce the length of the text while preserving its meaning for easier analysis and evaluation.

1. **Similarity Calculation Algorithms:**

* These algorithms determine the similarity between the student's response and the correct answer.
* The system uses several measures, including cosine similarity, Jaccard similarity, bigram similarity, and synonym similarity.
* Each measure provides a unique perspective on the similarity between the texts, contributing to a comprehensive evaluation process.

1. **Language Check:**

* The system uses the "language-check" Python tool to identify and tally spelling and grammatical mistakes in the text.
* This tool ensures that the evaluation process considers the language accuracy of the student's response, in addition to its content.
* Language check provides a more holistic assessment of the answer script, considering both correctness and language accuracy.

**Details of Design**

The design of the automated answer script evaluation system involves several key components and considerations to ensure its functionality and efficiency. The system is designed to automate the grading process for descriptive answers using natural language processing (NLP) techniques. The following are the main details of the system's design:

* **Overall Architecture:** The system follows a modular architecture, with distinct components for text extraction, summarization, preprocessing, similarity calculation, and grading. Each component is designed to perform a specific task in the evaluation process, contributing to the overall functionality of the system.
* **Text Extraction:** The text extraction component uses Pytesseract, an OCR tool, to extract text from images of answer scripts. This component ensures that handwritten or printed text is converted into a machine-readable format for further analysis.
* **Text Summarization:** After text extraction, the system employs keyword-based text summarization techniques to condense the extracted text. This component focuses on identifying key concepts and information in the text to provide a concise summary for analysis.
* **Text Preprocessing:** NLTK is used for text preprocessing, including tokenization, stopword removal, lemmatization, and part-of-speech tagging. These preprocessing steps are essential for preparing the text for similarity calculation and grading.
* **Similarity Calculation:** The system calculates four similarity measures - cosine similarity, Jaccard similarity, bigram similarity, and synonym similarity - to determine the similarity between the student's response and the correct answer. Each similarity measure provides a different perspective on the similarity between the texts, contributing to a comprehensive evaluation process.
* **Grading:** Based on the calculated similarity measures and the presence of grammatical and spelling errors, the system assigns grades to the answer scripts. The grading process aims to provide an objective and efficient evaluation of descriptive answers, taking into account both content and language accuracy.
* **User Interface:** The system includes a user-friendly interface for inputting answer scripts, viewing grades, and accessing evaluation results. The interface is designed to be intuitive and easy to use, allowing users to interact with the system effectively.

Overall, the design of the automated answer script evaluation system is focused on efficiency, accuracy, and user-friendliness, ensuring that it meets the needs of educators and students alike in the evaluation of descriptive answers.

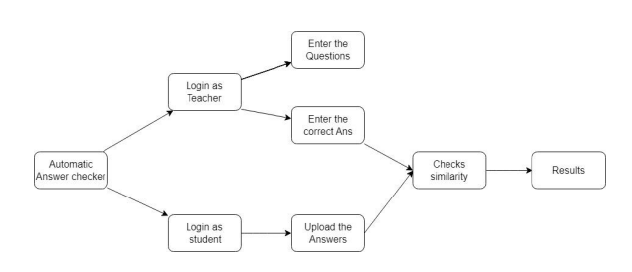


Fig. 1 Basic Working of Model

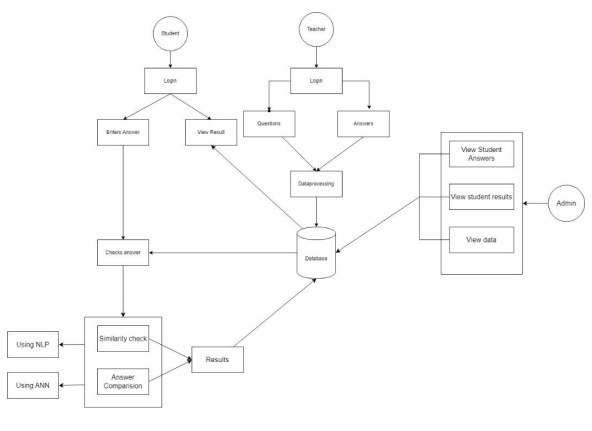


Fig. 2 Working of Answer Checker

**Conclusion**

1. **Summary of Main Finding:**

The automated answer script evaluation system developed in this study demonstrates the feasibility and effectiveness of using natural language processing (NLP) techniques for grading descriptive answers. The system leverages Python programming language, NLTK library, Pytesseract for OCR, and various similarity calculation algorithms to automate the grading process.

* **Text Extraction and Summarization:** The system successfully extracts text from images of answer scripts using Pytesseract and summarises the text using keyword-based methods. This allows for the conversion of handwritten or printed text into a machine-readable format and condenses the text for easier analysis.
* **Text Preprocessing:** NLTK is used for text preprocessing, including tokenization, stopword removal, lemmatization, and part-of-speech tagging. These preprocessing steps are essential for preparing the text for similarity calculation and grading.
* **Similarity Calculation:** The system calculates four similarity measures - cosine similarity, Jaccard similarity, bigram similarity, and synonym similarity - to determine the similarity between the student's response and the correct answer. These measures provide a comprehensive evaluation of the text from different perspectives.
* **Grading:** Based on the calculated similarity measures and the presence of grammatical and spelling errors, the system assigns grades to the answer scripts. The grading process aims to provide an objective and efficient evaluation of descriptive answers, taking into account both content and language accuracy.

Overall, the system demonstrates the potential of NLP techniques to automate the grading process for descriptive answers, offering a more efficient and reliable alternative to manual grading methods. Further improvements and refinements can be made to enhance the system's accuracy and scalability for broader application in educational settings.

1. **Implications:**

The automated answer script evaluation system developed in this study has several implications for educational assessment and natural language processing (NLP) research:

* **Efficiency and Time Savings:** The system offers a more efficient and time-saving method for grading descriptive answers compared to manual evaluation. By automating the grading process, educators can save time and focus on other aspects of teaching and learning.
* **Consistency and Objectivity**: Automated grading ensures consistency and objectivity in the evaluation process. The system applies predefined criteria consistently, reducing the potential for bias or subjective judgement in grading.
* **Scalability:** The system can be scaled to handle a large number of answer scripts, making it suitable for use in large-scale assessments such as exams or standardised tests. This scalability can help educational institutions streamline their assessment processes.
* **Feedback and Improvement:** The system provides detailed feedback to students based on their answers, highlighting areas of improvement. This feedback can help students understand their mistakes and improve their writing skills.
* **Integration with Learning Management Systems (LMS):** The system can be integrated with existing learning management systems (LMS) to streamline the assessment process further. This integration can facilitate seamless grading and feedback delivery to students.
* **Advancements in NLP Research:** The system contributes to advancements in NLP research by demonstrating the practical application of NLP techniques in educational assessment. The system's design and implementation can serve as a foundation for future research in automated grading systems.

Overall, the implications of the automated answer script evaluation system extend beyond educational assessment, offering insights into the potential of NLP in improving efficiency, objectivity, and scalability in various domains.

1. **Directions for Future Work:**

The automated answer script evaluation system developed in this study lays the groundwork for future research and development in the field of natural language processing (NLP) and educational assessment. Several directions for future work are suggested based on the findings and limitations of the current study:

* **Enhanced NLP Techniques:** Explore advanced NLP techniques, such as deep learning models like BERT or transformers, to improve the system's accuracy and performance in text analysis and summarization.
* **Multilingual Support:** Extend the system to support multiple languages, enabling its use in diverse linguistic environments and educational settings.
* **Integration with Learning Platforms**: Integrate the system with popular learning management systems (LMS) to streamline the grading process and provide seamless feedback to students and educators.
* **Feedback Mechanisms:** Implement interactive feedback mechanisms that allow students to receive personalised feedback on their answers, helping them understand their mistakes and improve their writing skills.
* **Real-time Evaluation:** Develop real-time evaluation capabilities to enable instant feedback to students, facilitating immediate learning and improvement.
* **Scalability and Performance**: Improve the system's scalability and performance to handle large volumes of answer scripts efficiently, ensuring reliable grading in high-stakes assessments.

1. **Final Thoughts:**

The development of an automated answer script evaluation system based on natural language processing (NLP) techniques has shown significant promise in improving the efficiency, objectivity, and scalability of the grading process for descriptive answers. By leveraging NLP algorithms and machine learning models, the system offers a viable solution to the challenges of manual grading, particularly in large-scale assessments.

Moving forward, further research and development in NLP and educational technology will continue to enhance the capabilities of automated grading systems, making them more accurate, reliable, and user-friendly. The integration of these systems into educational settings has the potential to revolutionise the way assessments are conducted, providing educators with valuable insights into student performance and enabling students to receive timely and constructive feedback on their work.

Overall, the development and implementation of automated grading systems represent a significant step forward in the field of educational assessment, offering a glimpse into the future of intelligent, data-driven evaluation methods.

**References**

[1] A. Mathur, S. Pradhan, P. Soni, D. Patel and R.Regunathan, "Automated Test Case Generation Using T5 and GPT-3," 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023

[2] Merien Mathew, Ankit Chavan, Siddharth Baikar,'ONLINE SUBJECTIVE ANSWER CHECKER',International Journal of Scientific & Engineering Research, Volume 8, Issue ISSN 2229-5518

[3] Ronika Shrestha, Raj Gupta and Priya Kumari,'Automatic AnswerSheet Checker',April 5, 2022

[4] Vasu Bansal,M.L. Sharma, Krishna Chandra Tripathi ,'Automated Answer-Checker',International Journal for Modern Trends in Science and Technology, 6(12): 152-155, 2020,ISSN: 2455-3778 online

[5] Riza, L.S., Firdaus, Y., Sukamto, R.A. et al. 'Automatic generation of short-answer questions in reading comprehension using NLP and KNN.' Multimed Tools Appl (2023).

[6] Rahman, Md & Siddiqui, Fazlul. 'NLP-based Automatic Answer Script Evaluation'. 4. 35-42.(2018)

[7] P. Bhongade, S. Girhay, A. M. Sheikh, R. Ghata, S. Ambadkar and C. Dusane, "Internet of Things - Enabled Smart Shoes for Blind People," 2022 IEEE Delhi Section Conference (DELCON), New Delhi, India, 2022

**Appendix**

Appendix A: Glossary of Terms

* Code Snippets: Include relevant code snippets or scripts used in the development of the automated answer script evaluation system.
* Sample Output: Provide sample output from the system, showcasing its ability to extract text from images, perform text summarization, and calculate similarity measures.
* System Architecture: Illustrate the system architecture, detailing the components and their interactions in the automated grading process.
* Dataset Description: Describe the dataset used for training and testing the system, including the source of the data and any preprocessing steps applied.
* Additional Figures and Tables: Include any additional figures, tables, or diagrams that support the findings and methodology of the study.
* References: Provide a list of references cited throughout the report, including research papers, books, and online resources.