IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Integrating Generative AI and OCR-Based Scoring Algorithms for Evaluating Educational Responses: A Bloom's Taxonomy Approach

[1]Abbas Pathan, [2]Vrushabh Sasane, [3]Sohum Thakre,[4] Shreyash Wadikar, Hitesh Chaudhari

[1]Department Of Computer Engineering,

[1]Sinhgad College Of Engineering ,Pune,India

Abstract: In the realm of educational technology, the assessment of student answers remains a significant challenge, particularly with the increasing integration of digital submissions. This study proposes a novel methodology that harnesses the power of generative artificial intelligence (AI), trained using Bloom's Taxonomy, to create model answers for educational questions. It then employs Optical Character Recognition (OCR) and Comma-Separated Values (CSV)-based scoring algorithms to automate the evaluation process of student responses submitted as JPEG images. This methodology not only simplifies the extraction and analysis of text from image-based submissions but also introduces a scoring system that evaluates answers based on the presence of relevant keywords and the adequacy of answer length, ensuring a comprehensive assessment of student understanding and learning outcomes. By automating the grading process, this approach aims to improve the efficiency and objectivity of educational assessments. Furthermore, the integration of generative AI with OCR technology exemplifies a significant advancement in educational assessment techniques, offering a scalable, efficient solution for educators and institutions. This research contributes to the field by demonstrating the feasibility and effectiveness of combining AI with image processing technologies to enhance the assessment of student learning, potentially revolutionizing educational practices and pedagogical strategies.

Index Terms - Generative AI, Bloom's Taxonomy, OCR Technology, Automated Assessment, Educational Technology, JPEG Image Analysis, CSV Data Processing, Scoring Algorithms, Keyword Matching, Educational Assessment.

I. INTRODUCTION

A. Background and Significance

The rapid advancement of educational technologies has underscored the necessity for innovative assessment methodologies that are both efficient and pedagogically sound. Traditional evaluation methods often struggle to keep pace with the dynamic and diverse learning landscapes facilitated by digital platforms. This backdrop necessitates the exploration of automated, AI-driven assessment tools capable of providing nuanced insights into student learning outcomes. The significance of integrating Bloom's Taxonomy into these tools lies in its foundational role in educational theory, offering a structured framework for categorizing educational goals and objectives. By harnessing generative AI trained on this taxonomy, alongside sophisticated image processing techniques, this research contributes to the field by proposing a novel, automated solution for academic evaluation. This approach not only promises to streamline the assessment process but also ensures that evaluations are aligned with cognitive learning objectives.

B.Problem Statement

Despite advancements in educational technology, traditional assessment methods face challenges in scalability, efficiency, and alignment with cognitive learning objectives. The problem lies in the need for an automated, reliable system capable of evaluating complex student responses, particularly those captured in non-textual formats like images. Additionally, there is a critical gap in ensuring that such evaluations are pedagogically grounded, leveraging frameworks like Bloom's Taxonomy to assess not just the content, but the cognitive depth of student answers. This research addresses these challenges by proposing a novel methodology for automated academic evaluation, integrating generative AI with image processing techniques.

C.Proposed Solution and Objectives

The proposed solution involves developing an AI-based system that employs generative AI trained on Bloom's Taxonomy for creating model answers, coupled with an automated evaluation process using OCR for text extraction from JPEG images and data analysis techniques. This system aims to objectively assess student responses by comparing them to AIgenerated model answers, considering both content relevance and cognitive depth. The objective is to enhance the accuracy, efficiency, and pedagogical alignment of academic evaluations, ensuring they reflect comprehensive learning outcomes and cater to diverse educational needs.

D. Scope and Comparative Analysis

The scope of this research encompasses the development and validation of an automated evaluation system integrating generative AI with OCR and data analysis to assess academic answers from images. It focuses on the accuracy, efficiency, and cognitive alignment of assessments with educational objectives, particularly leveraging Bloom's Taxonomy. The comparative analysis will benchmark this system against traditional and existing automated evaluation methods, highlighting its contributions in terms of scalability, pedagogical validity, and the ability to provide detailed, nuanced feedback on student learning outcomes.

Literature Review

The "Taxonomy for Learning, Teaching, and Assessing" [1] is a revised version of Bloom's Taxonomy developed by David R. Krathwohl and Lorin W. Anderson, offering a modernized framework for educational objectives. Organized into six hierarchical levels—Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating—the taxonomy guides educators in promoting increasingly complex cognitive processes. It serves as a versatile tool for curriculum development, lesson planning, and assessment design, emphasizing the cultivation of higher-order thinking skills like analysis, evaluation, and creation. Widely adopted in education, this taxonomy aids educators in fostering deeper understanding and critical thinking among students, ultimately enhancing the quality and effectiveness of teaching and learning experiences across diverse educational contexts. In the context of software engineering education, "Bloom's Taxonomy in Software Engineering Education: A Systematic Mapping Study" [2] investigates the integration of Bloom's taxonomy within software engineering education. Conducted by scholars in the field, the study systematically maps existing literature to analyze how Bloom's taxonomy is utilized in software engineering education. The research likely explores various aspects such as how the taxonomy informs curriculum design, instructional strategies, and assessment methods in software engineering programs. By synthesizing findings from multiple studies, this mapping study provides insights into the effectiveness and relevance of Bloom's taxonomy in enhancing software engineering education. It likely identifies trends, gaps, and opportunities for further research or improvement in this area, contributing valuable knowledge to both academia and practice in software engineering education. The application of Bloom's Taxonomy in programming fundamentals assessment is critically examined by [3], revealing challenges in aligning Bloom's cognitive levels with programming assessment tasks. This research sheds light on the complexities of using Bloom's Taxonomy in programming education and suggests a need for better alignment with learning objectives. [4], "Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome)" likely delves into the application and significance of the SOLO taxonomy in assessing the quality of learning outcomes. Authored by John B. Biggs and Kevin F. Collis, this book is likely to provide a comprehensive overview of the SOLO taxonomy, which categorizes students' learning outcomes into five levels: Pre-structural, Uni-structural, Multi-structural, Relational, and Extended Abstract. It probably explores how educators can use the SOLO taxonomy to evaluate the depth of students' understanding and the complexity of their learning outcomes across different subjects and educational contexts. By offering practical guidance and examples, the book likely serves as a valuable resource for educators seeking to enhance their assessment practices and promote higher-order thinking skills among students.

II. METHODOLOGY

Image Processing and OCR (Optical Character Recognition)

In our project, Image Processing and OCR (Optical Character Recognition) play a crucial role in converting JPEG images of student answers into analyzable text data. Initially, image processing techniques are applied to improve the clarity and readability of these images, addressing challenges such as varying handwriting styles, font sizes, and image quality. Following this, OCR technology is employed to accurately extract written content from the processed images. This extracted text is then transformed into a digital format, enabling further analysis and evaluation against the generative AI-generated model answers, based on Bloom's Taxonomy criteria.

Data Storage - CSV File

In the context of our project, after the OCR technology extracts text from JPEG images of student answers, the text data is segmented into individual words and stored in a CSV (Comma-Separated Values) file. This step is vital for organizing the extracted data in a structured, tabular format that facilitates easy access and analysis. The CSV file serves as a repository, allowing for efficient manipulation, analysis, and comparison of the textual data against predefined keywords and criteria, which are essential for the subsequent scoring and evaluation phase of the project.

Text and Data Analysis

In our project, the Text and Data Analysis phase involves a comprehensive examination of the textual content extracted from JPEG images. This step includes two primary analyses: first, a word count to determine the length of each student's answer, ensuring it meets the required depth and detail. Second, the extracted text is compared against a set of predefined keywords, which are integral to the question's learning objectives as outlined by Bloom's Taxonomy.

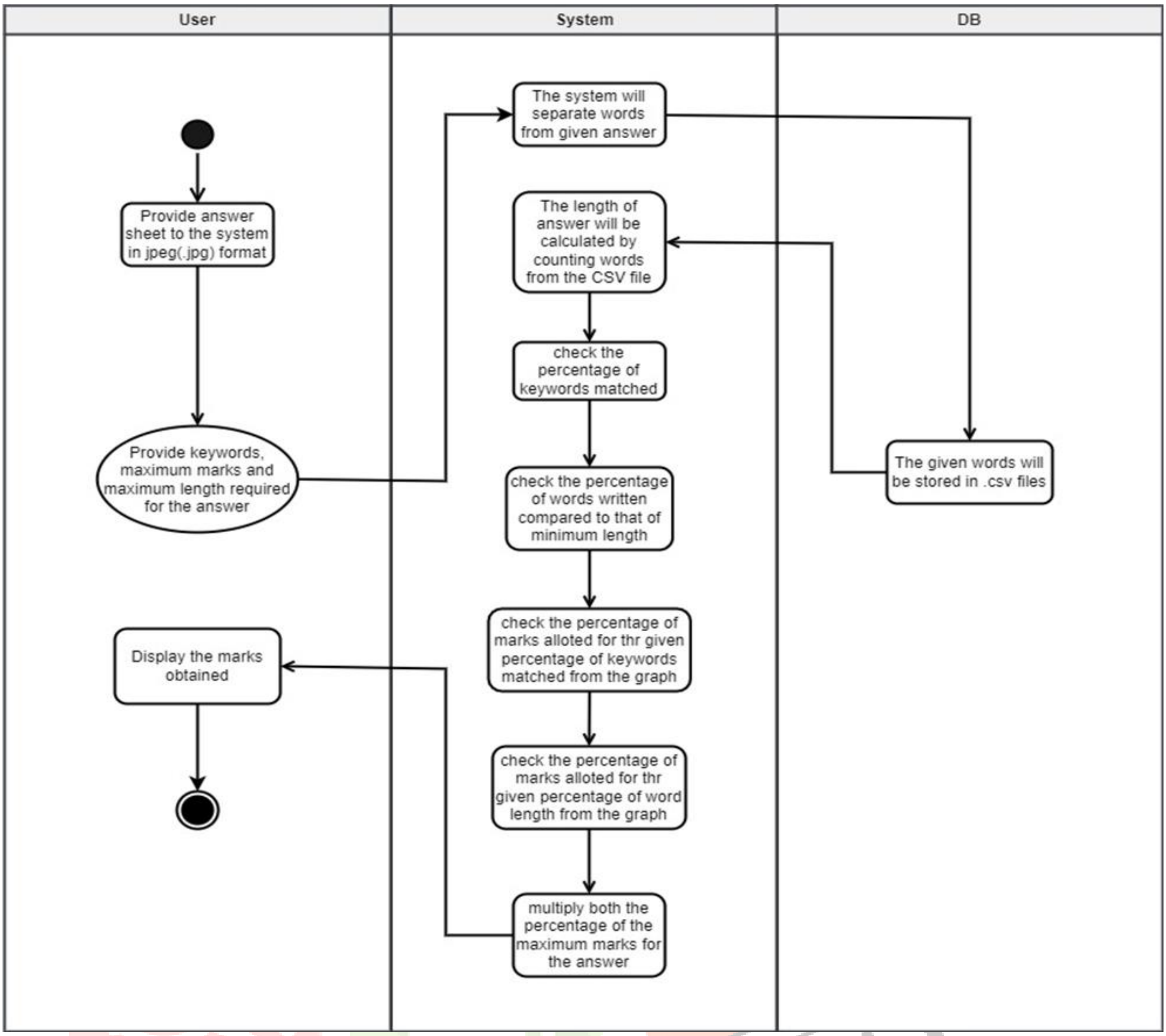

Fig. 1. Activity Diagram.

This dual approach allows for a nuanced evaluation of the answers, assessing not only the quantity but also the quality and relevance of the content provided by the students. Analysis encompasses:
- Counting words to assess answer length.
- Comparing extracted words against predefined keywords for keyword match percentage.

Scoring and Evaluation Criteria

The Scoring and Evaluation Criteria in our project are centered around two main factors: keyword matching and length evaluation. For keyword matching, the system calculates the percentage of predefined keywords present in the students' answers, reflecting the depth of coverage of key concepts. For length evaluation, the system assesses whether the answers meet a predefined minimum word count, ensuring the responses are sufficiently detailed. These criteria are quantitatively analyzed to assign a score, offering a balanced assessment of both the content's relevance and comprehensiveness.

Scoring is based on:
- Keyword Matching: Calculating the percentage of keywords found within the text.
- Length Evaluation: Ensuring the answer meets a minimum word count for detail adequacy.

Calculation of Final Score

In our project, the Calculation of the Final Score is a systematic process that combines scores from keyword matching and length evaluation. Each criterion is assigned a specific weightage based on its importance to the overall assessment objectives. The final score is calculated by aggregating these weighted scores, providing a comprehensive evaluation of each student's answer. This method ensures that the scoring reflects both the relevance of content to the question and the depth of the student's understanding, offering a fair and nuanced assessment of their performance.

Results Presentation

Marks are displayed with feedback on keyword usage and answer length, providing comprehensive insights. Software and System Integration The methodology integrates OCR software, data processing scripts, and user interfaces for results display, illustrating a multi-system approach.

III. Inputs And Evaulaution Criteria For Answer evaluation System

Answer Sheet in JPEG Format

- Primary input: Answer sheet in JPEG (.jpg) format containing text for evaluation.
- Used for text extraction via Optical Character Recognition (OCR).

Keywords
- Predefined list of keywords relevant to the subject matter of the answer.
- Evaluates content relevance and coverage of key concepts.

Maximum Marks

- Represents the highest score possible for the answer.
- Crucial for calculating the final score as a percentage of these marks.

Minimum Length Requirement

- Required minimum number of words in the answer for length criteria. Evaluation Criteria (Graphs or Algorithms for Scoring)
- Predefined criteria or algorithms for converting keyword match percentage and minimum length achievement into scores.

CSV File for Data Storage

- Intermediate storage format for words from extracted text for analysis. Software/System for OCR and Data Analysis
- Tools or software for performing OCR on the JPEG image and analyzing text.

IV. Process Description

Calculating Scores from JSON Files

1) Reads keywords, maximum marks (max_marks), and minimum length (min_length) from each JSON file.
2) Calculates the number of words matching the keywords by comparing words in the CSV file against the keywords from the JSON file.
3) Computes the keyword match percentage based on the ratio of matched keywords to the total number of keywords.
4) Calculates the final score for each criterion (like "remember" or "understand") as a percentage of the maximum marks, based on the keyword match percentage.

Creating Pie Charts

• Uses create_pie_chart function with inputs: keyword_matches, keyword_score, and analysis_type (like "Remember" or "Understand").
• Sets up a pie chart with segments representing the keyword match percentage, the score percentage, and the number of matches.
• Focuses on the visual representation of the calculated scores without performing complex calculations.

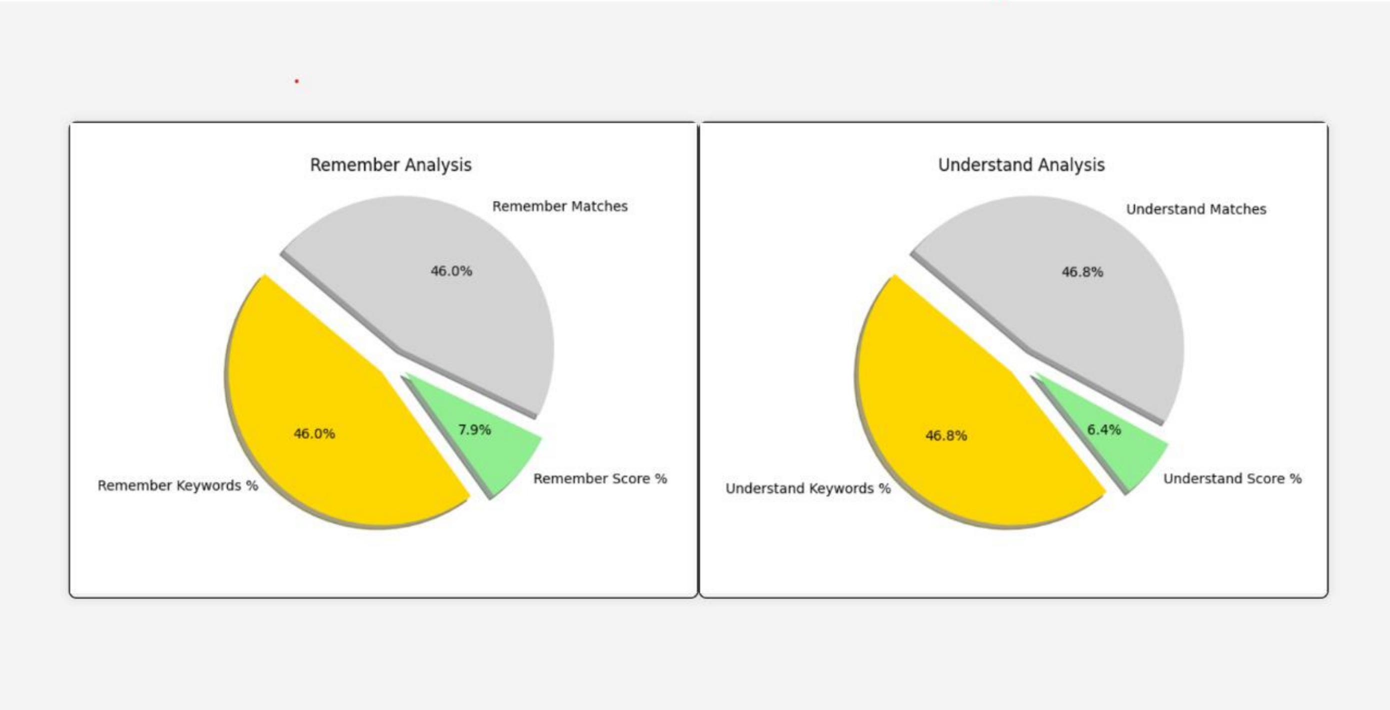

Fig. 2. Pie Chart.

Combining Scores for Final Output

• Suggests a comparison between scores from different JSON files for collective analysis or combination.
• The specific method for combining scores or final scoring is not detailed but implies aggregation for comprehensive evaluation.

FORMULAS USED

Calculating Keyword Match Percentage

The formula to calculate the keyword match percentage is given by:

Keyword Match Percentage = keyword matches/ len(keywords) × 100 (1)

where keyword_matches is the number of keywords found in the CSV file, and len(keywords) is the total number of keywords defined in the JSON file.

Calculating Final Score

Based on Keyword Matches The final score based on keyword matches is calculated using the formula:

Final Score = (keyword marks percentage /100) × max marks (2)

After determining the keyword match percentage (keyword_marks_percentage), this formula adjusts the score according to the maximum marks (max_marks) available.

Pie Chart Data Calculation

For pie charts, the segments represent:

- The keyword match percentage.
- The score percentage.
- The number of matches. The size of each segment is proportional to the values calculated, such as keyword_matches and keyword_score.

V. Acknowledgment

We, the authors, would like to express our sincere gratitude to everyone who contributed to the success of this project. This endeavor wouldn't have been possible without the support, guidance, and assistance of several individuals and organizations. First and foremost, we extend our heartfelt thanks to Prof. Hitesh Chaudhari for their invaluable guidance throughout the project. Their expertise, encouragement, and constructive feedback played a crucial role in shaping the project and enhancing its quality. Furthermore, we want to thank Sinhgad College of Engineering for providing the necessary resources and facilities that facilitated the smooth progress of the project. Finally, we appreciate the support from our families and friends who stood by us during the ups and downs of the project, offering encouragement and understanding. This project has been a rewarding journey, and we are thankful to everyone who contributed to its realization.

References

[1] L. Anderson, D. Krathwohl, P. Airasian, K. Cruikshank, R. Mayer, P. Pintrich, R. Raths, and M. Wittrock, A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Addison-Wesley Longman, 2001.

[2] R. Brito and M. Usman, "Bloom's taxonomy in software engineering education: A systematic mapping study," in Proceedings of the Frontiers in Education Conference (FIE 2015). IEEE Xplore, 2015.

[3] R. Gluga, J. Kay, R. Lister, S. Kleitman, and T. Lever, "Overconfidence and confusion in using bloom for programming fundamentals assessment," in Proceedings of the 43rd ACM Technical Symposium on Computer Science Education, 2012, pp. 147–152.

[4] J. Biggs and K. Collis, Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome). Academic Press, 1982.