International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Subjective Answer Sheet Evaluation

Shweta Kambare[1], Yashraj Pisal[2], Harish Pawar[3], Harshal Pendor[4], Ayushi Kamble[5], Srushti Kasurde[6]

[1,2,3,4,5,6]Department of Artificial Intelligence and
Data Science Engineering
Vishwakarma Institute of Technology
Pune, India
[1]shweta.kambare@vit.edu, [2]yashraj.pisal23@vit.edu,[3]harish.pawar23@vit.edu,
[4]harshal.pendor23@vit.edu, [5]ayushi.kamble23@vit.edu, [6]srushti.kasurde23@vit.edu

Abstract-

The personal answer sheet marking is required to increase the efficiency and accuracy in the educational systems. In this paper, we present a machine-based model with a primary element: text removal and marking. The text extension component appoints OCR through OpinCV and Tesseract, in addition to preprocessing methods such as gracekel conversion and sound eradication, overcomes such issues such as handwriting styles and noise layouts. The marked component uses the bias classification trained on the corpora labeled, and includes obscure -matching methods for the evaluation of the meaningful equality of the students' answers of the model's answers. Designed using backnd, flask and pirate base, supports real-time handling of data with educational platforms and easy integration. This system proposes a solid, skewable and user-friendly approach to mark the individual's answer in educational institutions.Keywords: fuzzy matching, Flask, Gaussian Naive Bayes, machine learning, subjective response evaluation, automation, semantic similarity, and educational technology.

Keywords: Subjective answer

evaluation, machine learning, Optical Character Recognition, Gaussian Naive Bayes, fuzzy matching, Flask, educational technology, automation, semantic similarity.

I. INTRODUCTION

Manually evaluating the personal answer sheets for educational institutions is a long -term hard -working and time -consuming process.With the increasing application of learning digital tools, automation of this process became a necessity, especially in the case of a large number of exams. Although automation is straightforward in terms of objective-type questions, subjective answers provide different challenges such as handwriting variations, meaningful processes, and the need to consider the variations in the words of the same ideas. To address these challenges requires an intelligent and dynamic solution with advanced image processes, machine learning and natural language processes.

This function presents a fully automatic, end-to-end system for automatic personal answer sheet grading based on two main modules: text extension and grading. Text Extraction Module uses optical character recognition by pyatsaract with preprocessing steps made by OpenCV. These include operations and thresholding, including grasscale thresholding, cleansing and noise removal, which significantly improves the quality of the text, such as system writing variations and confusion or complex answer sheets are able to easily handle the format.

Instead of evaluating responses according to the word matches, the grading system uses supervisory education, more precisely the Gausi Bhole Bias (Gausiyanab) classification. The alternative words and paragraph replies are used to ensure grade consistency in this system of obscure string or fziezi when recognizing efficiently.

Flask and pyrebase were used in the construction of Backonde to enable an intact connection with the academic platform. It is possible to create automatic grading and complete feedback by supporting the real-time uploading and processing of the scanned answer sheets. Due to its modular design, the system is extremely scaling and adjustable, which makes it suitable for the range of educational requirements and test formats.

II.LITERATURE REVIEW

According to many studies to enhance evaluation accuracy and efficiency, an automatic marking system has also been examined to implement personal responses to web-based tests, machine learning (ML) and natural language process (NLP).

•E. Johari etc. Designed a system "evaluation" that obtains personal answers through NLP and Cementic Learning, which offers automatic feedback to self -evaluate. It also has accessibility features like hand-to-speech, speech-to-text and special-enabled students to help. This system enhances efficiency by giving visual statistics to teachers and students, reducing evaluation stress and gaps in online exams.

• M. Bashir Att Al. Wordnet, Word 2 VEC, Word MoD distance (WMD), cosine equality, multinomyal Bhole Bice (MNB) and TF-IDF-using machine learning-based grading systems. Solution statements and keywords are classified with 88% points in the system with a 1.3% increase using MNB. The study states that descriptive answers WMD is better than the equality of the Cosme for grading.

• P. Patil at al. An automatic evaluation system developed that overcomes human transformation through tokenization, part-language tagging, lametization and cementic analysis. The system is divided into two parts: the first part scan the data from the documents and organizes it, and the second part enforces the ML and NLP technique for evaluation. The solution saves up to 90% of the evaluation time with good accuracy.

• G. Parashar at al. To automatically mark the written exam, a solution was proposed to scan the manuscript answers and mark the answers with the smart engine. The system provides points of 0 to 9, which then converts to percentage, reduces marked errors and saves up to 90% of the marker.

•K Sirsat It Al. An NLP-based grading system is proposed that pre-process the text and uses ML algorithms to analyze multi-focus descriptive responses. The goal is to improve the accuracy of grading, reduce teachers' workloads and increase total efficiency in university-level personal tests

III. METHODOLOGY

The system has been designed to transform the scanned answer sheets into the machine-reading text and to evaluate the answers using the **machine learning (ML)** and **Natural language processing (NLP)** technique. The process consists of two main modules: text extraction and grading system.

1. Text Extraction Module

This module is designed to convert scanned handwritten answer sheets into clean, digital text through three main phases:

Step 1: Image Preprocessing

Before extracting text, the scanned images undergo a few enhancement techniques to boost accuracy:

- **Grayscale Conversion**: Coloured images are converted into grayscale to reduce complexity and highlight the text more effectively.

- **Thresholding**: This method changes the grayscale image into a sharp black-and-white format, improving the contrast between the writing and the background.

- **Noise Removal**: Unwanted artifacts or marks in the image are cleaned using filters like median filtering, along with morphological techniques like erosion and dilation. These adjustments help make the text clearer and easier to extract.

Together, these steps help prepare a clean image so the text extraction process delivers more reliable results.

Step 2: Text Extraction Using OCR

Once the image is pre-processed, Optical Character Recognition (OCR) is applied to read the handwritten content. The tool used here is **Pytesseract**, which identifies characters, words, and lines, converting them into machine-readable digital text. While this process works well, challenges like diverse handwriting styles or complex layouts can sometimes affect the output.

Step 3: Text Segmentation

After OCR converts the image into raw text, the system organizes it into meaningful sections. These can include questions, answers, or paragraphs. Segmentation also helps remove unnecessary parts like headers or footers, and prepares the data for the grading stage.

2. Grading System Module

Once the text is extracted and organized, the grading module steps in to evaluate the student responses by comparing them with a standard (model) answer. This is achieved using machine learning and semantic similarity techniques.

Step 1: Feature Extraction

To make a fair comparison, both the model and the student answers are broken into smaller analyzable parts:

- **Tokenization**: The content is split into individual words or sentences.

- **Vectorization**: These tokens are then converted into numerical representations using methods such as TF-IDF or word embeddings. This allows the system to analyze text mathematically.

Step 2: Model Training with Gaussian Naive Bayes

The extracted features are used to train a machine learning model—in this case, **Gaussian Naive Bayes**. This algorithm helps the system learn how to evaluate and score answers based on similarities to the model response. The trained model then classifies new student responses, providing scores that reflect their semantic closeness to the ideal answer.

The grading system employs a **Gaussian Naive Bayes (Gaussian-NB) classifier** trained on labelled data consisting of **questions, model answers, and corresponding grades**. The model learns patterns from these examples, enabling it to predict grades for new responses based on extracted features.

Step 3: Semantic Similarity Evaluation (Fuzzy Matching)

Since students may phrase answers differently while conveying the same meaning, the system incorporates **fuzzy matching** using **Fuzzywuzzy** for approximate string matching. This technique computes a

similarity score between the student's answer and the model answer, accounting for **paraphrasing and alternative word choices**.

Step 4: Final Grading and Feedback Generation

The final score for each answer is calculated by combining both the results of the machine learning model and meaningful equality analysis. These two components are balanced using loaded scoring to ensure that the evaluation is correct and accurate for all students.

Once the score is finalized, the system automatically generates personalized feedback. This feedback only highlights the power and weakness of the answer, while guides the student to improve. The feedback includes:

- **Correctness**: Indicates how closely the student's answer aligns with the model or ideal response.
- **Clarity**: Notes whether the response is easy to understand and well-structured. If not, suggestions for better organization and expression are offered.
- **Improvement Suggestions**: Tips on how to write more clearly, manage time better, or enhance specific areas of the answer.

This smart grading process helps to maintain consistency in scoreing, saves trainers a significant time, and ensures that each student gets creative feedback depending on their performance.

Complete System Workflow

Here's an overview of how the system processes and evaluates a scanned answer sheet from start to finish:

1. **Input**: The system receives a scanned image of the student's answer sheet.
2. **Image Preprocessing**:
   - Converts the image to grayscale.
   - Applies thresholding to increase contrast.
   - Removes noise using techniques like filtering and morphological operations.

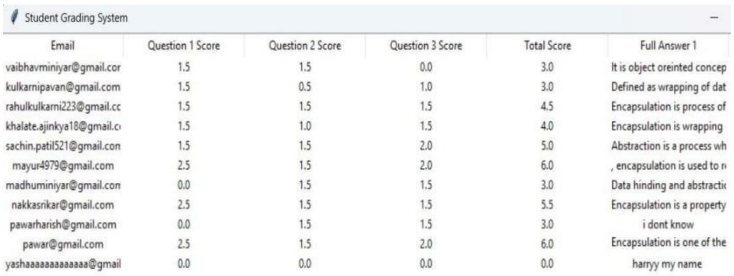
These steps improve the quality of the image for accurate text extraction.

3. **OCR (Optical Character Recognition)**:
   - The cleaned image is passed through an OCR engine (like Pytesseract).
   - This extracts raw handwritten text and converts it into a machine-readable format.
4. **Text Cleaning and Segmentation**:
   - The extracted text is segmented into logical units such as answers, questions, and paragraphs.
   - Unnecessary parts (like headers or page numbers) are removed.
5. **Grading Phase**:
   - **Feature Extraction**: Both the model and student answers are broken down through tokenization and vectorized into numerical formats using methods like TF-IDF or word embeddings.
   - **Machine Learning Evaluation**: A Gaussian Naive Bayes classifier is used to predict a score based on these features.
   - **Semantic Similarity Check**: Fuzzy matching is applied to assess how closely the student's response matches the model answer.

6. **Final Scoring and Feedback**:

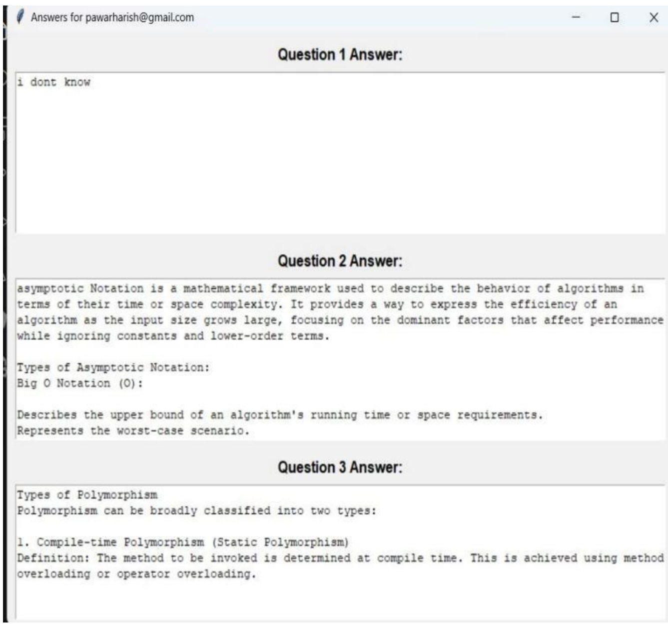
○ The scores from the ML model and semantic analysis are combined using a weighted average.

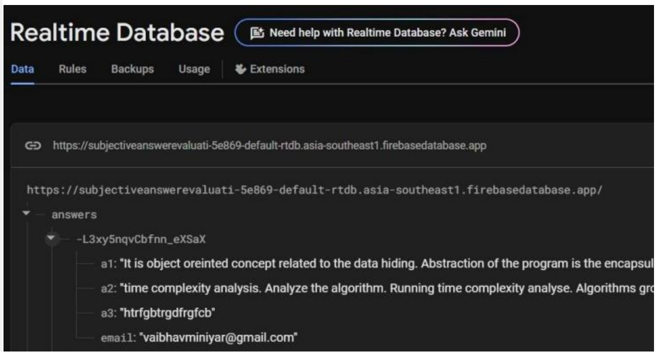
RESULTS:



Student Grading System

Email	Question 1 Score	Question 2 Score	Question 3 Score	Total Score	Full Answer 1
vaibhavminiyar@gmail.cor	1.5	1.5	0.0	3.0	It is object oreinted concep
kulkarnipavan@gmail.com	1.5	0.5	1.0	3.0	Defined as wrapping of dat
rahulkulkarni223@gmail.cc	1.5	1.5	1.5	4.5	Encapsulation is process of
khalate.ajinkya18@gmail.c	1.5	1.0	1.5	4.0	Encapsulation is wrapping
sachin.patil521@gmail.con	1.5	1.5	2.0	5.0	Abstraction is a process wh
mayur4979@gmail.com	2.5	1.5	2.0	6.0	, encapsulation is used to r
madhuminiyar@gmail.con	0.0	1.5	1.5	3.0	Data hinding and abstracti
nakkasrikar@gmail.com	2.5	1.5	1.5	5.5	Encapsulation is a property
pawarharish@gmail.com	0.0	1.5	1.5	3.0	i dont know
pawar@gmail.com	2.5	1.5	2.0	6.0	Encapsulation is one of the
yashaaaaaaaaaaaa@gmail	0.0	0.0	0.0	0.0	harryy my name





E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

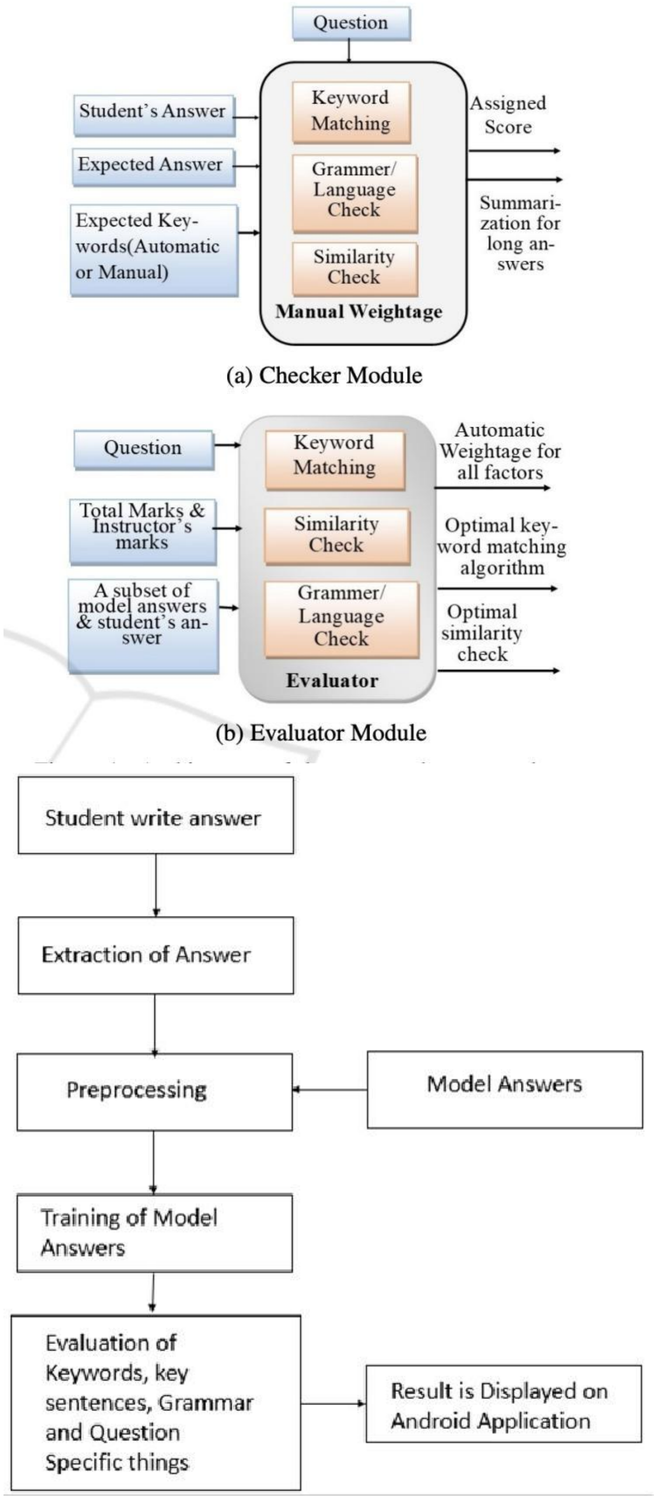


Fig. 1-Workflow diagram

V. CONCLUSION

The proposed machine to evaluate the personal answer sheets gives a promising solution to the general deficiency of manual grading such as learning-driven system inconsistencies and inefficiency. By combining effective text extraction methods, identifying the system manuscript response and evaluating their materials by using tools such as Pyatasract and OpenCV with intelligent grading systems such as Gausian Bhole Bias and Fluencies.

Prepoles and noise reduction, such as conversion, threshold and noise, plays an important role in improving OCR performance, especially when dealing with different handwriting styles and suboptimal scan quality. Aside from grading, the meaning of the system's meaningful equality analysis of the system recognizes the meaning of different phrases, when students use different writing styles, allowing proper evaluation.

The backnd architecture, made with flask and pyrebes enables real-time and scaleable evaluation, which makes it suitable for integration with current digital education systems. The experimental test shows the improved speed and efficiency, the ability to handle large batch of answer sheets at the same time.

This research underlines the possibility of automation in the personality of the personality, offers a system that not only accurate and consistent, but also reduces the burden of work on teachers. It is said that future developments can focus on refinement to handwriting inputs for bad written inputs, increase the ability of a more complex or long -term reaction system, and a wide dataset training to handle the wide types of the answer style.

Overall, the system shows significant progress in educational technology, paving the way for a faster, better and more efficient grading process.

REFERENCES

1. W. C. Smith and J. Holloway, "School testing satisfaction," Educational Assessment, Evaluation and Accountability, vol. 32, no. 4, pp. 461–479, 2020.
2. P. M. Nadkarni, L. Ohno-Machado, and W. W. Chapman, "Natural language processing: an introduction," Journal of the American Medical Informatics Association, vol. 18, no. 5, pp.544–551, 2011.
3. E. D. Liddy, "Natural language processing," 2001.
4. A. Rokade, B. Patil, S. Rajani, S. Revandkar and R. Shedge,"Automated Grading System using Natural Language Processing", Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies, 2018.
5. D.R. Tetali, G. Kiran Kumar and L. Ramana, "A Python Tool for Evaluation of Subjective Answers (APTESA)", International Journal of Mechanical Engineering and Technology (IJMET), vol. 8, pp.

247– 255, July 2017.

6. M. Syamala Devi and H. Mittal,
"Machine Learning techniques with Ontology for subjective answer evaluation", International Journal on Natural Language Computing (IJNLC), vol. 5, April 2016.
7. C. Roy and C. Chaudhari, "Case-Based Modeling of answer points to expedite semi-automated evaluation of subjective papers",IEEE 8th
International Advance Computing Conference (IACC), pp. 85–90, 2018.
8. K. Surya, E. Gayakwad, Nallakaruppan and M.K.,
"Deep learning for
Short Answer Scoring", International Journal of Recent Technology and Engineering (IJRTE), vol. 7, issue 6, 2277–3878, 2019.
9. D. Cer, Y. Yang, S.Y. Kong, N. Hua, N. Limtiaco, R.S. John, N.
Constant, M. Guajardo- Cespedes, S. Yuan, c. Tar, Y.H. Sung, B.
Strope and R. Kurzweil, "Universal Sentence Encoder", 2018.
10. A. Dhokrat, R.G. Hanumant and C. Mahender,
"Automated Answering for Subjective Examination", International Journal of Computer Applications, vol. 56, pp. 14–17, 2012.
11. S.A. Saany, A. Mamat, A. Mustapha, L.S. Affendey and M.N.A.
Rahman,
Semantics question analysis model for question answering system", Applied Mathematical