PAPER EVALUATION USING ARTIFICIAL INTELLIGENCE

#1 K .UDAY KIRAN, #2 A.KRISHNAPRIYA

#1 ASSISTANT PROFESSOR #2 MCA SCHOLAR
DEPARTMENT OF MASTER OF COMPUTER
APPLICATIONS QIS COLLEGE OF ENGINEERING
& TECHNOLOGY , ONGOLE
. VENGAMUKKALAPALEM(V), ONGOLE, PRAKASAM DISTRICT., ANDHRAPRADESH

Abstraction

The traditional method of evaluating academic answer sheets, especially descriptive or subjective ones, is a manual and labor-intensive task that often suffers from inconsistencies, human bias, and delays in providing results. As education continues to evolve with the integration of digital tools, there is a critical need for an intelligent, efficient, and scalable solution to automate the paper evaluation process. This project proposes an AI-based system that leverages cutting-edge technologies in Natural Language Processing (NLP), machine learning, and semantic analysis to assess written student responses against pre-defined model answers. The system is designed to understand the context, grammar, coherence, vocabulary usage, sentence structure, and semantic similarity between the student's answer and the ideal key. It also incorporates spelling and grammar correction modules and can offer feedback to enhance learning. The AI model is trained on a large dataset of evaluated papers to understand scoring patterns and apply them uniformly across new submissions. This eliminates subjectivity, reduces evaluation time, and ensures fairness in assessment. Furthermore, the system can be adapted for multiple languages, subjects, and

Academic levels, making it highly versatile and beneficial, for schools, universities, online examination platforms, and certification bodies.

The implementation of this solution aims to revolutionize academic assessment, making it more efficient, transparent, and student-friendly.

Introduction

Assessment is a fundamental component of any educational system, designed to measure student understanding, learning progress, and overall academic performance. Traditionally, this evaluation process, particularly for descriptive or long-answer questions, has been carried out manually by human examiners. While human evaluation allows for contextual judgment and comprehension, it is highly time-consuming, labor-intensive, and prone to subjective bias. Factors such as fatigue, personal perceptions, and inconsistency in applying rubrics can lead to inaccuracies in grading, which can affect a student's academic growth and morale.

In recent years, the integration of technology into education has paved the way for the digitization of assessments. However, the majority of automated grading tools focus primarily on objective question types like multiple-choice, fill-in-the-blanks, or true/false, where predefined answers can

be easily matched. Descriptive answer evaluation still lags behind due to its complexity in understanding language, context, and the semantic correctness of responses. With the advancement in Artificial Intelligence, particularly Natural Language Processing (NLP) and machine learning, it is now possible to design systems that can comprehend human language, evaluate textual content, and generate human-like feedback and scores.

This project aims to build an AI-driven paper evaluation system that can analyze and score subjective answers by comparing them with model responses and applying context-aware techniques. The system uses NLP to understand the meaning behind student answers and applies machine learning algorithms trained on human-evaluated scripts to mimic the scoring process. By doing so, it ensures consistent and fair evaluation while significantly reducing the turnaround time for result declaration. This intelligent evaluation system not only helps educational institutions streamline their assessment workflows but also empowers students by offering instant, personalized feedback that enhances their learning experience.

Literature Survey

Title:*Automated Essay Scoring with e-rater® V.2*

Author: Jill Burstein, Daniel Marcu, Kevin Kukich

Description: This paper presents the e-rater system developed by ETS for automatessay scoring using natural language processing techniques. The system evaluates essays based on syntactic variety, discourse structure, and topical content. The study shows that automated evaluation can yield scores comparable to human raters in many cases.

Title:*Using Machine Learning Algorithms for Automatic Assessment of Short Answers* **Author:** Maria Wolska and Manfred Pinkal **Description:** The authors explore supervised machine learning algorithms to assess short student responses. The study demonstrates how a feature-based model using text similarity, grammar patterns, and keyword presence can improve the accuracy and reliability of automatic grading systems.

Title:*Automatic Short Answer Grading System (ASAGS)*

Author: Rakesh Agrawal, Piyush Arora **Description:** This system leverages Natural Language Processing to assess short and descriptive answers. It compares student responses with model answers using semantic similarity and key concepts. It concludes that automatic systems can significantly reduce the human workload while maintaining scoring accuracy.

Title: *A Neural Network-Based Automated Essay Scoring System*

Author: Taghipour, K., & Ng, H. T. **Description:** This study introduces a deep

learning approach using recurrent neural networks for essay scoring. The system captures both the surface and semantic features of the essays, producing results close to human raters. It highlights the importance of using deep learning in capturing writing style and context.

Title: *Automated Essay Grading using Machine Learning*
Author: Akshay Kulkarni, Vinayak Agham
Description: This research applies decision trees, SVMs, and logistic regression for evaluating student essays. The study emphasizes the use of linguistic and syntactic features and concludes that machine learning can be an efficient tool for scalable and unbiased paper evaluation.

Title: *A Review on Automated Essay Grading System using NLP and ML Techniques* **Author:** Dr. Smita Nirkhi, Anjali Jain **Description:** This review paper provides a comparative analysis of various NLP and machine learning techniques used in essay grading systems. It outlines the strengths and limitations of traditional and neural models and emphasizes the need for hybrid approaches.

Title: *Semantic Similarity Measures Applied to Automatic Text Grading*
Author: Daniel Bär, Torsten Zesch, Iryna Gurevych
Description: This work explores semantic similarity algorithms like LSA and WordNet in evaluating text similarity between student responses and model answers. It shows that semantic similarity can be a robust indicator for fair evaluation.

title: Automated Essay Scoring Using Machine Learning Techniques

Authors: EllisB. gePa
Description: Introduced Project Essay Grade (PEG), one of the earliest AI-based essay scoring systems. Used statistical features like word length, sentence complexity, and grammar indicators to predict human-like scores.

System Analysis Existing System

The current methods of paper evaluation, particularly for subjective and descriptive answers, are predominantly manual. Teachers and evaluators read each answer script and assign marks based on their understanding and experience. While this process allows human interpretation and comprehension, it is inherently time-consuming and lacks standardization. Evaluators may interpret answers differently based on their mood, background knowledge, or fatigue, resulting in inconsistent grading. In large-scale examinations, this becomes a major challenge as it increases both evaluation time and the chances of error.

Some semi-automated systems exist, but they are primarily used for evaluating objective-type questions such as multiple-choice or fill-in-the-blanks. Optical Mark Recognition (OMR) systems and digital test platforms are limited to pattern recognition and do not extend to evaluating content-rich answers. Even in digital assessment platforms, the descriptive answers are often manually

reviewed by human assessors. These systems, while reducing paper usage, still fail to address the deeper problem of scalability and unbiased evaluation when it comes to subjective questions.

Furthermore, existing systems do not offer real-time feedback or learning insights to students. Once scores are released, students typically do not receive specific guidance on where they went wrong or how to improve. This restricts the scope of formative assessment and limits opportunities for academic growth. There is also a lack of adaptability in existing systems to accommodate different languages, educational levels, and contextual relevance of answers. Thus, there is a significant need for an intelligent, automated, and scalable solution that can evaluate descriptive answers efficiently, fairly, and accurately.

Disadvantages of Existing Systems: Time-Consuming and Labor-Intensive:
Manual evaluation of descriptive answers requires a significant amount of time and effort, especially in large-scale academic settings. This delays result declaration and adds pressure on educators.

Subjectivity and Human Bias: Scores assigned by human evaluators can vary based on individual judgment, fatigue, and interpretation. This subjectivity can lead to inconsistent grading and unfair evaluations.

Lack of Scalability: Manual grading systems struggle to scale with increasing student populations. Institutions with limited staff face difficulties in maintaining quality and speed of evaluations during peak examination periods.

Limited to Objective Question Evaluation: Most existing automated tools are restricted to multiple-choice or objective questions. They are incapable of understanding context or evaluating the semantics of descriptive answers.

No Instant Feedback: Current systems do not provide students with immediate feedback or improvement suggestions, limiting the opportunity for learning and growth based on their performance.

Inflexibility across Subjects and Languages: Traditional systems do not adapt easily to multiple subjects or different languages, reducing their effectiveness in diverse educational environments.

Proposed System

The proposed system introduces an **AI-powered paper evaluation platform** designed to assess descriptive and subjective answers with high accuracy, consistency, and speed. By leveraging Natural Language Processing (NLP) and Machine Learning (ML) techniques, the system analyzes student answers in terms of grammar, content relevance, structure, coherence, and semantic similarity to model answers. This ensures that each response is graded not just based on

keyword matching, but on a deeper understanding of meaning and intent, much like a human evaluator.

The system is trained on a large dataset of previously evaluated answer scripts to learn patterns in scoring and acceptable variations in student responses. Advanced models such as **Transformers, BERT**, or **GPT-based architectures** are used to process natural language and understand context, enabling the system to handle a wide variety of subjects and question types. The AI can also detect vague answers, off-topic content, and even writing issues, offering both scores and constructive feedback to help students improve their performance over time.

Moreover, the proposed system supports **real-time and scalable evaluation**, making it suitable for schools, universities, and online education platforms. It includes a user-friendly interface for both evaluators and students, with dashboards to visualize scores, analytics, and performance trends. The system is designed to be language-independent, with multilingual support, and can adapt to institution-specific grading rubrics. This AI-driven approach aims to revolutionize academic assessments by providing **fast, fair, and feedback-rich evaluations**.

Advantages of the Proposed System:

Automated and Time-Efficient: The AI system significantly reduces the time required for evaluating descriptive answers by automating the process, allowing institutions to deliver results faster and manage large volumes of scripts efficiently.

Consistent and Unbiased Evaluation: Unlike human evaluators, the AI model provides standardized grading by eliminating subjective bias, fatigue, or inconsistency, ensuring every student is assessed fairly and uniformly.

Contextual Understanding of Answers: With advanced NLP capabilities, the system can understand the semantic meaning of student responses, not just keywords, making it suitable for evaluating complex, context-rich answers across subjects.

Real-Time Feedback and Performance Insights: Students receive instant feedback, detailed analysis, and suggestions for improvement, promoting continuous learning and helping them focus on their weak areas.

Scalable and Cost-Effective: The system can evaluate thousands of answer scripts simultaneously, making it highly scalable and cost-effective for institutions, especially during peak exam seasons.

Multilingual and Domain-Adaptable: The platform supports multiple languages and can be trained to handle different academic disciplines and question formats, ensuring flexibility and broader applicability.

Data-Driven Decision Making: With integrated analytics, educators can track class-wide performance trends, identify common errors, and tailor their teaching methods to better address learning gaps.

Implementation

The implementation of the Paper Evaluation System using Artificial Intelligence focuses on automatically analyzing and evaluating answer sheets, assignments, examination papers, and descriptive responses using AI, Machine Learning, and Natural Language Processing (NLP) techniques. The system helps educational institutions reduce manual evaluation effort and improve accuracy, consistency, and speed in assessment processes.

The proposed system can evaluate objective and subjective answers intelligently and generate marks, feedback, and performance analysis automatically.

1. Data Collection

The first stage involves collecting educational and examination-related data from schools, colleges, universities, and online learning platforms.

The collected dataset may include:

- Student Answer Sheets
- Question Papers
- Model Answers
- Marks Distribution
- Subject Information
- Student Performance Records
- Assignment Submissions
- Handwritten or Typed Responses

These datasets help train the AI model for automated paper evaluation.

2. Data Digitization and Input Processing

Student answer sheets are converted into digital format using:

- Scanners
- Mobile Camera Input
- PDF Uploads
- Online Examination Portals

For handwritten answers, Optical Character Recognition (OCR) techniques are used to extract text from images.

OCR Tools Used

- Tesseract OCR
- Google Vision API
- OpenCV-based OCR systems

This converts handwritten or printed text into machine-readable format.

3. Data Preprocessing

The extracted textual data is cleaned and prepared before evaluation.

Preprocessing Steps

- Noise removal
- Spelling correction
- Tokenization
- Stop-word removal
- Lemmatization and stemming
- Sentence segmentation
- Grammar normalization

This improves the quality of textual analysis.

4. Feature Extraction
Important textual and semantic features are extracted from student answers.

Features Used
Linguistic Features

- Grammar correctness
- Vocabulary usage
- Sentence structure Semantic Features

- Meaning similarity with model answers
- Concept coverage
- Context understanding Statistical Features

- Keyword frequency
- TF-IDF values
- N-gram analysis

Feature extraction improves answer evaluation accuracy.

5. Artificial Intelligence Model Development

AI and Machine Learning models are used to evaluate answers automatically.

AI Techniques Used
Natural Language Processing (NLP)

Used for understanding textual answers and semantic similarity.

Machine Learning Algorithms

Used for answer classification and scoring.

Deep Learning Models
Used for advanced contextual understanding.

Transformer Models

Used for semantic analysis and intelligent grading.

Examples:

- BERT
- GPT-based models
- LSTM Networks

6. Answer Similarity Analysis

The system compares student answers with model answers using semantic similarity techniques.

Similarity Methods

- Cosine Similarity
- Word Embeddings
- Sentence Transformers
- Semantic Matching Algorithms The AI

model evaluates:

- Correctness of concepts
- Relevance of answer
- Depth of explanation
- Writing quality

Methodology

The methodology of the proposed Paper Evaluation System follows an Artificial Intelligence and NLP-based automated educational assessment approach.

Step 1: Problem Identification

Traditional paper evaluation methods are time-consuming, inconsistent, and prone to human errors. Manual assessment may also lead to bias and delayed result processing. The proposed system aims to automate paper evaluation using Artificial Intelligence techniques.

Step 2: Requirement Analysis

The following requirements are analyzed:

- Examination dataset requirements
- OCR and text extraction requirements
- NLP processing requirements
- Automated grading requirements
- Educational analytics requirements

Step 3: Dataset Preparation

Student answer sheets, model answers, and grading records are collected and divided into:

- Training Dataset
- Validation Dataset
- Testing Dataset

Relevant educational attributes are selected for analysis.

Step 4: OCR and Text Processing The methodology includes:

1. Scan answer sheets
2. Extract textual content using OCR
3. Clean and preprocess text
4. Normalize grammar and sentence structure

This prepares answers for AI evaluation.

Step 5: AI-Based Evaluation Implementation The AI evaluation workflow includes:

1. Input student answers
2. Extract semantic and linguistic features
3. Compare answers with model solutions
4. Analyze content relevance and correctness
5. Generate marks automatically
6. Provide feedback and performance reports

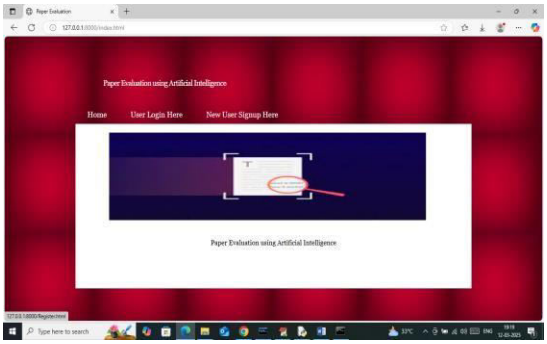
Step 6: Performance Evaluation The system is evaluated based on:

- Evaluation accuracy
- Semantic understanding capability
- Consistency of grading
- Processing speed
- Student feedback quality

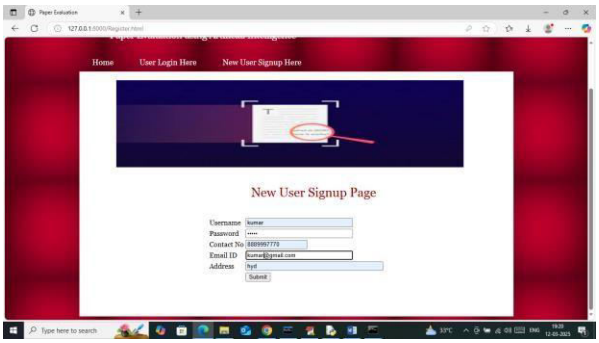
Technologies Used

- Python
- Artificial Intelligence
- Natural Language Processing (NLP)
- Machine Learning & Deep Learning
- TensorFlow / PyTorch
- OpenCV
- OCR Tools (Tesseract)
- Scikit-learn
- Flask / Django
- MySQL / MongoDB

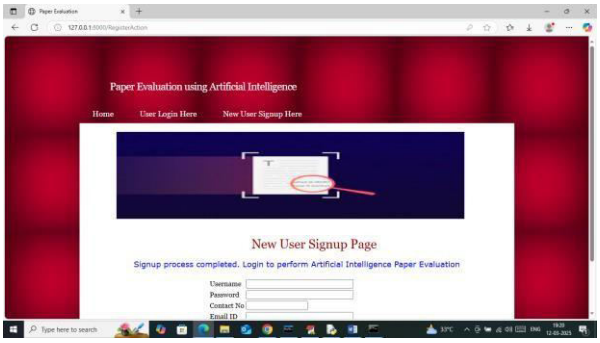
Results



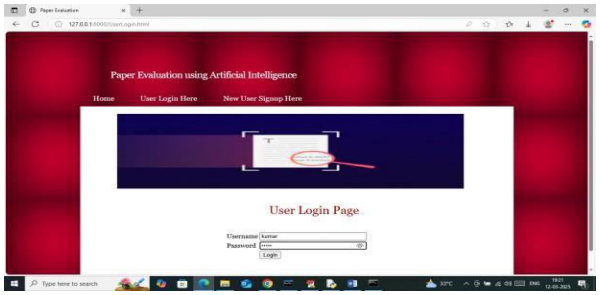
This image shows the homepage of a **Paper Evaluation System using Artificial Intelligence**. The web application provides features such as **Home**, **User Login**, and **New User Signup**. The interface has a red-themed background with a central panel displaying a sample paper evaluation area. The system is designed to automate and improve the process of evaluating examination papers using AI technology, making assessment faster, more accurate, and efficient.



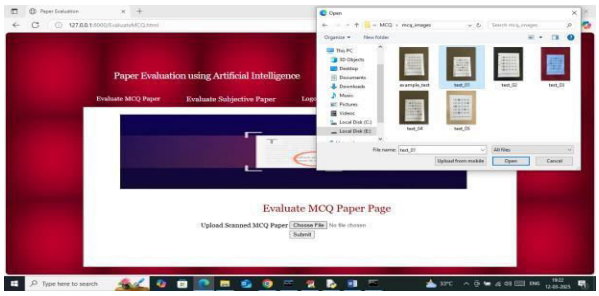
This image shows the **New User Signup Page** of the **Paper Evaluation System using Artificial Intelligence**. The page allows new users to create an account by entering details such as **Username, Password, Contact Number, Email ID, and Address**. A **Submit** button is provided to complete the registration process. The signup page serves as the entry point for users to access the AI-based paper evaluation system and its features.



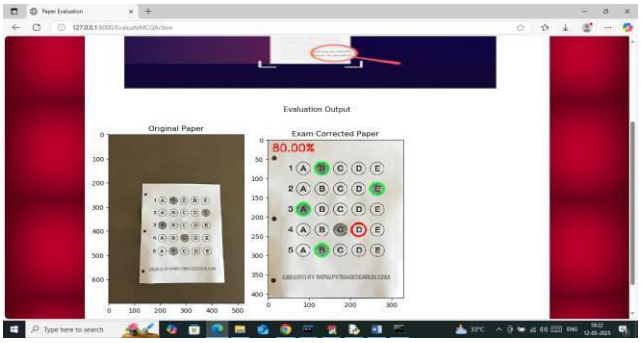
This image displays the **New User Signup Page** of the **Paper Evaluation using Artificial Intelligence** system after successful registration. A confirmation message, **"Signup process completed. Login to perform Artificial Intelligence Paper Evaluation,"** is shown to inform the user that account creation has been completed successfully. The page includes fields for Username, Password, Contact Number, Email ID, and Address, along with navigation options such as Home, User Login, and New User Signup. This page guides users to proceed to the login section and access the AI-based paper evaluation services.



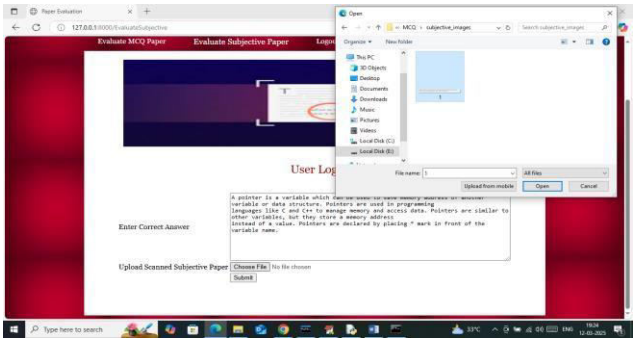
"This page is used for user authentication. Users enter their username and password to log in and access the AI-based paper evaluation system securely."



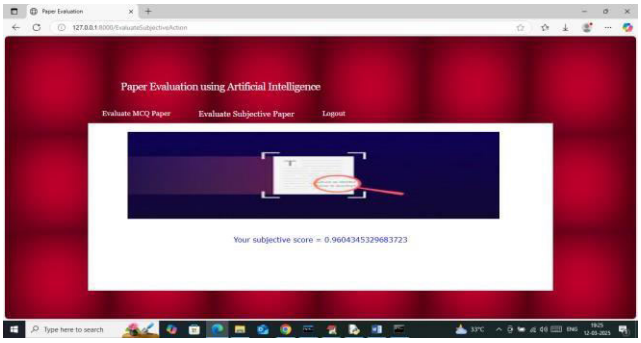
"This page appears after the user logs in successfully. It provides options to evaluate MCQ papers and subjective papers using Artificial Intelligence. The user can select the required evaluation type and view the results generated by the system. It also includes a logout option for secure session management."



In above screen first image is the uploaded exam image and second image is the correction image along with marks where green circle represents correct answer and red circle represents incorrect answer given by student. Similarly you can upload and test other papers. Now click on 'Evaluate Subjective Paper' link to get below page.



In above screen in first text area entered the correct subjective answer and then upload student subjective scanned image and then press button to get below page



In above screen student got 96% score and similarly you can checked with any image.

References

1. Zhang, Y., & Wallace, B. C. (2015). A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. *arXiv preprint arXiv:1510.03820*.

2. Devlin, J., Chang, M. W., Lee, K., & Toutanova,
3. K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL-HLT*.

4. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 5998-6008.

5. Shermis, M. D., & Burstein, J. (2013). *Automated essay scoring: A cross-disciplinary perspective*. Routledge.

6. Alikaniotis, D., Yannakoudakis, H., & Rei,
7. M. (2016). Automatic text scoring using neural networks. *Proceedings of the 54thAnnual Meeting of the Association for Computational Linguistics*.

8. Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater® V.2. *The Journal of Technology, Learning and Assessment*, 4(3).

AUTHORS PROFILE

Mr. K. Uday Kiran is an Assistant Professor in the Department of Master of Computer Applications at QIS college of Engineering gole, Andhra Pradesh. He earned his Master of Computer Applications (MCA) from Bapatla Engineering College, Bapatla. His research interests include Machine Learning, Programming Languages. He is committed to advancing research and fostering innovation while mentoring students to excel in both academic and professional pursuits.

STUDENT PROFILE

Mrs.A.KRISHNAPRIYAis a postgraduate student pursuing a MCA in the Department of computer Applications at QIS College of Engineering& Technology, Ongole autonomous college in prakasam dist. She completed undergraduate degree in MPCs (computer science) from ANU. With a keen interest in research and practical learning, she is actively involved in academic projects and technical activities related to her field