

Automated Answer Evaluation Based on Deep Learning

Mrs.S.Mahalakshmi S M.E,
Assistant Professor,
Department of Computer Science And
Engineering,
Shree Venkateshwara HiTech Engineering
College,
Gobichettipalayam.
E-mail:mahalakshmitelnet@gmail.com

Ms.R.Dhanya,
Student,
Department of Computer Science And
Engineering,
Shree Venkateshwara Hi Tech Engineering
College,
Gobichettipalayam.
E-mail:csedhanyar@gmail.com

Ms.B.Maheswari,
Student,
Department of Computer Science And
Engineering,
Shree Venkateshwara Hi Tech Engineering
College,
Gobichettipalayam.
E-mail:mahesbala0121@gmail.com

Ms.A.Nancy,
Student,
Department of Computer Science And Engineering,
Shree Venkateshwara Hi Tech Engineering
College,
Gobichettipalayam.
E-mail:nancyannadurai@gamil.com

ABSTRACT-Automated answer evaluation based on deep learning is a novel approach to stream line the grading process in educational settings .This project aims to develop a system capable of assessing students' answers to open-ended questions accurately and efficiently. Leveraging deep learning techniques, such as natural language processing and neural networks, the system will analyze text responses, identify key concepts, and assign scores based on predefined criteria. By automating the evaluation process, educators can save time, provide more timely feedback to students, and

ensure consistency in grading. This paper discusses the architecture, implementation, and evaluation of the automated answer evaluation system, highlighting its potential to revolutionize the assessment process in education.

KEYWORDS- Automated assessment, Deep learning, Natural language processing, Neural networks, Answer grading, Educational technology, Machine learning, Evaluation system, Open-ended questions, Educational assessmen

I.INTRODUCTION

Optical character recognition (OCR) is the process of identifying individual letters and words within a digital image. It involves employing a classification algorithm to analyze each character and assemble them into coherent words. This method relies on algorithms that group similar words together, comparing the outcome with the expected text in the image. OCR technology converts images of printed, typed, or handwritten text into machine-encoded text, facilitating its extraction and manipulation by computers. For businesses, OCR is invaluable for quickly extracting and transforming text from scanned documents and images into a readable, editable, and searchable format. It enables computers to interpret written language much like the brain and eyes work together to comprehend text from images. Despite initial challenges, automated OCR systems like Tesseract have become widely

available for integration into various programs, enhancing their functionality. Meanwhile, the traditional method of manually grading subjective responses, common in educational assessments, is increasingly impractical, particularly in the current remote working environment exacerbated by the pandemic. Although automated systems effectively evaluate multiple-choice or objective questions, they fall short in assessing subjective responses, which are crucial for understanding a student's comprehension and depth of understanding. Manual evaluation of such responses is time-consuming and labor-intensive, leading to inconsistencies in scoring and potentially affecting student performance. To address this issue, machine learning techniques, particularly deep learning algorithms, are being explored to automate the grading process, particularly in areas like natural language processing and image recognition. These algorithms can discern intricate patterns and correlations within data, making them well-suited

for assessing subjective responses and providing accurate grades efficiently.

A.RECURRENT NEURAL NETWORKS (RNNs)

Recurrent Neural Networks (RNNs) are a class of neural networks particularly well-suited for sequential data processing, making them a popular choice for tasks like automated answer evaluation. Unlike feedforward neural networks, which process input data independently, RNNs have connections that form a directed cycle, allowing them to maintain internal state and capture dependencies across time steps. In the context of answer evaluation, RNNs can analyse textual answers word by word, considering the sequential nature of language. This enables them to capture the contextual information and dependencies present in the answers. For example, when evaluating a paragraph-long response to a question, an RNN can analyse each word in the response while retaining information from previous words to understand the overall meaning and coherence. One common variant of RNNs is the Long Short-Term Memory (LSTM) network, which addresses the vanishing gradient problem encountered during training. LSTMs have mechanisms called gates that regulate the flow of information, allowing them to capture long-range dependencies more effectively than traditional RNNs. In automated answer evaluation systems, RNNs can be trained on labelled datasets containing pairs of student answers and corresponding scores. By learning from these examples, the RNNs can develop the ability to predict scores for new answers based on their content and structure. Additionally, RNNs can be integrated with other components such as attention mechanisms to focus on relevant parts of the answer during evaluation, further enhancing their performance. Overall, RNNs offer a powerful framework for automated answer evaluation by leveraging the sequential nature of textual data.

II.EXISTING SYSTEM

Automated answer evaluation systems leveraging deep learning undergo a multi-stage process. Initially, textual answers are preprocessed, involving tasks such as cleaning and tokenization to ensure data uniformity and readability. Subsequently, feature extraction is employed to convert the text into numerical representations, often utilizing methodologies like word embeddings to capture semantic meanings effectively. Following this, the model training phase commences, wherein deep learning architectures such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), or transformer models are deployed to discern intricate patterns within the textual data. These architectures are trained on large datasets to learn how to accurately assess the quality of answers. Finally, the system

undergoes evaluation, where predicted scores are compared against ground truth labels to gauge the system's efficacy.

III DATA PREPROCESSING MODULE:

The Data Preprocessing Module is responsible for preparing the input data for the automated answer evaluation system. It includes the following tasks:

Text Cleaning and Normalization: Remove special characters, punctuation, and non-alphanumeric characters from the text. Convert text to lowercase to ensure consistency. Remove extra whitespaces and trim the text. **Tokenization:** Split the text into individual words or sub words. Create tokens to represent each word or sub word separately. **Padding or Truncation:** Ensure all input sequences have the same length. Pad shorter sequences with zeros or truncate longer sequences. **Word Embedding:** Convert words into dense vectors using techniques like Word2Vec or Glove. Use pre-trained embeddings to capture semantic relationships between words. **Handling Out-of-Vocabulary Words:** Replace out-of-vocabulary words with a special token or unknown token.

IV EVALUATION AND SCORING MODULE

Comparison: Compare student responses with model answers using similarity metrics such as cosine similarity or Levenshtein distance. **Scoring:** Assign scores or labels to student responses based on their similarity to model answers. **Evaluation Metrics:** Calculate evaluation metrics such as accuracy, precision, recall, and F1-score to measure the performance of the automated evaluation system.

V ABBREVIATIONS AND ACRONYMS

LSTM- LONG SHORT TERM MEMORY

AI -ARTIFICIAL INTELLIGENCE

SVM- SUPPORT VECTOR MACHINE

HMM- HIDDEN MARKOV MODEL

GMM- GAUSSIAN MIXTURE MODELS

RPN- REGION PROPORTIONAL NETWORK

UCF-UNIVERSITY OF CENTRAL FLORIDA

CSS- CASCADING STYLE SHEETS

HTTP-HYPertext TRANSFER PROTOCOL

VI CONCLUSION AND FUTURE ENHANCEMENT

An automated evaluation system plays a crucial role in assessing open-ended inquiries like short responses and essays, providing several advantages such as simplifying manual grading processes, saving time, effort, and resources, and ensuring fairness in evaluating students' responses. This paper introduces the concept of Automatic Arabic Short Answer Grading, utilizing Latent Semantic Analysis (LSA), a commonly used corpus-based similarity method. The system is applied to AR-ASAG, a limited Arabic dataset available to the public. Two experiments are conducted, employing different strategies for weighting data representation: local weighting and a hybrid approach that combines local and global weighting.

VII FIGURES AND TABLES

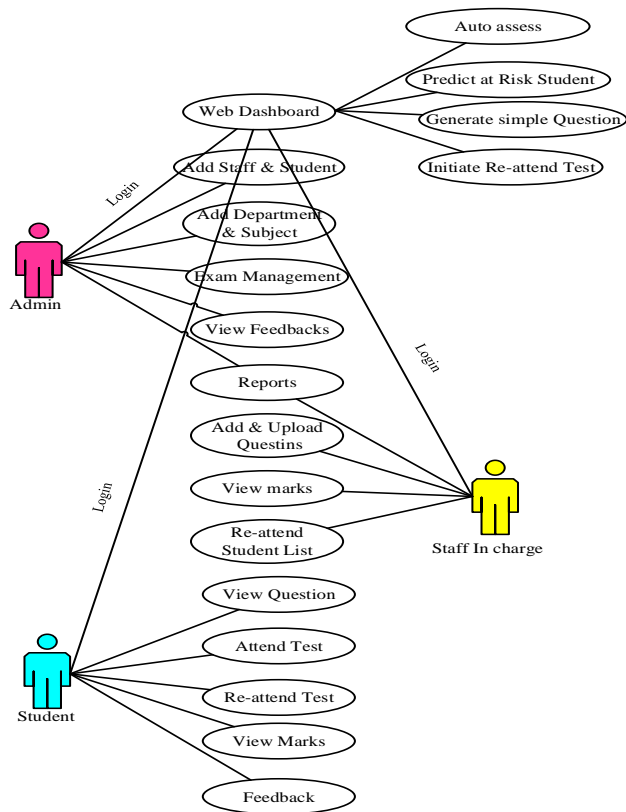


FIG NO:1 USECASE DIAGRAM

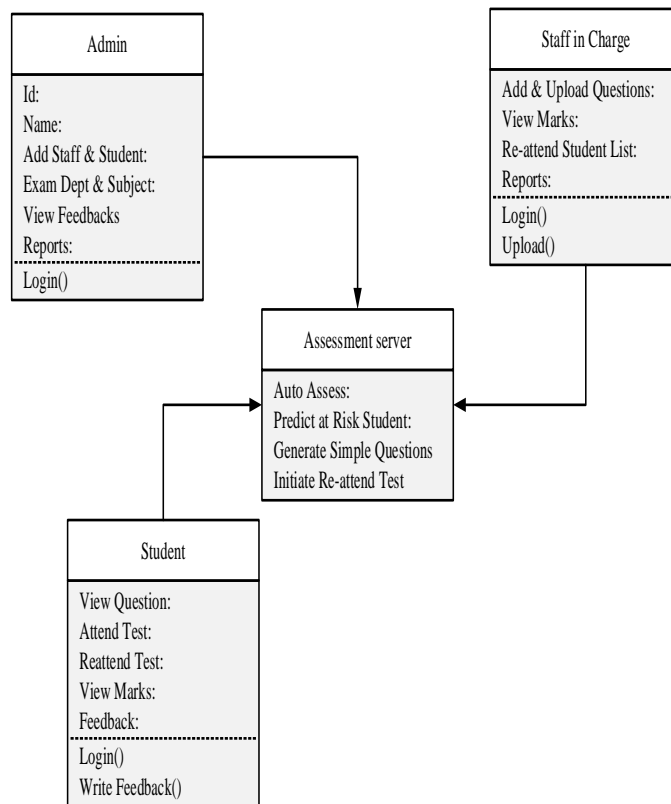


FIG NO:2 CLASS DIAGRAM

VIII ACKNOWLEDGEMENT

We are very grateful to Dr.T. SENTHIL PRAKASH ME PhD, Professor and Head Department of Computer Science and Engineering, for the aspiring suggestion, invaluable constructive criticism and friendly advice.

We wish to express our gratefulness to our guide Mrs. MAHALAKSHMI ME., Assistant Professor of Computer Science and Engineering, for his invaluable guidance and constructive suggestions

IX REFERENCES

[1] S. W. N. Cheung, S. C. Ng, and A. K. F. Lui, "A framework for effectively utilizing human grading input in automated short answer grading," Int. J. Mobile Learn. Organization, vol. 16, no. 3, p. 266, 2022.

[2] W. H. Gomaa and A. A. Fahmy, "A survey of text similarity approaches". J. Compute. Appl., vol. 13, no. 1, pp. 11–23, 2013.

[3] R. Mihalcea, C. Corley, and C. Strap Arava, "Corpus-based and knowledge-based measures of text semantic similarity," in Proc. AAAI, vol. 6, 2006, pp. 775–780.

[4] R. A. Farouk, M. H. Khafagy, M. Ali, K. Munir, and R. M.

Badry, "Arabic semantic similarity approach for Farmers' complaints," Int. J. Adv.

Compute. Sci. Appl., vol. 12, no. 10, 2021.

[5] E. Rslan, M. H. Khafagy, K. Munir, and R. M. Badry, "English semantic similarity based on map reduce classification for agricultural complaints,"

Int. J. Adv. Compute. Sci. Appl., vol. 12, no. 12, pp. 1–8, 2021.

[6] W. H. Gomaa and A. A. Fahmy, "Automatic scoring for answers to Arabic test questions," Compute. Speech Lang., vol. 28, no. 4, pp.

[7] N. Y. Habash, "Introduction to Arabic natural language processing," Synth. Lectures Human Lang. Technol., vol. 3, no. 1, pp. 1–187, Jan. 2010.

[8] T Senthil Prakash, V CP, RB Dhumale, A Kiran., "Auto-metric graph neural network for paddy leaf disease classification" - Archives of Phytopathology and Plant Protection, 2023.

[9] T Senthil Prakash, G Kannan, S Prabhakaran., "Deep convolutional spiking neural network fostered automatic detection and classification of breast cancer from mammography images" - Research on Biomedical Engineering,

[10] TS Prakash, SP Patnayakuni, S Shibu., "Municipal Solid Waste Prediction using Tree Hierarchical Deep Convolutional Neural Network Optimized with Balancing Composite Motion Optimization Algorithm" - Journal of Experimental & Theoretical Artificial ..., 2023

[11] TS Prakash, AS Kumar, CRB Durai, S Ashok., "Enhanced Elman spike Neural network optimized with flamingo search optimization algorithm espoused lung cancer classification from CT images" - Biomedical Signal Processing and Control, 2023

[12] C Aswath, T Prakash, P Kumari, N Thakur, R Sharma., " Effect of Gamma Radiation on Pollen Viability and Pollen Germination of Marigold Cultivar" - Think India Journal, 2019.

[13] R. Senthilkumar, B. G. Geetha, (2020), Asymmetric Key Blum-Goldwasser Cryptography for Cloud Services Communication Security, Journal of Internet Technology, vol. 21, no. 4, pp. 929-939.

[14] Ramana, S., et al. "Atmospheric Change on the Geographical Theme Finding Of Different Functions on Human Mobility". International Journal of Scientific Research in Computer Science and Engineering 6.2 (2018):

134-151.

[15] Senthilkumar, R., et al. "Pearson Hashing B-Tree With Self Adaptive Random Key Elgamal Cryptography For Secured Data Storage And Communication In Cloud." Webology 18.5 (2021): 4481-4497