International Journal of
Human Computations & Intelligence
Vol. 03, Issue. 04, 2024

RESEARCH ARTICLE OPEN ACCESS

Transformer Model to Evaluate Subjective Script

Madhushree[1] . Lavanya N L[2] . Shashi Rekha G[3] . Nagarathna C[3]

[1]Department of Computer Science and Engineering, M.S.Ramaiah University of Applied Science, # 470-P, Peenya Industrial Area 4th Phase, Peenya, Bengaluru.
[2]Department of Computer Science and Engineering, East West College of Engineering, Yelahanka new town Bengaluru, India.
[3]Department of Computer Science and Engineering, Sapthagiri College of Engineering#14/5, Chikkasandra, Hesaraghatta main road, Bangalore, India.

DOI: **10.5281/zenodo.14557335**
Received: 29 November 2024 / Revised: 20 December 2024 / Accepted: 22 December 2024

Abstract—Boards and institutions frequently provide offline subjective exams to a significant student body. It would be highly time-consuming to manually look into such a big volume of papers. There have been initiatives in the past to grade student responses utilising computer science. However, most of these efforts rely on typical counts or particular word counts to achieve the goal. On the contrary, the proposed method employs a Transformer Neural network model is used to process natural language. Components of Transformer model are position encoding, Word vector embedding, Attention mechanism and Feed forward neural network. Various Natural Language Processing techniques such as Tokenization, Word2Vec, Lemmatization and Semantic Checking.

Index Terms – Positional Encoding, Word vector Embedding, Attention Mechanism, Feed Forward neural network

I. INTRODUCTION

In current years, automated evaluation of subjective scripts with machine learning models has increased popularity. This is a result of the time consuming process involved in manually evaluating scripts using the outdated method, which could lead to uneven judgment. Despite the fact that there are several systems that can quickly evaluate multiple-choice or objective questions, they are inappropriate for subjective exams, which serve as the foundation of college- and board-level evaluations. These tests are attempted by a significant amount of students, and the moderator’s award grades based on how well they believe the students' descriptive answers reflect their knowledge. The final score for the student may be affected by the moderator's mood, the technique of evaluation, and their connection with the student.

MILESTONE RESEARCH.IN OPEN ACCESS

ISSN (Online): 2583-5696
Int. Jr. of Hum Comp. & Int.

Manual examination of these subjective replies is a tedious and expensive operation. This technique apply machine learning (ML) and natural language processing (NLP) methods to automate the subjective answer evaluation procedure which is one possible way to solve this problem. Automating the assessment of subjective scripts is a benefit to teaching fraternity, which can be achieved through recent developments in natural language processing (NLP) and machine learning. The invention of transformers, a kind of neural network design is used and transformed NLP developed, it is one of the most important developments in this area.

Transformers are proven efficient technique for a variety of NLP tasks, including question-answering, text classification, and language translation. They excel at tasks requiring lengthy text sequences, which are frequent in subjective scripts. Building transformers expressly for the subjective script evaluation has attracted increasing interest. Large datasets of annotated scripts is used to train the model, after building the model it can then be used to assess essays on a variety of criteria, including coherence, structure, and evidence utilisation. Education professionals can save time and work while obtaining more reliable and impartial findings by automating the examination of subjective scripts. Additionally, feedback on subjective can be achieved through this model, allowing them to improve their skills and performance. Overall, building transformers for subjective script evaluation has the potential to revolutionize for academician and help students achieve better outcomes.

II. LITERATURE SURVEY

Numerous researches have been reported in the field of subjective script Evaluation. In [1], The paper model involve k-nearest neighbors technique is applied to categorize text in this paper's method forsubjective scoring. The approach suggests a new similarity metric based on both word matching and word ordering and is made to operate with the English language. This method might be a useful tool for instructors to evaluate unclear responses. The expected accuracy, however, may be less accurate in circumstances with a high number of target classes or big score disparities. In [2], The paper various similarity algorithms is applied, where Euclidean Distance and Manhattan Distance, are used to generate an occurrence matrix which consists s the model answer and the student answer. The score is then determined through assessing the occurrence matrix. In this tailored method an accuracy level of 80–85% of human expectations and outcomes through several studies including both human and machine evaluation.

In [3], The methodology of the paper is designed to be resilient to changes in the language used by the writer. The system proper weights are used to verify the subjective answer and weights doesn't affect the addition of new words or changes to the weights of existing terms. This ensures that the system remains accurate and consistent even after applied to new or unexpected language. Moreover, the system has the potential to be improved in the future. Example, the grammar checking module could be modified to confirm to accepted standards, improving system with good accuracy and reliability for further. Additionally, new features could be added to the system to improve its ability to evaluate subjective answers, semantic analysis is preferred to better analysis and understand of the text meaning. In [4], The Idea behind this paper involves creation of software that uses OCR and an assessment algorithm to evaluate theoretical responses and assign grades based on keyword matches. The assessment system evaluates the answer paper, Considering the grammatical sense of text phrase, length of the response, and number of significant keyword matches, and assigns marks appropriately. TheOCR is used to extract text from submitted files. This programme speeds up the examination of responsesand generates findings right away, reducing human labour and saving time.

In [5], This paper proposed a novel method for automatically scoring students' brief subjective responses. The system evaluates the text using the text's semantics and uses a model or template response as a reference. As long as the answer's semantic and syntactic structure are right, it can take any of a number of correct forms. The author uses Conceptual Graphs (CG) as a logical framework to explain links between concepts and convey meaning. This enables the system to properly interpret the student's response and contrast it with the model response. CGs offer an effective method for automated grading that can identify the differences in language and emotion while preserving the objectivity and correctness of the assessment. In [6], This study introduces an automated grading system that is knowledge-oriented and goes beyond simple keyword matching. The system combines dictionary mapping and Latent Semantic Analysis (LSA) to allocate marks to pertinent replies after using ontology to map out the domains linked to a given term.The algorithm additionally checks the grammar and syntax to make sure the response is sound both grammatically syntactically, and semantically. It even ensures that students get fair and accurate evaluations of their responses, this multifaceted approach to automated grading depicts how natural language processing techniques have the potential to completely transform the area of education.

In [7] , This work proposes a system that analyses subjective replies using a semantic learning algorithm and gives students feedback on how to strengthen their responses by emphasising essential information from a model answer. Online tests focused at self-improvement benefit greatly from this automated technique. The system also caters to the demands of students with special needs by offering several ways of speech-based usability features, like the capacity to audibly respond to inquiries and queries. In [8], In this work, an automated evaluation approach for subjective issues is proposed. This type of system is plagued by many fundamental issues that lower the quality of the assessment, like the problem of trickiness, which is brought on by students amassing irrelevant knowledge when responding to the questions, which results from the fact that there are many ways to express a known concept as most words typically have multiple meanings. To address the issue of trickiness a reference unit vector is used and latent semantics indexing and basic topic ontology are offered as solutions to these difficulties of synonyms.

In [9], The Approach in this paper utilizes the cosine similarity function identifies the degree of similarity between the script to be evaluated and the scheme of evaluation. The lexical similarity is calculated by applying cosine similarity function between the two texts, based on which marks are allocated to the student. This approach allows for a more objective script evaluation, as the system is able to compare the script and scheme of solution to an established benchmark. By using NLP techniques to automate the grading process, this system has the potential to significantly reduce the workload of teachers and provide more consistent and accurate evaluations of student performance.

III. METHODOLOGY

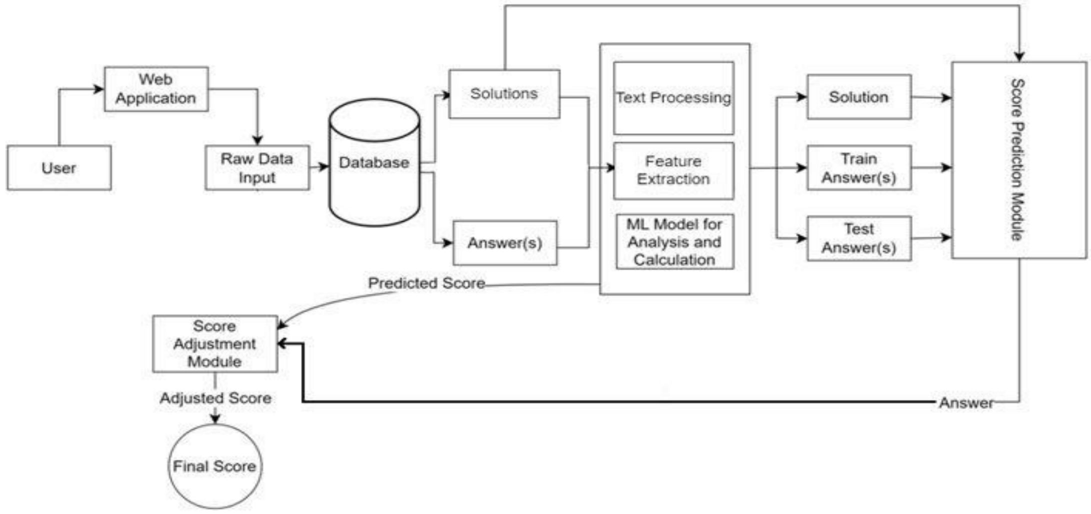

Fig. 1: System Architecture

The suggested technique forecasts a subjective script's score using a transformer model. When it comes to NLP activities like sentiment analysis, question answering, and language translation, the Transformer model is a one of the neural network type that has been frequently employed. It is used to assess subjective scripts, such essays or written answers to questions.

The primary characteristics Transformer model's are listed below:

1. Self-Attention: The model performs better on tasks which ask for a thorough comprehension of the input sequence thanks to the self-attention mechanism, which enables the model concentrate on various elements of the input sequence.
2. Multi-Head Attention: The model may focus on several elements of the input sequence at once thanks to the multi-head attention mechanism, which improves performance on challenging tasks.
3. Positional Encoding: The model may capture sequential information by accounting for the position of each character in the input sequence thanks to the positional encoding approach.
4. Feed forward Neural Networks: The model may represent non-linear interactions among the input layer and output layer sequences using feed forward neural networks.
5. Residual Connections and Layer Normalization: By minimizing the effect of vanishing gradients and easing the training process, residual connections and layer normalization help the model learn more successfully.

To implement a Transformer model for evaluating subjective scripts, the following steps are followed:

Data Preprocessing:

Preprocessing is a crucial NLP phase where raw text input is converted into a format which can be quickly analysed by computers. The many phases in data preparation are listed below:

1. Text Cleaning: In this phase, unnecessary characters from the text are removed, including punctuation, special characters, and numbers. Additionally, it entails changing the entire text to lower- or uppercase characters. By standardizing the material, this makes it simpler for the algorithms to comprehend.
2. Tokenization: Tokenization is the division of a textual passage into smaller units known as tokens. Themost typical kind of tokenization is dividing the text into words and spaces between each word. Depending on the needs, tokenization may also entail dividing the text into phrases, sentences, or even chapters.
3. Stopword removal: Stopwords are often used words in a English language like "the," "is," "and," "of," etc. that have little to no significance and may be eliminated from the text without risk. Stopword elimination reduces the size of data and increases algorithmic performance.
4. Stemming and lemmatization: The process of stripping words of their suffixes, such as "-ing," "-ed," and "-s," and returning them to their original root form. For instance, the words "running" and "runned" can both be spelt "run". Lemmatization is a more sophisticated kind of stemming that split down words into simplest form, or lemma, using a lexicon and morphological analysis. Lemmatization, for instance, might make the word "better" into the word "good".
5. Part-of-Speech Tagging: This method categorizes by each word is assigned with a grammatical tag in the document depending on its phrase, such as whether word is a noun, verb, adjective, etc. Many NLP tasks, like sentiment analysis and named entity identification, employ this to grasp the text's context.
6. Named Entity Recognition: It is the process of locating and categorizing named entities in the text, such as individuals, locations, businesses, etc. numerous NLP applications, including information extraction and question answering, benefit from this.
7. Feature Extraction: Feature Extraction is a procedure of converting text into a numerical format so thatmachine learning algorithms may utilise it is known as feature extraction. The document is represented as a set of features that capture the key aspects of the data using techniques like bag-of-words, TF-IDF, and word embeddings.

Training the transformer Model:

The Transformer model must then be trained. This entails specifying the model's architecture, including themodel's layer count, attention head count, and size of hidden layer. To train the model a sizable corpus of annotated data with scores or grades assigned to each script is used. The model gains the ability to gauge a script's grade based on its content. A suitable Preprocessing of the incoming data must be done to achieve the task at hand. Different regularization approaches, such dropout or weight decay, may be used during training to avoid over fitting the model to the training data.

Fine Tuning the Transformer Model:

A deep learning Technique called fine-tuning is in cooperated to raise a trained model's performance on a given task. Natural language processing (NLP) fine-tuning is training a pre-trained Transformer model on a limited quantity of task-specific data for you to tailor the model to the particular job.

The following are stages in perfecting the pre-trained Transformer model:

1. Task-Specific Data Preparation: Gathering and preprocessing the task-specific data is the first phase. Based on the objective, the data should be labeled with the proper labels, such as sentiment labels, named entities, or parts of speech.
2. Model Architecture: Typically, the base model is the pre-trained Transformer model, and task-specific layers are built including it. quantity of levels and size of the hidden layers in the task-specific layers can be adjusted based on difficulty in the work and computational resources.

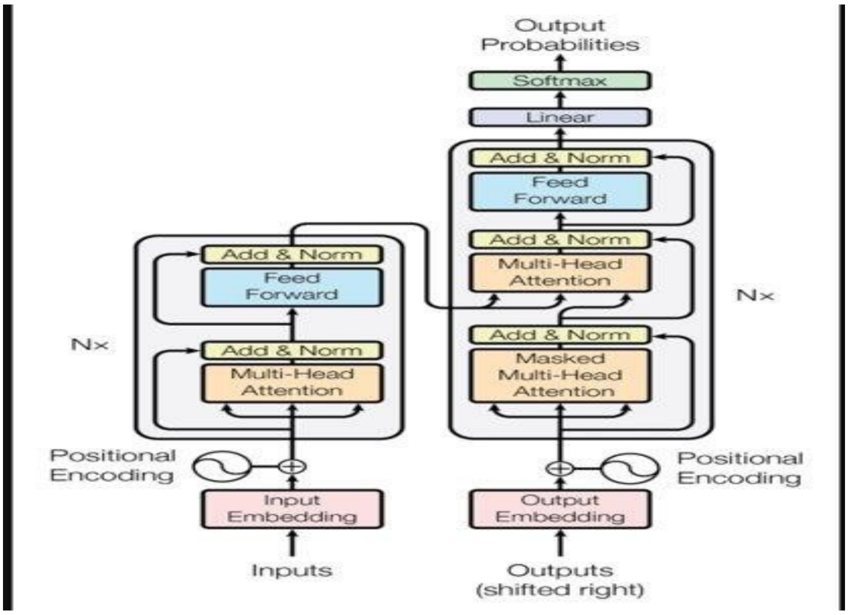

Fig. 2: Transformer Model Architecture

Evaluating Model:

After training and fine-tuning, the model may be used to assess new subjective scripts. Tokenizing the script creates numerical vectors that are subsequently input into the model. The algorithm generates a forecasted grade or score for the script.

Model Optimization:

Finally, the model may be enhanced in relation with performance by modifying its architecture or hyper parameters, or by combining numerous models using ensemble approaches.

Result:

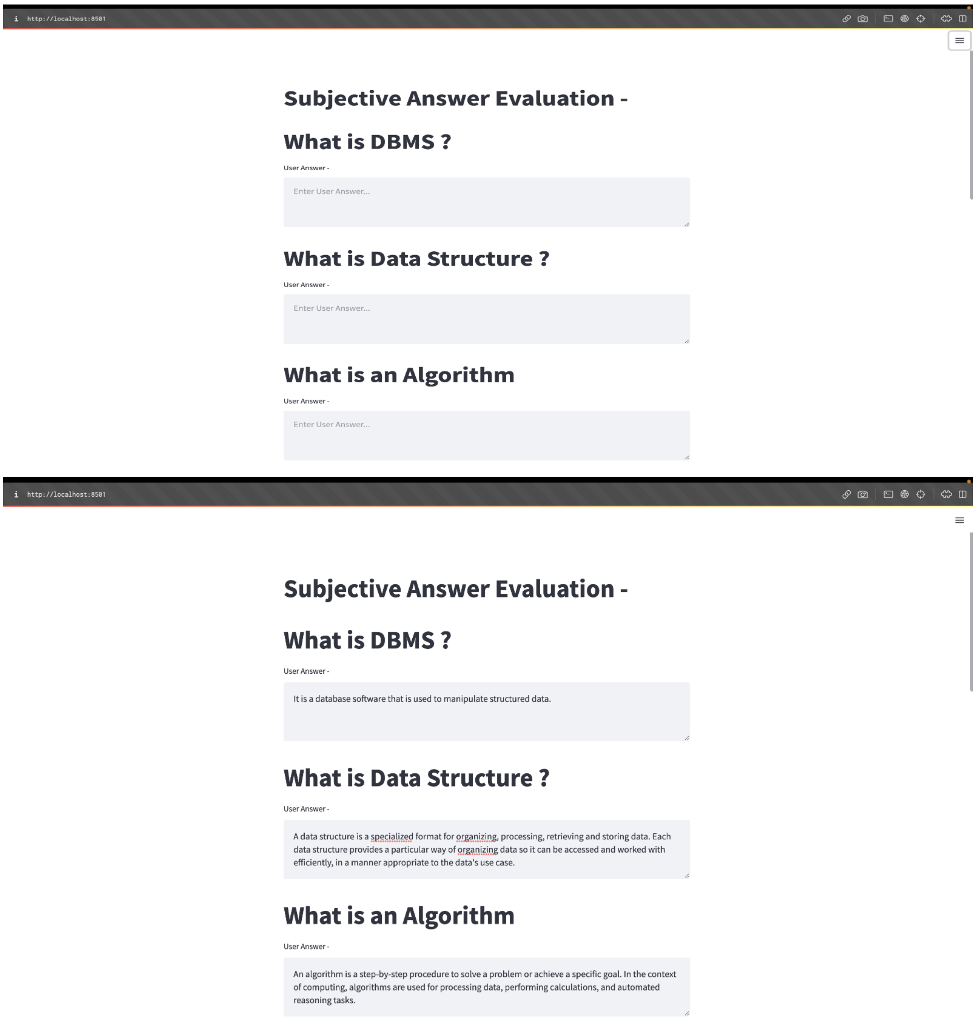

Fig. 3: Screenshots from the implementation process

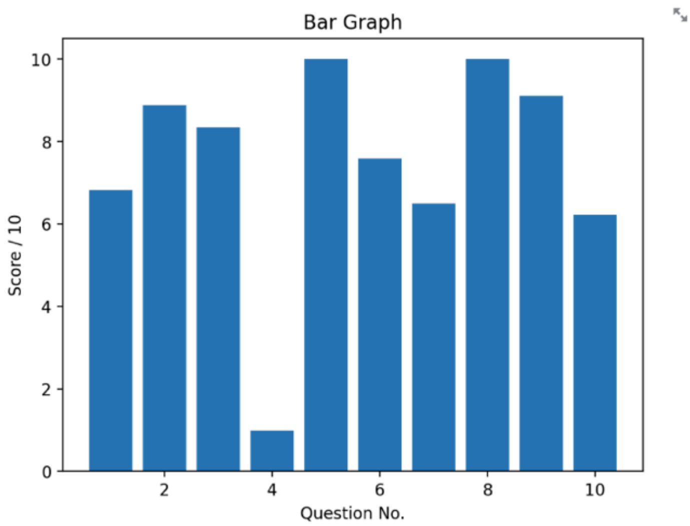

Fig. 4: Prediction numbers and the score representation

IV. CONCLUSION

In conclusion, automated grading may be achieved by assessing subjective answer scripts using machine learning (ML) and natural language processing (NLP) approaches. To improve the accuracy of grading procedure the ML model uses a good volume of training data. While ML and NLP approaches are useful for grading objective questions, it is vital to note that in mind it might not be as effective for analyzing complicated and nuanced replies that demand a deeper grasp of the topic matter. Therefore, to get the best results when assessing subjective answer scripts, a combination of automated and human grading may be required. To enhance the capabilities of ML and NLP models in analyzing subjective scripts, advancements include multilingual support, integration of external information sources, and the creation of explainable AI methodologies. These improvements make it possible to compare languages, grasp things and connections better, and give comprehensible explanations for similarity ratings. In the end, these developments improve user comprehension and performance while assessing subjective scripts.

REFERENCES

[1] Sriwanna, K. (2018, February). Text classification for subjective scoring using K-nearest neighbors.In 2018 International Conference on Digital Arts, Media and Technology (ICDAMT) (pp. 139-142). IEEE.

[2] Dave, N., Mistry, H., & Verma, J. P. (2017, February). Text data analysis: Computer aided automated assessment system. In 2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT) (pp. 1-4). IEEE.

[3] Hu, X., & Xia, H. (2010, April). Automated assessment system for subjective questions based on LSI.In *2010 Third International Symposium on Intelligent Information Technology and Security Informatics* (pp. 250-254). IEEE.

[4] Rosy, S. V. D., Viola, G. V. P., & Sathya, R. (2020). Intelligent short answer assessment using machine learning. International Journal of Engineering and Advanced Technology, 9(4), 1111-1116.

[5] Jain, G., & Lobiyal, D. K. (2017). Conceptual graphs based approach for subjective answers evaluation. International Journal of Conceptual Structures and Smart Applications (IJCSSA), 5(2), 1-21.

[6] Rokade, A., Patil, B., Rajani, S., Revandkar, S., & Shedge, R. (2018, April). Automated grading system using natural language processing. In 2018 Second international conference on inventive communication and computational technologies (ICICCT) (pp. 1123-1127). IEEE.

[7] Johri, E., Dedhia, N., Bohra, K., Chandak, P., & Adhikari, H. (2021, May). ASSESS-Automated Subjective Answer Evaluation Using Semantic Learning. In Proceedings of the 4th International Conference on Advances in Science & Technology (ICAST2021).

[8] Jagadamba, G. (2020, October). Online subjective answer verifying system using artificial intelligence. In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 1023-1027). IEEE.

[9] Mandge, V. A., & Thalor, M. A. (2021, March). Revolutionize cosine answer matching technique for question answering system. In 2021 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 335-339). IEEE.

[10] Sreedhar Kumar, S., Ahmed, S. T., Mercy Flora, P., Hemanth, L. S., Aishwarya, J., GopalNaik, R., & Fathima, A. (2021, January). An Improved Approach of Unstructured Text Document Classification Using Predetermined Text Model and Probability Technique. In *ICASISET 2020: Proceedings of the First International Conference on Advanced Scientific Innovation in Science, Engineering and Technology, ICASISET 2020, 16-17 May 2020, Chennai, India* (p. 378). European Alliance for Innovation.

[11] Kumar, S. S., Ahmed, S. T., Xin, Q., Sandeep, S., Madheswaran, M., & Basha, S. M. (2022). Unstructured Oncological Image Cluster Identification Using Improved Unsupervised Clustering Techniques. *Computers, Materials & Continua*, *72*(1).

[12] Ahmed, S. T., & Basha, S. M. (2022). *Information and communication theory-source coding techniques-part II*. MileStone Research Publications.