"Subjective Lens: Exploring Subjective Responses"

1st Prithvi Gohil
[1]Student,[1]Artificial Intelligence and Data Science,[1]KJ Somaiya Institute of Technology, Mumbai, India

2nd Anish Naik
[2]Student,[2]Artificial Intelligence and Data Science,[2]KJ Somaiya Institute of Technology, Mumbai, India

3rd Bhavik Sachani
[3]Student,[3]Artificial Intelligence and Data Science,[3]KJ Somaiya Institute of Technology, Mumbai, India

4th Prof. Devanand Bathe
[4]Professor,[4]Artificial Intelligence and Data Science,[4]KJ Somaiya Institute of Technology, Mumbai, India

Abstract: In evaluating students' comprehension and analytical skills, subjective response evaluation is essential. However, subjectivity and inefficiency are common with older manual assessment procedures. This research suggests a system to automate and enhance the assessment of subjective responses that makes use of machine learning (ML) and natural language processing (NLP) techniques.

By taking grammatical structure, contextual comprehension, and semantic significance into account, the system seeks to thoroughly evaluate students' responses. It processes and interprets textual input using machine learning (ML) models, making it easier to find important patterns and meanings in answers. Semantic similarity measures, summarization algorithms, sentiment analysis, and other NLP techniques are used to probe further into the replies' content and context.

The suggested model consists of many phases, including digitizing handwritten or scanned replies, identifying pertinent characteristics, matching student responses to model answers, and developing a grading system. Methods like summarization, keyword matching, and sequence alignment help measure how well student replies match the predicted model responses in terms of coherence and resemblance. By offering a more uniform and impartial review process, the system seeks to reduce subjectivity and inconsistency. This framework aims to improve subjective response assessment accuracy, efficiency, and fairness by combining ML and NLP, giving teachers and students a more dependable way to measure performance.

Key Words: Machine Learning, Natural Language Processing, Subjective Answer

I. INTRODUCTION

Within the field of education, subjective response evaluation is a basic yet complex component of evaluating students' understanding and critical thinking skills. Subjectivity, inconsistent replies, and the demanding requirements of large-scale evaluations are common problems for the old techniques of manually grading subjective responses. But the incorporation of state-of-the-art technology like Natural Language Processing (NLP) and Machine Learning (ML) has the potential to revolutionize this crucial procedure.

This study explores the innovative field of using ML and NLP approaches to transform the assessment of subjective responses. A paradigm shift in the field of academic evaluations is expected as a result of leveraging the inherent capabilities of these cutting-edge technologies.

There is a previously unheard-of chance to automate and improve the evaluation of subjective responses because to the confluence of ML and NLP. The system can evaluate large volumes of textual data, identify complex patterns, and extract subtle insights from student replies thanks to machine learning algorithms. Simultaneously, the system is enabled by NLP approaches to decipher the subtleties of meaning, syntactical patterns, and contextual depths that are encoded in these subjective responses.

This combination is important because it can improve subjective response evaluation's efficiency, objectivity, and correctness. This method aims to reduce the inherent biases, inconsistencies, and laboriousness of manual evaluations by utilizing advanced algorithms.

The purpose of this study is to clarify the architecture, process, and expected results of a system intended to use ML and NLP for the analysis of subjective answers. It looks at how this collaboration may change the face of assessment and provide teachers and students with a more consistent, trustworthy, and perceptive way to evaluate performance.

By investigating this novel framework, a transition toward a more effective, equitable, and perceptive method of evaluating subjective responses is anticipated, constituting a noteworthy advancement in the field of educational assessments.

II. RELATED WORKS

This section will serve as a review of previous research on comparable systems that were in use when the system was being developed.

Aditi Tulaskar, Aishwarya Thengal and Kamlesh Koyande [1] developed a system for Annual boards and universities exams that are conducted offline, requiring manual evaluation of subjective questions. This process is lengthy and time-consuming. Competive and entrance exams use objective or multiple choice questions, making evaluation easy and error-free. However, there is no provision for subjective questions. Automating the evaluation process for descriptive answers would be beneficial for educational institutions.

Shihui Song and Jason Zhao [2]develop a model for Automated Essay Scoring Using Machine Learning. They explain that An automated essay scoring system was developed to evaluate 13,000 essays from an online Machine Learning competition on Kaggle.com. The system focused on essay style and maturity, evaluating Linear Regression, Regression Tree, Linear Discriminant Analysis, and Support Vector Machines.

Dharma Reddy Tetali, Dr. Kiran Kumar G[3] developed, ApTeSa a tool developed for automated evaluation of subjective answers using PyQt, Python, and its modules. It uses a systematic technique and works in semi-automated or complete automated modes. The tool allows faculty to reevaluate answers and update results, with the semi-automated mode yielding better results than the complete automated mode. The paper provides an example evaluation of a subject titled 'Structural Engineering'.

Nisarg Dave, Harsh Mistry and Jai Prakash Vera [4]discusses the challenges of computer-aided automated assessment, focusing on subjective evaluations in text analytics. It proposes a model to handle these issues, allowing faculty to avoid manual evaluations and automatically grade students and provide a summary report.

Xinming Hu and Huosong Xia[5] proposed a paper stating it uses latent semantic indexing to present an automated assessment system for subjective questions. Transferring reference answers to a term-document matrix through Chinese automatic segmentation techniques and subject ontology, the matrix is then projected to a k-dimensional LSI space via singular value decomposition. Trickiness is reduced by introducing a reference unit vector. The similarity between the projected vectors is used by the system to assess the quality of the solution. The viability of this theoretical architecture and flow for automated evaluation of subjective questions is demonstrated by experimental results.

The paper introduced by **Hu, H., Liao, M., Zhang, C., & Jing, Y.[6]** stated an independently recurrent neural network (IndRNN) that solves gradient problems by combining recurrent weight matrices with an attention model and long short-term memory (LSTM). The outcomes demonstrate how well these models adjust to text classification tasks, resolving issues like explosion and gradient vanishing.

III. PROBLEM STATEMENT

Subjectivity and inconsistencies make analyzing subjective replies in the context of academic examinations a challenging process. While manual grading is a conventional approach, it can result in subjective biases and differing interpretations among assessors, thereby leading to disparities in students' grades. Due to these limitations, it is more challenging to carry out the in-depth examination needed to ascertain how effectively pupils actually comprehend and apply topics. Additionally, the drawn-out manual assessment procedure slows down evaluation and hinders prompt feedback, both of which have an impact on the learning trajectory. Additionally, the amount of arbitrary replies in tests and larger courses cannot be efficiently managed by the current systems due to their inability to scale. Given these grave flaws, a novel approach that leverages machine learning (ML) and natural language processing (NLP) skills is desperately required. With the goal of revolutionizing subjective answer analysis, this cutting-edge approach promises impartiality, effectiveness, and deeper insights into students' replies. This project aims to build a complete, standardized, and trustworthy assessment system that has the potential to change the assessment paradigm in educational contexts by merging ML algorithms with NLP approaches.

IV. PROPOSED MODEL:

Our system accepts student answer sheets as well as model answers as inputs, utilizing AI, ML, and NLP. It consists of three main tasks: matching keywords from student to model replies, sequencing the responses, and summarizing both model and student answers. Both of the summarized inputs are examined using keyword alignment and sequencing methods. AI-powered understanding can improve the work of text evaluation. NLP and optical character recognition (OCR) are two technologies that help achieve this improvement. The final results of the students are based on the weighted average of these tasks.

As demonstrated in Figure 1, the model first converts scanned student answer sheets into readable text by processing them in PDF format. It then compares student responses to model answers by ensuring response sequencing is correct, aligning keywords, and summarizing content. The student's final results are ultimately determined by a weighted average of these procedures.

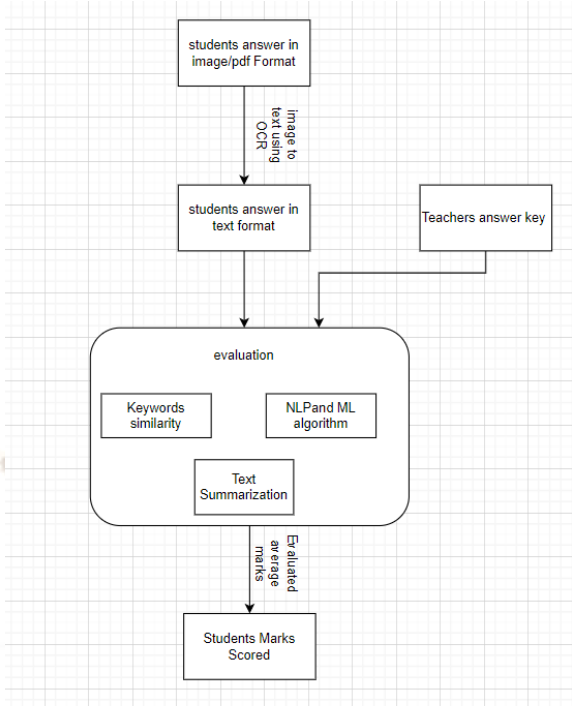

Fig. 1. Workflow of proposed system

The following section discusses the most effective techniques and algorithms that can be used to the various system phases:

(1)Hidden Markov Model (HMM):

A Markov model is a stochastic (probabilistic) model that is applied to a framework in which the present state is the only factor that determines future states. We undertake the work under the supposition that we can use a limited state change organization to address the Markov model for the sake of POS labeling. In the limited state progress network, each hub corresponds to a state, and each coordinated edge that leaves the hubs corresponds to a possible transition from one state to another. Every edge has a number assigned to it that indicates the probability that a particular advancement would reach its current condition and, in addition, that a change from a random state will inevitably result in a value of 1.

In Hidden Markov P(t) and P(w|t) are defined by the model as follows: w is a sequence of words, and 't' is a sequence of tags.

Where the odds of a transition are:

$$P(\mathbf{t}) = \prod_i P(t_i \mid t_{i-1})$$

$$P(\mathbf{t}) = \prod_i P(t_i \mid t_{i-1}, t_{i-2})$$

Emission probabilities:

$$P(\mathbf{w} \mid \mathbf{t}) = \prod_i P(w_i \mid t_i)$$

For lexical computations in HMM, the well-known search method Viterbi is employed. This approach uses dynamic programming to generate a list of the phrases that have been observed in a given order. The n-grams approach, which this algorithm employs, works by making a number of assumptions [7].

(2)Recurrent Neural Network (RNN):

Recurrent neural networks (RNNs) are a type of neural network in which the contribution to the current advance is determined by taking into account the yield from previous advances. Normally, all the information inputs and outputs in a neural organization are independent of each other. Nevertheless, there are situations where memory for the past words is necessary, such as when one needs to predict the next sentence's expression. As a result, RNN emerged and, with the help of a Hidden Layer, resolved this problem. The core and most important part of an RNN is its hidden state, which retrieves some information about a cluster. In a feed-forward neural network with a recursive hidden state, a recurrent neural network's hidden state, 'ht,' is activated by the preceding state at time-step t. Recurrent neural networks may therefore handle data of any length and dynamically model contextual knowledge [6].

RNN have a "memory" which recollects all data about what has been determined. It utilizes similar boundaries for each contribution as it plays out similar assignment on every one of the information sources or covered up layers to create the yield. This diminishes the intricacy of boundaries, in contrast to other neural organizations.

(3)Semantic Textual Similarity (STS):

Semantic equivalence gauges how closely related two objects' meanings are, assessed using methods like word embeddings or natural language processing. These techniques measure the similarity in content or context between the objects, such as with cosine similarity in NLP for sentences or documents. It is an essential linguistic approach in Natural Language Processing (NLP). Among its many uses are text classification, information processing, sentiment analysis, query response, semantic search, and plagiarism detection. Accuracy is vital during the semantic similarity step [8].

The Semantic Textual Similarity (STS) task is utilised to evaluate sentence embeddings in a one-way fashion. According to STS, a good sentence representation should be able to distinguish between dissimilar and comparative sentences by encoding the semantic information of the phrase. In practise, we measure a model's ability to recognise similar sentences by first recording the cosine comparability of multiple sentence pairs, each with a closeness score between 0 and 5, and then observing the relationship between these scores and figured cosine similitudes.

(4)Latent Semantic Indexing (LSI):

A theory and method called Latent Semantic Indexing (LSI) describes how to extract and define a term's contextual meaning from a large body of text by performing mathematical calculations. LSI is the name for the algebraic extension of the traditional vector model. By translating documents (and symmetry terms) into a vector space with reduced dimensionality (using SVD to minimize the dimensions), the latent semantic space, LSI aids in the resolution of synonymy and polysemy issues that beset automated information retrieval systems [9].

Latent semantic analysis is a useful technique for text analysis since it takes the language's meaning into account. Latent Semantic Analysis (LSA) is used to find the hidden topics that the paper or text represents. Then, identical texts are grouped using the hidden subjects. Since LSA is an unsupervised method, the precise subject of the document is unknown. The easiest method for identifying identical texts is to use vector representation of the text and cosine similarity. In a vector representation, every paper is represented as a vector. This vector is known as the document-term matrix.

Term Co-occurrence Matrix: (vocabulary size) * (vocabulary size) is the matrix's dimension. It displays the frequency with which the dataset's terms occur together. We can determine which words belong together with the help of the matrix. Consequently, the LSA returns definitions that represent the supplied text in place of topics. The principles are a list of words that, to the best of their abilities, sum up the text.

(5)Word2Vec:

Text data is processed by Google's Word2vec neural network. Word2Vec is a collection of learning models that consists of Skip-gram and Continuous Bag of Words (CBOW). Whereas Skip-gram predicts words based on their context, CBOW predicts words based on their meaning. Word2Vec uses a single learning model to process the text corpus and generate word vectors. Word2Vec builds a vocabulary from the training text corpus before learning each word's vector representations. Word2Vec will also calculate the cosine distance between every word. Therefore, it is beneficial to cluster related phrases based on their distances from one another [10].

Word2vec is a natural language processing technique. Through the use of a neural network interface, the word2vec algorithm discovers This model, like Word2vec, learns word connections from vast text data. Once trained, it can propose synonyms or related terms for a given sentence. Each word gets linked to a numerical vector, and these vectors are designed to help measure how closely related words are by using a basic math formula (cosine similarity) between the vectors.Word2vec is a collection of related word embedding models. Word language settings can be replicated by these shallow two-layer neural networks thanks to their acquired knowledge. Word2vec receives a large text corpus as input and produces a vector space with several hundred dimensions, where each unique word is associated with a corresponding vector. Terms with comparable contexts in the corpus are positioned adjacent to one another via term vector placement in the vector space.

Word2vec is a neural network with two layers that processes text by "vectorizing" words. It uses a corpus of textual data as input and outputs a feature vector sequence that characterizes phrases in the corpus. Word2vec transforms text into a numerical representation that deep neural networks can understand, despite the fact that it is not a deep neural network.

Clustering vectors in vector space that include similar terms is the aim of Word2vec. Put another way, it uses mathematics to search for correlations. Word2vec generates vectors, which are distributed numerical representations of word properties, such as the meaning of individual phrases. It accomplishes this without needing human help.

Each object in the vocabulary created by the Word2vec neural net has its own vector, which can be queried to find word associations or input into a deep learning network.

(6)Bag of Words:

Clustering vectors in vector space that include similar terms is the aim of Word2vec. Put another way, it uses mathematics to search for correlations. Word2vec generates vectors, which are distributed numerical representations of word properties, such as the meaning of individual phrases. It accomplishes this without needing human help.

Each object in the vocabulary created by the Word2vec neural net has its own vector, which can be queried to find word associations or input into a deep learning network.

Each text is split up into a list of terms (words) by the BOW, which then computes feature vectors depending on how frequently those terms occur.

Since there are many terms in every natural language, feature vectors with millions of dimensions are usually produced by the BOW solution. It is evident that most classifiers cannot handle this number, hence dimensionality reduction methods must be used. Next, a training set and a testing set are created from the dataset. The former becomes a model after it is put into a classification algorithm. The adequacy of the construction model is then evaluated using the trial collection [11].

The most fundamental technique for representing numerical text is the Bag of Words (BoW) model. Similar to the expression itself (a string of integers), a phrase can also be expressed as a bag of words vector.

Consequently, the text is preprocessed using the Bag of Words model, which converts it into a bag of words and counts the total number of times the most frequently used terms appear.

Drawbacks of BoW:

• Our vocabulary and the length of the vectors will both grow if the new sentences contain new terms.

• Moreover, a sparse matrix will be produced by the vectors' high number of zeros, which is what we want to avoid.

• No information regarding the sentence structure or word order within the text is preserved.

(7)TF-IDF:

A statistical test called the TF-IDF evaluates a term's significance to a text within a collection of documents. This is achieved by multiplying two metrics: the reciprocal reference frequency of a word across a group of papers and the number of times the term appears in a document.

It has many uses, such as phrase ranking in machine learning algorithms for Natural Language Processing (NLP) and autonomous text interpretation.

The TF-IDF (term frequency-inverse document frequency) algorithm was created for database search and retrieval. It functions by rising in proportion to a word's frequency of occurrence in a document, while being counterbalanced by the quantity of documents that contain the word. Therefore, terms that are frequently used in all documents—such as this, what, and if—rank poorly since they don't convey much information about the document specifically.

Techniques for grouping data can only be applied to ordered data. Unstructured data must therefore be converted into ordered data. Documents are represented by rows, while terms are represented by columns, as it anticipates extracting words from the records. The dimension curse, a problem caused by the large word count, lowers algorithm performance. The utilization of a function filtering approach can help reduce these expressions. The most common words in the corpus are eliminated using the TF-IDF approach, leaving just the most crucial phrases [12].

(8)WordNet:

WordNet comprises word databases in the English language. WordNet is arranged by definition and context, not alphabetically like a dictionary would be. Conventional dictionaries were created with people in mind, but modern lexical resources must be more computer-friendly. Here's when WordNet can be useful. WordNet is a network of words that is connected both lexically and semantically. Adjectives, adverbs, verbs, and nouns are grouped together as synsets, which are sets of cognitive synonyms that each convey a unique meaning. Sync Sets are connected by conceptual, semantic, and lexical links.

With the aid of WordNet's lexical knowledge base, two terms, w1 and w2, can have their semantic similarity sim(w1; w2). These term groupings are arranged into collections of synonyms, or synsets, that represent the idea that all words have similar meanings. A synset's knowledge includes definitions of these concepts and links to other synsets that are comparable. WordNet's synonyms are arranged in a hierarchical tree-like structure, with a few generic keywords at the top and a number of particular words at the bottom [13].

(9)Context- free grammar (CGF):

A set of rules that characterizes the set of all well-formed sentences in a language is known as a context-free grammar (CFG). Each law, when read from left to right, identifies the alternating component portions on the right side and explains a syntactic category on the left. In the language that a CFG describes, a sentence is a collection of concepts that may be created by methodically applying the rules, beginning with the rule that has the letter s on the left. A sentence parse is a series of rule implementations where a syntactic category is replaced by the right-hand side of a rule with that category on its left-hand side, and the sentence itself is produced by the last rule implementation [14].

V. FUTURE SCOPE

The future scope of subjective lens is exceptionally promising and holds significant potential for advancement in various fields. As technology continues to evolve, there is a growing need for automated systems that can effectively evaluate and analyze subjective responses, such as essays, reviews, and open-ended questions. Machine learning algorithms, coupled with NLP techniques, offer a robust framework for enhancing the accuracy and efficiency of subjective answer evaluation. Future research could focus on refining existing models, developing novel algorithms, and incorporating advanced linguistic features to improve the precision of subjective assessment. Additionally, exploring applications in educational settings, job recruitment, and content analysis for various industries can be pivotal. As the demand for automated evaluation tools rises, understanding the nuances of human language and context becomes crucial. Future endeavors may involve addressing challenges related to cultural variations, bias mitigation, and adapting models for different domains. Collaborative efforts between researchers, educators, and industry professionals can further propel the development of innovative solutions that contribute to the broader landscape of artificial intelligence in education and beyond. The potential impact of this research extends to automating labor-intensive tasks, enhancing decision-making processes, and fostering a more efficient and objective assessment of subjective responses in diverse contexts.

VI. CONCLUSION

In this research paper, the focus is on transforming the assessment of subjective responses in education using Machine Learning (ML) and Natural Language Processing (NLP). The paper reviews existing systems, highlighting their contributions, and identifies the limitations of manual grading. The proposed model integrates various advanced techniques such as Hidden Markov Models, Recurrent Neural Networks, Semantic Textual Similarity, and more to create a comprehensive system for automated evaluation. The goal is to address subjectivity, inconsistency, and scalability issues in traditional evaluation methods, promising a paradigm shift towards more efficient, objective, and accurate subjective response assessment in education.

VII. REFERENCES

[1] Aditi Tulaskar, Aishwarya Thengal and Kamlesh Koyande.: Subjective Answer Evaluation System. In: International Journal of Engineering Science and Computing, April 2017, Volume 7 Issue Number 4

[2] Shihui Song and Jason Zhao.: Automated Essay Scoring Using Machine Learning. In: https://nlp.stanford.edu/courses/cs224n/2013/reports/song.pdf

[3] Dharma Reddy Tetali, Dr. Kiran Kumar G and Lakshmi Ramana.: A Python Tool for Evaluation of Subjective Answers (APTESA). IJMET Volume 8, Issue 7, July 2017, pp. 247–255, Article ID: IJMET_08_07_029

[4] Nisarg Dave, Harsh Mistry and Jai Prakash Vera, Assist. Prof.: Text Data Analysis: Computer Aided Automated Assessment System. In: IEEE-CICT 2017, 978-1-5090-6218-8/17/$31.00 ©2017 IEEE

[5] Xinming Hu and Huosong Xia.: Automated Assessment System for Subjective Questions Based on LSI. 978-0-7695-4020-7/10 $26.00 © 2010 IEEE

[6] Hu, H., Liao, M., Zhang, C., & Jing, Y.: Text classification based recurrent neural network. In: sentence through a convolution filter, and can learn short652 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC 2020) 978-1-7281- 4323-1/20/$31.00 ©2020 IEEE.

[7] Sindhya K Nambiar, Sonia Jose, Antony Leons and Arunsree.:Natural Language Processing Based Part of Speech Tagger using Hidden Markov Model. In: Third International Conference on I- SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2019)IEEE Xplore Part Number:CFP19OSV-ART; ISBN:978-1-7281-4365-1

[8] SUNILKUMAR P and ATHIRA P SHAJI.: A Survey on Semantic Similarity. In: 2019 International Conference on Advances in Computing, Communication and Control (ICAC3).

[9] Jingmin HAO, Lejian LIAO and Xiujie DONG.: Improving Latent Semantic Indexing with Concepts Mapping Based on Domain Ontology. In: 2008 International Conference on Natural Language Processing and Knowledge Engineering 10.1109/NLPKE.2008.4906768.

[10] Long Ma and Yanqing Zhang.: Using Word2Vec to process big text data. In: 2015 IEEE International Conference on Big Data (Big Data). doi:10.1109/bigdata.2015.7364114

[11] Kholoud Alsmearat; Mahmoud Al-Ayyoub and Riyad Al-Shalabi.:An extensive study of the Bag-of-Words approach for gender identification of Arabic articles. In: 2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA). doi:10.1109/aiccsa.2014.7073254

[12] Prafulla Bafna; Dhanya Pramod and Anagha Vaidya.: Document clustering: TF-IDF approach. In: 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT).doi:10.1109/iceeot.2016.7754750

[13] Xiao-Ying Liu; Yi-Ming Zhou and Ruo-Shi Zheng.: Measuring Semantic Similarity in Wordnet. In: Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 19-22 August 2007. 1-4244-0973-X/07

[14] J.F. Power and B.A. Malloy.: Metric-based analysis of context-free grammars. In: Proceedings IWPC 2000. 8th International Workshop on Program Comprehension. doi:10.1109/wpc.2000.852491