iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35044

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Computerized Answer Grading

Anurag Langan[1], Avineet Bhat[2], Apoorva Singh[3], Ayush Bali[4], Prof. Nagashree Nagaraj[5]
[1, 2, 3, 4, 5] *Vidyavardhaka College of Engineering (Mysore, India)*

Abstract: Grading student answers is a tedious and time-consuming task. A study had found that almost on average around 25% of a teacher's time is spent in scoring the answer sheets of students. This time could be utilized in much better ways if computer technology could be used to score answers. This system will aim to grade student answers using the various Natural Language processing techniques and Machine Learning algorithms available today.

I. INTRODUCTION

Innovation in education has come a long way in improving the efficiency of grading. Short answers are open-ended questions. Unlike multiple-choice questions which have a given set of answers, short answers can have a range from a single word to a few sentences. Multiple-choice questions can be easily scored while short answers are scored by a human grader because of their nature. Therefore, scoring short answers is time-consuming and can also lead to human errors.

II. RELATED WORKS

Various techniques have been used to score answers.
Mohler and Mihalcea [1] compared various corpus-based text similarity measures and evaluated the impact of size on corpus-based metrics.
Gomma and Fahmy [2] proposed an unsupervised bag-of-words method. They tested different string-based and corpus-based similarity measures and then combined them to improve performance.
Alotaibi and Mirza [3] proposed an integrated information extraction (IE), decision tree learning (DTL), machine learning (ML) method. Their IE technology uses parser and dictionaries, while their ML technology uses classification rules extracted from DTL to automatically mark-free text.
Pulman and Sukkarieh [4] explored computational linguistic techniques for automatically marking short free-text responses. Initially, they solved the problem using IE. However, because artificially designed IE requires skills and expertise in domains and tools, they use several ML techniques (such as DTL and Bayesian learning) to learn IE models.

III. METHODOLOGY

This study aims to evaluate the short answers given by a student and automatically assign marks to the respective questions. NLP is used to extract text from the answer script and process the data. Various similarity measure has been implemented that is used as the parameter for assigning marks.

A. Text Preprocessing

Preprocessing a text simply means to bring a text into a form that is predictable and analyzable for a particular task.It transforms text into a more digestible form so that machine learning algorithms can perform better. It helps to get rid of unhelpful parts of the data, or noise, by converting all characters to lowercase, removing punctuation marks, and removing stop words and typos. Removing noise comes in handy when you want to do text analysis on pieces of data. The various steps involved in preprocessing are as follows.

1) *Tokenization:* Tokenization refers to dividing a document into a list of individual words. The word_tokenize method belonging to the nltk library can be used to return set of individual words.
2) *Stop Words Removal:* Stop words are very commonly used words (a, an, the, etc.) in the documents. These words do not really signify any importance as they do not help in distinguishing two documents.
3) *Lemmatization:* Lemmatization refers to reducing a word to its root form, as found in the dictionary. Lemmatization is different from stemming. On the other hand, in lemmatization, a word is reduced to its meaningful representation, as found in a dictionary.

B. Model

The model used is the Doc2Vec model. This model creates a numerical representation of the sentence or a paragraph. It computes a feature document for every document in the corpus. The vectors generated by Doc2Vec can be used to find the similarity between sentences. It is general and applicable to paragraphs of any lengths. A vocabulary is built from a sequence of sentences and this vocabulary is used to keep track of all the unique words in the model. Model should be trained for a few epochs.

C. Marks Scoring

The aim of this system is to automatically assign grades after evaluating the answers provided by students. It is the last step and its accuracy will determine the success of the system. In the system, answers have already been provided and these answers will be used to compare with the student answer for the similarity between them. Then the system will automatically assign grade based on the answer.

IV. RESULT

The main purpose with which this system was created was to automatically grade student answers using NLP and Machine Learning techniques. By automatically assigning grades we will reduce the time required to evaluate answers and also reduce the burden of teachers. To automatically assign grades we use a doc2vec model which creates numerical representations of the sentences. The sentences are converted to feature vectors and then these are used to compare the similarities in different sentences. Based on these similarities between the answers a grade is assigned to the answer.

V. CONCLUSION

The system that we have built uses NLP and ML based techniques to evaluate student answers. Our system includes the following steps (1) preprocess the text for further analysis; (2) find various similarity measures ; And (3) Score the student answers. Here, we accept the average common words as keywords, and ignore the most common and less common words. The abstract text was preprocessed with the help of NLTK, which is the leading platform for building python programs. Here, tokenization, StopWord deletion, lemmatization, two-letter group generation, and word frequency counting are performed as preprocessing. We also consider grammar and spelling errors to evaluate the answer script. After preprocessing, four similarity measures are calculated—synonym similarity, two-letter group similarity, cosine similarity and Jaccard similarity measures, which are used as the final mark scoring parameters. In order to score the marker, after investigating the best weight estimate, a weight value is assigned to each parameter. The weight value is multiplied by the parameter value to score the question. In this system, we considered three types of questions based on markers, and the answer script based on this question was manually evaluated. Compare manual tags with automatic scoring tags to validate the method we developed. In most cases, we find that our proposed method scores tags similarly to manually assigned tags. In rare cases, the automatically assigned mark is slightly higher or lower than the manually assigned mark. The limitation of our research is that we manually assign weight values to each parameter by conducting surveys.Therefore, our next goal is to introduce a machine Learning algorithm, the algorithm will be trained through various calculated parameters, and the algorithm will predict the score of the answer script. Also in the future, we will introduce some new technologies to effectively and accurately generate abstracts.

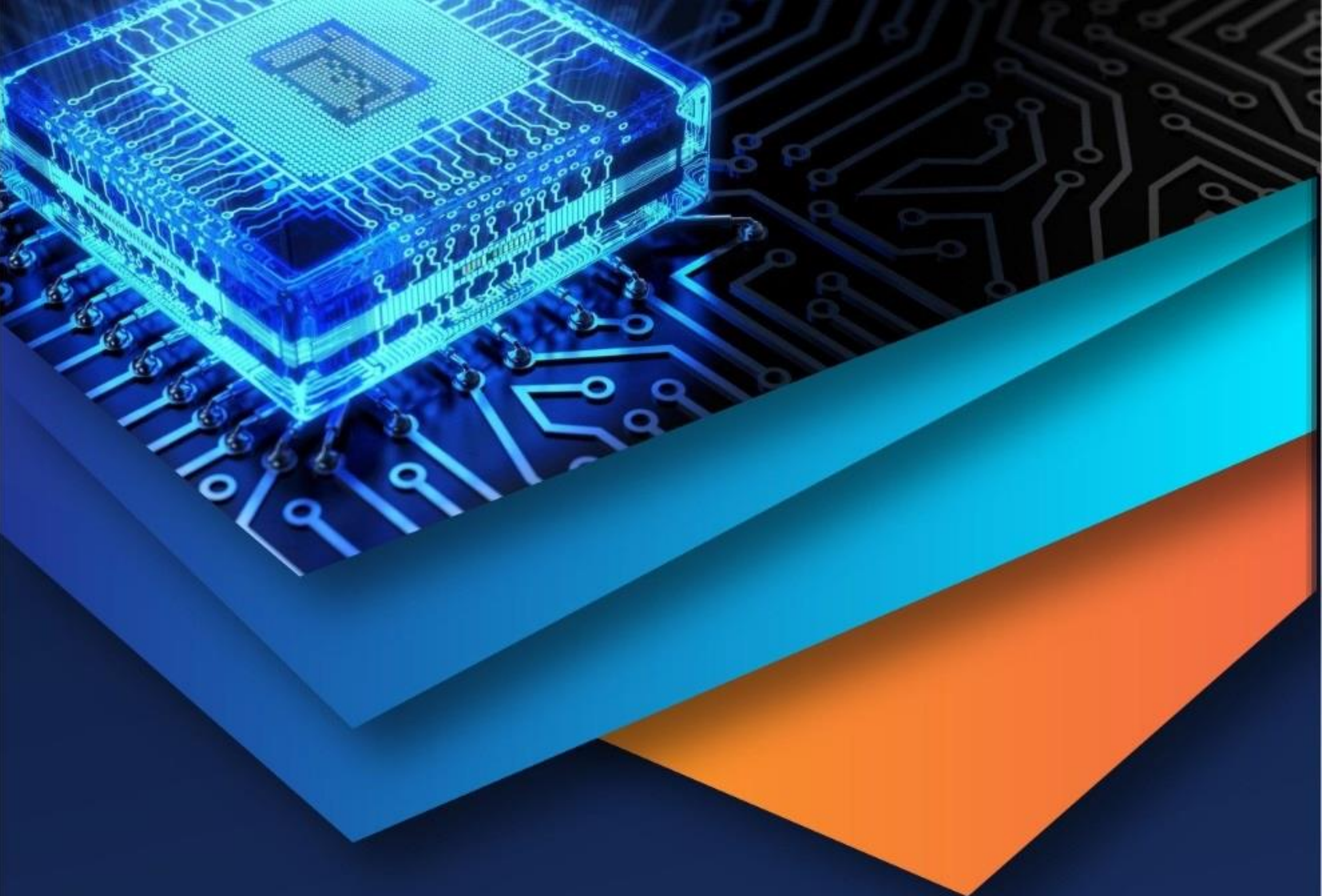
REFERENCES

[1] M. Mohler and R. Mihalcea, Text-to-text semantic similarity for automatic short answer grading, EACL: 12th Conference of the European Chapter of the ACL, 2009, pp. 567–575.

[2] W.H. Gomma and A.A. Fahmy, Short answer grading using string similarity and corpus-based similarity, International Journal of Advanced Computer Science and Applications 643 3(11) (2012).

[3] S.T. Alotaibi and A.A. Mirza, Hybrid approach for automatic short answer marking, SWDSI: 43rd Annual Meeting of the Southwest Decision Sciences Institute, 2012.

[4] S.G. Pulman and J.Z. Sukkarieh, Automatic short answer marking, EdAppsNLP:2ndWorkshop on Building Educational Applications Using NLP, 2005, pp. 9–16.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY