International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

Subjective Answer Evaluation using Natural Language Processing and Machine Learning

Abhishek Girkar[1], Mohit khambayat[2], Ajay Waghmare[3], Supriya Chaudhary[4]

[1, 2, 3] Student, [4] Professor, Dept. of Information Technology, Padmabhushan Vasantdada Patil Pratishthan's College of Engineering, Mumbai, Maharashtra – 400022

Abstract - *Every year boards and universities exams are conducted offline mode. Large number of student attend subjective type exam. For evaluation of such large number of paper manually required hard efforts. Sometimes quality of evaluation may change according to mood of evaluator. The evaluation work is very lengthy and time consuming. Competitive and entrance exams typically contain objective or multiple choice questions. These exams are evaluated on machine as they conducted on machine and therefore their evaluation is easy. It also saves multiple resources and human interaction and hence it is errorless. There are multiple system are available for evaluation objective (MCQ) type question but there is no provision for subjective (Descriptive) type question. It will be very helpful for educational institutions if the process of evaluation of descriptive answers is automated to capably assess student's exam answer sheets.*

Key Words: NLP, Contextual similarities, Semantic Analysis Grammatical Correction, WordNet, TFID, Cosine-Similarity, Contradiction, Antonyms, Synonyms, Machine Learning.

1. INTRODUCTION

In general, students' academic performance is assessed on the basis of the examination, either as a subjective or objective pattern. Nowadays during this Covid-19 pandemic, everyone is working in a virtual manner. In the present scenario, manual evaluation of subjective answers is a hectic task. There are multiple systems which can evaluate objective type or MCQ questions quickly. These techniques are evaluated in machines itself after providing a pre-defined correct answers. But it helps only in competitive or objective type exam evaluation. The subjective examinations are the backbone of all the universities and board level examinations. . Large number of student attend subjective type exam. On the basis of the descriptive answer, the moderator will know how much knowledge the student has gained during his academic, on which moderator assign marks. Manual evaluation of subjective answers is a very tedious, time-consuming task and requires lots of manpower. Answer evaluation varies from moderator to moderator according to their way of evaluation, mood at time of evaluation and interrelation between student and moderator. This affects the result of the student. The project aims to automate evaluation process for subjective answer using ML and NLP.

2. Literature Survey

[1]A model to evaluate subjective Answer papers using Semi-Automated Evaluation technique is created. For that first they create Question base which contains question type, sub type, question and marks. Then Answer base is created with model answer. Evaluated answer is mapped using hash index which referred as question number. The student answer is evaluated by considering semantic meaning and length of the sentence.

[2]The proposed system is designed to evaluate answers for five students providing five different answers. The standard answer is stored in the database with the keywords, meaning and the description of that answer. Then each answer is evaluated by matching the keywords as well as its synonyms with the standard answer. It will also check the grammar and spellings of the words. After the evaluation, the answer is graded depending on the correctness of it.

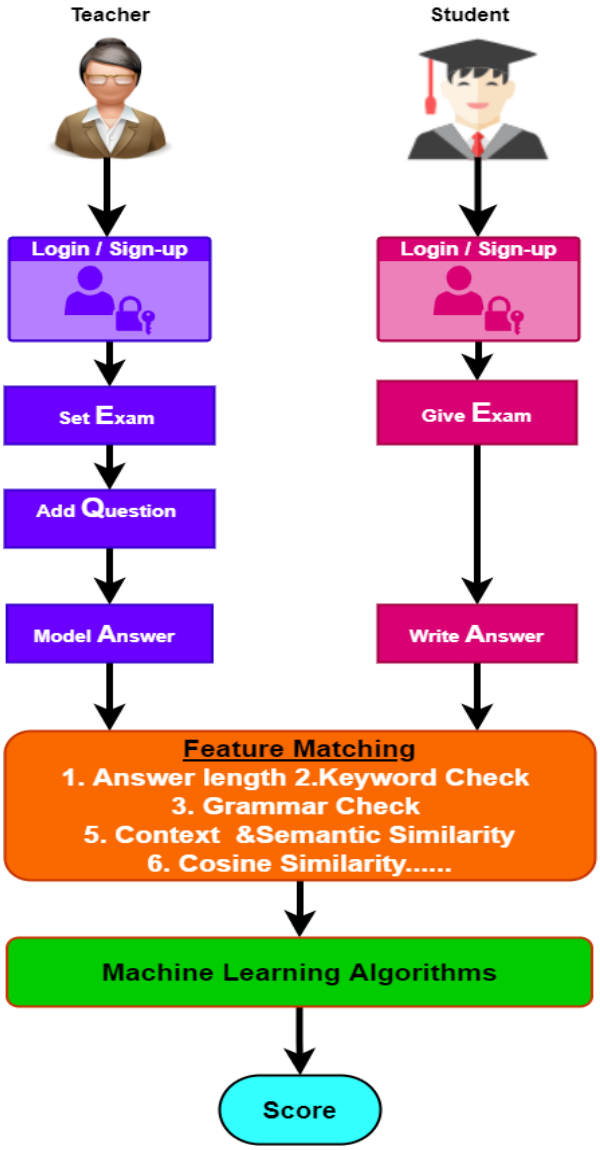
[3] E-assessment system developed to checks the answer sheet of the student and provides marks to the same. The system consists of an algorithm that compares the student's answer against three reference answers given by three different faculties and the answer with most close results and with highest precision is taken into consideration and marks are allocated accordingly. Algorithm based on TFIDF, Grammar check, WMD, cosine and Jaccard Similarity. Both the answers need not be exactly the same or word to word. This approach can be a quick and easy way for the examiners by reducing their workload.

[4]The proposed system receives solution sets from admin and student's answers. Then stop words are removed from them in order to generate keywords. After keyword generation it checks for similarity and by calculating similarity it also checks the relation of the keywords along with sentences with the sentences in the dataset which finds the exact similarity and correctness of the sentence with the datasets. If the sentences match with the datasets it generates similarity score as per the overlapping percentage. It also checks the synonyms and similar words before relating the keywords in order to increase the accuracy of the overlap. Data duplication technique is used to compare the previous answers submitted by students and on the basis of uniqueness of answers, grades are generated.

[5] Developed an algorithm which will evaluate theoretical answers and give marks according to the keyword matching which will reduce manual work and saves time with faster result evaluation. A person should collect the answer copy from the student and scan it. The machine will take the image as input and will evaluate the answer based on the length of the answer and important keywords covered Which are specified by the teacher with each answer which is to be evaluated.

[6]The handwritten text image, given as input to the handwriting conversion module is extracted and converted to machine encoded text using Optical Character Recognition (OCR) Algorithm. The machine encoded text is given as input to the evaluation module. Evaluate module evaluates the answer based on grammatical meaning of the sentences; number of keywords matched and gives marks as output.

3. Proposed System

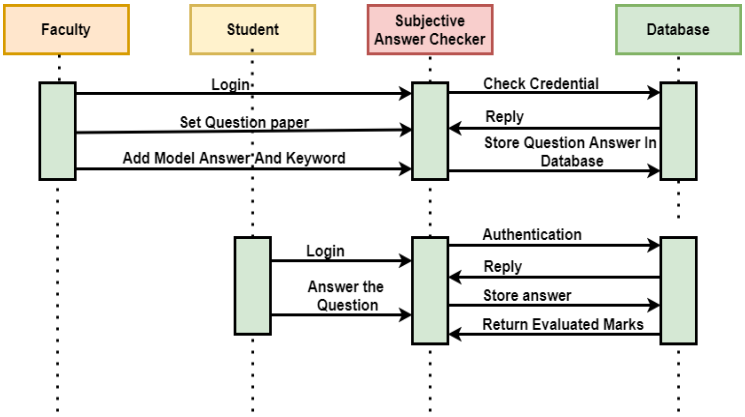


We proposed System which consist the following things: It have two sections one for teacher and the other one for the student. There is a login page for both the student and the teacher if not registered then they can create their account. In the teachers section they are able to set the paper for the exam for that they need to create a test by adding test name, date and time. After test creation teacher need to add question along with marks and model Answer for each question.

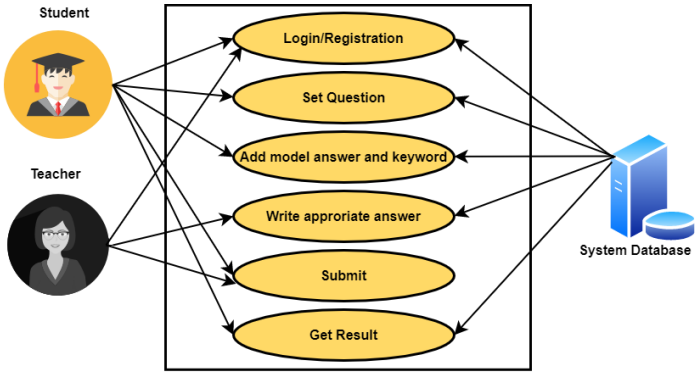
In the student section the student need to join the class to give the exam. Once the student joins the class the exam will be visible. There will be a timer where the exam will start on time and end on time. After attempting the exam the test should be submitted. After attempting the test by the student those answer will be checked with the model answer given by the paper setter. These answers will be checked on basis of the Answer length, Keyword Check, Grammar check, Context & Semantic Similarity and Cosine Similarity against model answer.
When we obtain all values for the above parameter, then it will be passed through machine learning model to obtain accurate score for student answer. In this way marks for each question evaluated ,after giving the output the marks will be allocated according to the accuracy of the student answer that will be checked to the model answer. If the teacher found that the answer is not accurate as per the model answer that was assigned by Algorithm then teacher can change the marks manually.

3.1 UML Diagram

Sequence Diagram



Use Case Diagram



Subjective Answer Checker

4. Natural Language Processing

Natural Language Processing (NLP) is a branch of AI that helps computers to understand, interpret and manipulate human languages like English or Hindi to analyze and derive its meaning. NLP helps developers to organize and structure knowledge to perform tasks like translation, summarization, named entity recognition, relationship extraction, speech recognition, topic segmentation.

1) Stemming

Stemming is the process of reducing inflection (prefix, suffix) in words to their root forms such as mapping a group of words to the same stem even if the stem itself is not a valid word in the Language. It reduces a word to its word stem that affixes to suffixes and prefixes or to the roots of words known as a lemma. Stemming will remove suffix and prefix from the word.

2) Lemmatization

To extract the proper lemma of word, it is necessary to look at the morphological analysis of each word. Rather than stemming, lemmatization provides lemma which has a dictionary meaning. The key to Lemmatization is linguistics.

3) Stopwords

Stopwords are the most common words in any natural language. For the purpose of analyzing text data and building NLP models, these stopwords might not add much value to the meaning of the document.

5. Weighting Module for Evaluation of Answer

5.1 Answer length

In the existing system, we observed that the student gets a full mark on writing only keywords. And so, answer length is also an important factor. As student may write all keywords and grammatically correct short sentences. Such answer would get full marks for keywords and grammar but less marks for short length.

5.2 Keyword Matching

Keywords is another important factor which we consider to check whether student cover all important concept (key-point) in his answer or not. If the keyword matches the maximum marks will be allotted as per the keyword considering its correctness. But if few the keywords differ with the model answer keywords then the marks will be deducted accordingly.

5.3 Contextual Similarity

We have designed an algorithms which give similarity between two answers and if two answers are not contextual similar then it will give 0 else it give percentage of similarity. We studied different python libraries which give high similarity ratio even though the answers are not contextually similar.

5.4 Semantic Similarity

If two answer are contextual similar then we have considered similarity between two answers with respect to contextual meaning by using API.

5.5 Contradiction

The Existing systems were unable to find contradicting sentences, because the system removed important stopwords (not, never, doesn't...etc.) while preprocessing and also not considering key antonyms present in the answer. We use POS tagger to find contradicting sentences from student answer against model answer.

5.6 Grammar Check

To check and detect grammatical mistakes and spelling errors Grammar, Textgear API is used in our system. When the text is sent to Grammar Textgear's API, it returns a number of grammatical error present in text

5.7 Cosine Similarity

Cosine similarity is a standard point of reference to measure how similar the documents are irrespective of their size. The similarity is represented as the dot product of two text vectors.

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

A and B are word vector

5.8 Antonyms and Synonyms

NLTK WordNet is a lexical database of English language. Nouns, verbs, adjectives and adverbs that are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. The lemmas() function of the synset. It returns synonyms as well as antonyms of that particular synset. Synonyms are considered to find semantic similarity

between two different answer having same contextual meaning. Antonyms are considered to find conflicting or contradicting sentences from student answer and model answer.

6. Machine Learning Algorithms

6.1 NAÏVE BAYES

Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is mainly used in text classification that includes a high-dimensional training dataset. Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object. The formula for Bayes theorem is given as

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

P(A|B) is Posterior probability: Probability of hypothesis on the observed event B.
P(B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.
P(A) is Prior Probability: Probability of hypothesis before Observing the evidence.
P(B) is Marginal Probability: Probability of Evidence.

It is a good and mostly used classification Algorithm. However it provides less accuracy for this model.

6.2 Decision Tree Classification

Decision Tree is a supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome. In a Decision tree, there are two nodes, which are the Decision Node and Leaf Node. Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches. The decisions or the test are performed on the basis of features of the given dataset. It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions. It is called a decision tree because, similar to a tree, it starts with the root node, which expands on further branches and constructs a tree-like structure. In order to build a tree, we use the CART algorithm, which stands for Classification and Regression Tree algorithm. A decision tree simply asks a question, and based on the answer (Yes/No), it further split the tree into subtrees.

7. CONCLUSIONS

We proposed Subjective Answer Checker System Using NLP and Machine Learning (Pariksha Software), would be helpful for online universities, schools and colleges level exam. During this COVID-19 pandemic, most of educational institutes conduct their examinations online, but these exams only contain multiple choice questions. Our Subjective Answer Evaluation software assign mark to subjective question based on Answer length, keyword matching, Grammar check, cosine similarity and Contextual similarity against Model answer provided by faculty and student answer. We also developed algorithm to find contradicting statements from student answer with model answer. Even answer written by student does not match word to word with model answer provided by teacher, but still our system capable to evaluate answer based on context.

ACKNOWLEDGEMENT

With deep sense of gratitude we acknowledge the guidance of our project guide Prof. Supriya Chaudhary. The time-to-time assistance and encouragement by her has played an important role in the development of our project. Our sincere thanks to all the Lecturers who helped us in our endeavor, gave valuable advice and thus made our journey possible.

REFERENCES

[1] Chhanda Roy, Chitrita Chaudhuri, "Case Based Modeling of Answer Points to Expedite Semi-Automated Evaluation of Subjective Papers", in Proc. Int. Conf. IEEE 8th International Advance Computing Conference (IACC), 2018, pp. 85-9.

[2] Aditi Tulaskar, Aishwarya Thengal, Kamlesh Koyande, "Subjective Answer Evaluation System", International Journal of Engineering Science and Computing, April 2017 Volume 7 Issue No.4.

[3] Saloni Kadam, Priyanka Tarachandani, Prajakta Vetaln and Charusheela Nehete,"AI Based E-Assessment System", EasyChair Preprint ,March 18, 2020.

[4] Vishal Bhonsle, Priya Sapkal, Dipesh Mukadam, Prof. Vinit Raut, "An Adaptive Approach for Subjective Answer Evaluation" VIVA-Tech International Journal for Research and Innovation Volume 1, Issue 2 (2019).

[5] Prince Sinha, Sharad Bharadia, Ayush Kaul, Dr. Sheetal Rathi ,"Answer Evaluation Using Machine Learning" Conference-McGraw-Hill Publications March 2018

[6] Rosy Salomi Victoria D, Viola Grace Vinitha P, Sathya R," Intelligent Short Answer Assessment using Machine Learning" International Journal of Engineering and Advanced Technology (IJEAT) , Volume-9 Issue-4, April 2020.

[7] Sakshi Berad, Pratiksha Jaybhaye, Sakshi Jawale,"AI Answer Verifier", International Research Journal of Engineering and Technology (IRJET), Volume: 06 Issue-01; Jan 2019.