TITLE – AUTOMATED SUBJECTIVE ANSWER EVALUATION USING MACHINE LEARNING

From KIET Group Of Institutions, Delhi-NCR,
Ghaziabad-Meerut Road, Ghaziabad-201206

Submitted By-

Ishika Aggarwal, Pallav Gautam, Gaurav Parashar

Contents

- Problem Statement
- Objective
- Introduction
- Literature Review
- Proposed Methodology
- Results and Discussions
- Conclusion

Problem Statement

- Solving the most time-consuming problem in the education sector
- Copy checking is the most time-consuming task taken by the teachers
- Almost 50% of the time of the teachers is taken in subjective evaluation
- This reduces human efforts and saves time with easy implementation
- Less accuracy in manual evaluation

Objective

- To provide a simple UI for evaluating subjective answers automatically
- The project uses a simple algorithm for the convenience of teachers.
- Unbiased and error-free evaluation of answer sheets.

Introduction

- Educational assessments are very important in the learning process for students. The evaluation and scoring is a tedious process that also takes most of the valuable time of professors.
- In this work, we present a system where the answer sheets are evaluated automatically, Automatic Subjective Evaluator.
- We used previous research to produce a solution where the evaluations can be done automatically and with precision and with the least error.
- We assessed the problems of evaluators and reduced them to almost 90% saving them a huge amount of time
- We converted handwritten notes to text and generated an engine that would take those answer texts and question papers as input and then evaluate and give marks.
- The result we get is on a scale of 0 to 9 then convert them intelligently to percentages

Literature Review

- An automatic subjective answer sheet evaluator, also known as an automatic essay grader or automated scoring system, is a tool that uses NLP(Natural Language Processing) and algorithms of machine learning to evaluate and grade written responses to open-ended questions. These tools are primarily used in educational settings to grade student essays, but they have also been used in other areas such as language proficiency testing.
- Research has shown that automatic subjective answer sheet evaluators are able to grade written responses with a high level of accuracy. Studies have found that the scores generated by these tools are very close to those produced by human graders.
- If we agree with these kinds of assumptions, there is an automatic trust that numerical measures, like test scores, show the actual reality of the condition and that the result of education is the outcome of the individual working and is not affected by greater societal contexts or family occurrences.
- Generally, many of the AWE technologies use (a) Natural Language Processing tools to extract linguistic, syntax-related, semantics-related, or different attributes of text related to writing quality and (b) statistics-related or machine-learning algorithms to create scores and feedback based on patterns discovered among those features

- The goal of NLP of to do human-like processing. Originally the field of NLP was called Natural Language Understanding (NLU). Some of the jobs of NLU were - paraphrasing an input text and translating the text into another language.
- An automated answer sheet evaluator helps students get automated scores and feedback, improving their work and writing both in terms of looks and content. It is equally beneficial for the teachers as it gives them the power to monitor and access students more efficiently.
- The answer is scored using the cosine similarity. For every question, there is an answer that the system gets itself from search engines like Google or provided by the teacher. Then both the answers, given by students and by the teacher or the search engine are sent forward to the engine for summarization and other evaluations. Then the similarity between the summarized text is calculated using some algorithms.

Proposed Methodology

- In this proposed model, the data set includes typed answers and paragraphs taken from students themselves and from various websites and blogs for comparison.
- The data set needs to be in massive amount, so it comes from students through google forms, we also targeted various websites that contain subjective questions answer to train our model.
- In our proposed system, we compare the marks given by a machine and a human. Evaluation is done by the algorithm proposed and assign the marks.
- The overall system consists of five sub-processes. All these five sub-process have their own working and methodology based on specific techniques.

- The idea behind the machine evaluation is that firstly we summarize the subjective question answers written by students and find the first ground truth.
- Then we search for the actual answer to the subjective question through various websites and platforms in order to find out the second ground truth. This process of extracting the text from URL is called Web Scraping.
- A comparison is made, and the similarity Index is calculated between both the ground truths obtained using the Cosine Similarity Index.
- Now, the classes are made based on some defined ranges. Using the value obtained from Cosine similarity, we decide the class and marks are given to the student's answer.
- After the machine evaluation, the same answers are evaluated by a teacher, and marks are given. Both the marks are compared in order to find the correctness of the proposed system.

Results and Discussion

- Division of classes in ranges and allocating Marks : To allocate the marks after calculating the cosine similarity, Classes are made based on the ranges on a scale of 0 to 9. These classes tell us the category of answer that whether it is Good, Very Good or Excellent.
- Calculation of Cosine similarity: Using the proposed system, cosine similarity is obtained using the data set from various students. The Cosine similarity is calculated between two tokenized texts. Approximate marks or score is given based on the comparison made between the student's answer and the actual answer. Cosine similarity is always found in the range [-1,1].
- Comparison between Human Evaluation and Proposed System: In this proposed system, we first show the result obtained by the algorithm developed using the tools and technology. After that, we also evaluated the student's responses manually with the help of teachers in order to check the consistency of the result. This step plays an important role in checking the accuracy of the proposed system.

THANK YOU