An NLP Based Algorithm for Subjective Answer Checking

Pratik Manoj Desai
desaipm18.comp@coep.ac.in
College of Engineering, Pune, Maharashtra

Rhugaved Rajendra Narmade
narmaderr18.comp@coep.ac.in
College of Engineering, Pune, Maharashtra

ABSTRACT

In the current pandemic situation, where not only the classes, but also the exams are being conducted in online mode, an artificial intelligence-based agent can evaluate the subjective type answers not only efficiently, but also fairly with considerably good accuracy. Using Natural language processing helps in reducing the manual labor for the evaluation of hundreds of similar answers and makes the process easier and faster. An agent to check answers automatically and provide an evaluation based on certain parameters are the basis for this paper. The tutor/teacher gives certain inputs to the agent which are looked for in the answers and accordingly the answers are graded by the agent. The grades given by the agent are as good as the grades given by a human teacher. In this paper, we've proposed a technique using Natural Language Processing to evaluate subjective answers.

Keywords— Automatic System, Automatic Subjective Answer Checking, Automatic Scoring, Automatic Evaluation, Artificial Intelligence, NLP

1. INTRODUCTION

In today's day and age, Artificial Intelligence has made radical changes and improvements in almost every field it has touched. Such enhancements have also been made in the field of Answer Checking over the past few years. CAA, also known as Computer Assisted Assessment is the term used to refer to the assessment carried out by the use of computers. The use of computers has changed the field of learning/education a lot over the past decades. In the current era, where the entire world is affected by the Pandemic, the use of computers in learning is ever more visible and is a need of the hour as offline/face-to-face learning is not possible. Due to online-only teaching-learning, the mode of examinations has also changed from pen-paper classroom-based examinations to online computer-based examinations. Although the CAA-based systems, which were started nearly 50 years ago and developed over the years, are capable of evaluating short textual answers like multiple choice questions, visual identification, short answers, matching the pairs, etc. Researchers in the teaching-learning field agree that evaluation of one's understanding and application of concepts is a very complicated and incomplete approach by simply using objective-type questions. The ability to recall, organize and express ideas, understanding, and application of learned concepts, etc., is possible mainly through descriptive or essay-type questions supplemented by objective-type questions. Using descriptive questions helps in better evaluation and serves the most useful purpose. Many researchers have claimed that the evaluation of essays by human graders and by assessment tools have a great variation in the grades awarded to the answers. As specific concepts are considered in evaluating an answer, only the presence of that concept leads to grades, else the answer is marked as incorrect. The system proposed here solves this problem.

Evaluation of answers at a higher level than that of Bloom's taxonomy with the consideration of assessing descriptive answers is the purpose to present the new system. We also discuss a variety of CAA techniques in this paper. Also, how Computer Assisted Assessment is approaching currently and its utilization for designing new systems.

There are three main techniques for automatic marking of texts: Natural Language Processing (NLP), Statistical Technique, Information Extraction (IE) Technique.

A. Full Natural language processing (NLP): Using the semantic meaning of the answer after we parse the answer and compare it with the teacher's answer to assign the grades

B. Statistical Technique: This method involves matching specific keywords that are given by the teacher in a list format to the student's answer. As this method does not take into consideration synonyms and antonyms, or the word order, it's considered a weak method

C. Information Extraction (IE) Technique: This involved breaking of answers into concepts and relationships which are then compared with the teachers' answers to give out grades.

2. PROBLEM DEFINITION

There are many tools on the internet that help to evaluate the answers. But the tools and software are for multiple-choice questions or one-line responses. Descriptive answers do not get checked using these tools that are available. Descriptive answers are very important to test the depth of knowledge. These existing solutions don't have a way to check answers grammatically or check any spell errors. There is also no way to check if the word order of the answer is different. To avoid all these cons of the existing solutions the proposed system will evaluate the descriptive answers by taking the word order as well as the sentence sequence into consideration. The spell errors will also be checked through this system. This proposed solution will also provide feedback for the answer.

3. PROPOSED SYSTEM

The proposed solution will use Computer Assisted Assessment for the evaluation of descriptive type answers. This system will also consider word order while checking along with the sequence of sentences. The answer to a particular question is answered differently by different students. So, this proposed system will have the primary task of finding out the semantic meaning of the answer. The following is the architecture of the system:

1. Teacher: The teacher will form questions along with expected answers here. He will also add some additional accepted words for the answers. There will also be a weight assigned to those words.
2. Student: The questions will be displayed here. The students will answer the questions which will be evaluated by the system and the students will get feedback.
3. Evaluation Module: The answer input by the student as well as the expected answers from the teacher will be processed here. Several NLP techniques will be used here. Some steps will be like tokenization.
4. Grading Module: Here the text processing will take place. Every word of the response will be compared with the expected answer along with several other synonyms. Weights will be used to evaluate how close the answer is to the expected answer. Accordingly, the grades will be assigned.
5. Display Module: All the scores will be accumulated and displayed to the student along with the feedback.

Algorithm Steps:
1. START
2. Take the correct answer input from the Teacher and store it somewhere.
3. Identify important words and assign them weights depending on their importance.
4. Find out Synonyms and antonyms for those words and store them.
5. Take the response from the student and store it elsewhere.
6. Initialize a score variable and set it to 0.
7. Find out the keywords from the response and match it with the keywords extracted from the correct response.
8. If keywords match then add the weight to score variable
9. If words don't match check for a match in the synonyms table.
10. Else assign 0 for the wrong answer.
11. Next check for nouns and verbs as well as word order.
12. Check the meaning of the sentence if that makes sense or not.
13. Also, check for sentence sequences.
14. Deduct some marks for incorrect sentence sequence.
15. Set a threshold for correct and incorrect answers.
16. If the sum of all weights is below the assigned threshold the response is incorrect.
17. Give feedback on the answer to the student.
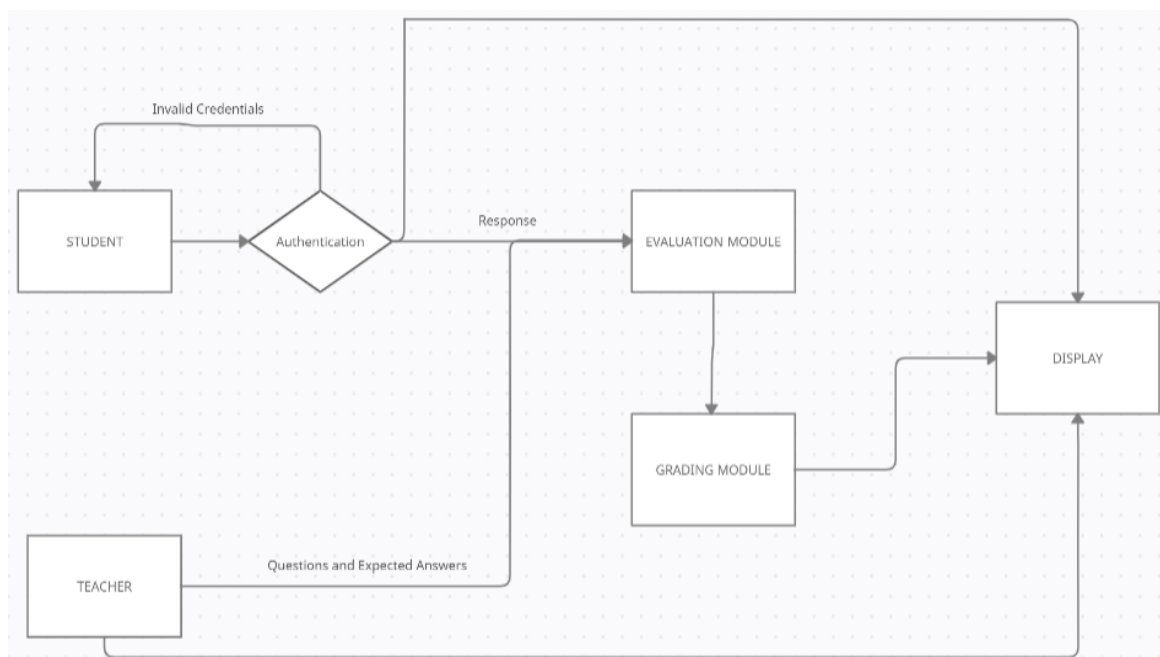
4. CONCLUSION

Computer Assisted Assessment has been one of the most interesting research areas in the field of answer evaluation since the 1970s. CAA has been proven to evaluate answers accurately and efficiently and provide an alternative to manual checking. When compared with human evaluators, most CAA systems provide promising results. Although CAA systems are fairly accurate and efficient, the existing systems are only capable of evaluating answers at a lower level like that of short answer questions or objective questions and not of long answer questions. The system we propose in this document tries to overcome this shortfall of existing CAA systems by evaluating answers at a higher level like that of long answer questions or descriptive type questions with multiple sentences. The system considers the multiple sentences collectively and evaluates them further. Other capabilities of the system include spell and grammar checking in the answers. On completion of the evaluation, the system will provide a report with feedback to the students

5. REFERENCES

[1] Perez D.: Adaptive Computer Assisted Assessment of free-text students' answers: an approach to automatically generate students' conceptual models. PhD thesis, Escuela Polit´ecnica Superior, Universidad Aut´onoma de Madrid (2007)

[2] Marina Nakano, Emi Masuda, Masaru Kamada, "A Structure Editor for the English Language", in Proc. Int. Conf. on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016, pp. 155-161

[3] K. Gaurav and Pradeep Kumar Bhatia, "Analytical Review of Preprocessing Techniques for Offline Handwritten Character Recognition", Int. Journal of Advances in Engineering Sciences, vol.3, no. 3, 2013, pp. 14 -22.

[4] Prince Sinha, Sharad Bharadia and Ayush Kaul, "Answer Evaluation Using Machine Learning", March 2018.

[5] Sk Asif Akram, Mousumi Saha and Tamasree Biswas, "Evaluation of descriptive answer sheet using machine learning", in International Journal of Engineering Sciences & Research Technology (IJESRT), April 2019, pp. 184-186.

[6] Yuan, Zhenming, et al. A Web-Based Examination and Evaluation System for Computer Education. Washington, DC: IEEE Computer Society, 2006.

[7] Laurie Cutrone, Maiga Chang and Kinshuk, "Auto-Assessor: Computerized Assessment System for Marking Student's Short-Answers Automatically", IEEE International Conference on Technology for Education,2011

[8] Mitchell T, Russell Broomhead P, and Aldridge N. 2002, "Towards robust computerized marking of free-text responses." In proceeding of the 6th computer Assisted Assessment Conference, 2002.

[9] V. Lakshmi and Dr. V. Ramesh, "Evaluating students descriptive answer using natural language processing and artificial neural networks", in International Journal of Creative Research Thoughts (IJCRT), Volume 5, Issue 4, December 2017, pp. 3168-3173.