

Automatic Evaluation of Descriptive Answers

Supriya Sanjay Sasane¹, Sayali Kiran Joshi², Aishwarya Gulab Thorat³

Student¹, Student², Student³
Department of Information Technology,
Bharati Vidyapeeth College of Engineering
Kharghar, Navi Mumbai

Abstract: Evaluation systems at school, college, and University levels are in Objective and Subjective form as well. Systems for automatic evaluation of the objective type of tests already exist. Automation evaluation of descriptive answer is aimed at fast processing, less manpower, ease in record keeping and extraction. Our approach is mainly based on keyword match, sequence match, and quantitative analysis. Firstly, we will extract the handwriting from the student's answer sheet in a Portable Document Format (PDF) form and then convert it to a text file. The system will evaluate the answer by extracting the important keywords from the student's answer and compare those keywords with the predefined set of solutions. If any word in student's answer does not match with teacher answer then a synonym of that unmatched word is checked in the dictionary if synonym of that word match with teacher answer then an appropriate assessment of marks will be given to the student. The similarity score will give us the parameter to judge or evaluate the degree of correctness in the student answer.

Keywords - NLTK tool, PyQt4, Tokenizing, stemming, filtering, Descriptive answers.

I. INTRODUCTION

With the growth in population and necessity of education, it is proving hard for assessors to check the relevance and accuracy of the answers written by students. Automatic evaluation of descriptive answer would be beneficial for the universities, schools, and colleges for the academic purpose by providing ease to faculties and the examination evaluation. So, our aim is to develop a system that will automatically check the descriptive answers based on keyword match, sequence match, and quantitative analysis. The proposed model aims to provide proper relevant labels or attributes which will enhance the precision of scoring. The system will calculate the accuracy of the answer by comparing these attributes with the original solution. Also, the proposed system provides User Interface (UI) for simple user interaction and understanding. The system will also consider the case of synonyms to evaluate the answers more effectively, If any word in student's answer does not match with teacher answer then a synonym of that unmatched word is checked in the dictionary if synonym of that word match with teacher answer then appropriate assessment of marks will be given to the student. If the density of matching is more than or equal to 50% then the answer was termed to be correct and it is known as positive confidence else negative confidence is provide for the mismatched and the answer is termed to be wrong. In School, Colleges and University Examination System objective and subjective type questions are assessed. The evaluations of the papers are very lengthy and time-consuming. Existing systems are only used for evaluation of objective answers which is not the best way for evaluation of one's subject understanding & knowledge. Exam reports and mark sheet takes more effort and physical space Report Card and Mark sheet utilize a lot of papers. This is often an area where significant money can be saved by automation as marks can be obtained on the line along with score card. The project aims to provide the system which will eradicate the inconsistency in the manual assessment. Descriptive answers have specific domain related questions and evaluation can be an easier task because of NLTK tool. Some aimed systems only focus on fast processing. Variation of lengthy words, a form of sentences all matters while evaluating the answer.

II. RELATED WORK

Evaluation of descriptive answers requires lots of manual efforts. It also consumes more time of the teachers to evaluate the answer sheets. Till now there have been many researches done by various authors. The author named Mena. K and Lawrence. R, tried to solve the problem in An automatic evaluation system which was used to make sure the confidentiality and neutrality of evaluation. An algorithm called Assessment algorithm was used which used semantic similarity for evaluation of detailed type answers. It eradicated the inconsistency in the manual assessment. Also, for preprocessing, an algorithm for pruning and stemming was used to reduce the size of the descriptive type answers. Stemmed words were converted to vectors using the semantic method, Latent Semantic Analysis (LSA).[1] The vectors obtained from the semantic method were clustered using the Self-organizing map. Based on the value returned by the similarity measure, marks were awarded. In the system proposed by authors of [2], implemented CAA for descriptive type answer. The previous systems checked single line text response without considering word order. So this system tried to avoid this problem by considering the collective meaning of multiple sentences. The primary focus of this newly system was to determine the semantic meaning of student answer with a consideration that student responses to question in number of ways. The system basically focused on multiple sentence response. It also used various techniques for automatic marking of text named Statistical Technique, Information Extraction and Full Natural Language Processing. It consists of various modules like Student Module, Tutors module, Preprocessing module, Answer Comparison and grade Assignment module. Later on, the paper named Pattern Matching Algorithm proposed by authors of [3], represented an approach to check the degree of learning of the student, by evaluating their descriptive exam answer sheets. By representing the descriptive answer in the form of graph and comparing it with the standard answer are the key steps. In this approach pattern matching algorithm was used for evaluation of the answer. Considering the structure of the text in natural language processing most of the work has been done for morphological and syntactic analysis but semantic, pragmatic and discourse were not explored previously which were introduced in this paper by the authors. Grammatical mistakes were ignored in this system. Mousumi Saha proposed a system which represented the concept of Intelligent Tutoring System using AI is to build any computer program that contains some intelligence and aims at providing immediate and useful instructions to novice learners. It was a system

that performed teaching the students and evaluating their performances with proper feedback. The system worked on the student's knowledge level and psychology. It checked the answers with appropriate spelling and grammar under restriction. It worked on the respected answer given and distinguishes them whether correct, error or elaboration. This intelligent online exam system detected the problems to a large extent and helped the instructors to decide knowledge. It measured the correctness of descriptive type answers with spelling checking and grammar checking under restricted condition [4]. Later the authors of [5] proposed a system which was capable of evaluating the descriptive answers. The online examination system checked the student answer by matching the answer with a predefined set of answer. The predefined answers were saved on the server and evaluation was done automatically using the automatic assessment tools. The machine learning approach was used to solve this problem by using text mining. Measuring the similarity between, sentences, words, documents, and paragraphs was an important component in various tasks such as text summarization, information retrieval, automatic essay scoring, document clustering, and machine translation and word-sense disambiguation. In this system JSON is used for transferring data between web application and server, serving as an alternative to XML.

COMPARISON OF EXISTING SYSTEM

Table 2.1 shows the comparison of existing systems which consist of four main parameters namely:

	[1] Meena.K	[2] Shweta Patil	[3] Pranali Nikam	[4] Mousumi Saha
Grammatical mistakes	Not Checked	Checked	Not Checked	Checked
Word Order	Was Considered	Was not Considered	Was not Considered	Was not Considered
Evaluation of detail answers	Yes	Yes	Yes	No
Algorithm Used	Assessment Algorithm	Natural Language Processing	Pattern Matching Algorithm	Machine Learning approach

table 2.1 comparison of existing systems

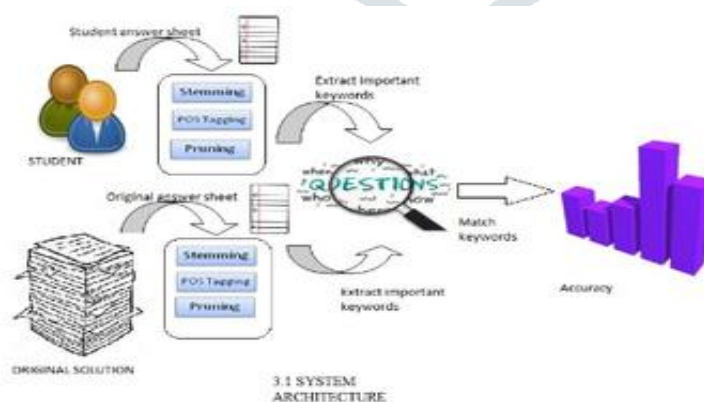
- Grammatical Mistakes: Grammatical mistakes in the sentences
- Word Order: Sequence of words in the sentence
- Algorithm Used: Algorithm used for evaluation

In the systems [1] and [3], Grammatical mistakes were not considered while in [2] and [4] it was considered. The system proposed in [3] uses two approaches for Pattern Matching Algorithm. Word Order in [2],[3] and [4] were not considered whereas in [1] the word order was Considered. Algorithms used in [1],[2],[3],[4] were Assessment Algorithm, Natural Language Processing, Pattern Matching Algorithm, and Machine Learning approach was used respectively.

III. PROPOSED SOLUTION

The proposed model aims to provide proper relevant labels or attributes which will enhance the precision of scoring. The system will calculate the accuracy of the answer by comparing these attributes with the original solution. Also, the proposed system provides a User Interface (UI) for simple user interaction and understanding. The system will also consider the case of synonyms to evaluate the answers more effectively, If any word in student answer sheet does not match with teacher answer then a synonym of that unmatched word is checked in the dictionary if synonym of that word match with teacher answer then appropriate assessment of marks will be given to the student. If the density of matching is more than or equal to 50% then the answer was termed to be correct and it is known as positive confidence else negative confidence is provide for the mismatched and the answer is termed to be wrong.

Figure 3.1 describes the proposed architecture. The student and original solution are present in the system where student answer sheet and original answer sheet are the main components which will be used for further preprocessing techniques



A. Pruning, POS Tagging, Stemming.

The student answer sheet and original answer sheet is used for preprocessing techniques like Pruning, POS Tagging and Stemming which extract the important keywords. Firstly all the punctuations will be removed from the sentences and then some common words will be removed and we will be left with some important keywords. Now a similar process will take place with the original answer sheet which will give us important keywords through pruning, stemming and POS tagging.

IV. PROPOSED SOLUTION

A. Student Answer Sheet

The student writes the answer on the answer sheet the Tesseract will be used to detect handwriting and convert it into pdf format which will then be converted to a (.txt) file

- *Tesseract:*

Tesseract is an OCR engine with support for Unicode and the ability to recognize more than 100 languages out of the box. It can be trained to recognize other languages. Tesseract is used for text detection on mobile devices, in the video, and in Gmail image spam detection.

B. Teacher Answer Sheet

The Teacher answer sheet is stored in the system.

The NLTK Tool has some preprocessing techniques that will be applied to this file which will remove punctuations if any, all normal words and output will be important keywords that will be compared with the original solution important keywords to get accuracy.

- *NLTK Tool:*

The Natural Language Toolkit, or more commonly NLTK is a suite of libraries and programs for symbolic and statistical natural language processing (NLP) for English written in the Python programming language. It contains text processing libraries for tokenization, parsing, classification, stemming, tagging and semantic reasoning. It also includes graphical demonstrations and sample data sets as well as accompanied by a cook book and a book which explains the principles behind the underlying language processing tasks that NLTK supports.

After extracting the handwriting from the student's answer sheet in a Portable Document Format (PDF) form and then convert it to a text file. After that System will evaluate the answer by extracting the important keywords from the student's answer and compare those keywords with the predefined set of solutions. For extracting these keywords Natural Language Toolkit (NLTK) is used to perform some preprocessing techniques like tokenizing, stemming, pruning, POS (Part Of Speech) tagging and removing stop words from student's answer.

V. RESULT

For every answer being entered it is important to reward points based on the usage of words and their relevance. If any word in student's answer does not match with teacher answer then synonym of that unmatched word is checked in dictionary if synonym of that word match with teacher answer then appropriate assessment of marks will given to the student. The similarity score will gives us the parameter to judge or evaluate the degree of correctness in the student answer.

CONCLUSION

In colleges, universities and various other educational institutes' examinations play a very vital role. Many aimed systems include online evaluation. But, these systems only include multiple -choice questions, which are efficient in checking the student's aptitude skills, in contrast, they fail to determine the theoretical knowledge a student possesses. The proposed system attempts to calculate the subjective answers. The proposed system calculates the student's answer based on the keywords. By judging against the original answer sheet keywords and the student's answer sheet accuracy is assigned to the student. Thus the proposed system could be of great effectiveness to the educational institutes, as it saves time and the trouble of checking bundles of papers. In future, a system can be developed to evaluate subjective answers with diagrams and mathematical expressions. In colleges, universities and various other educational institutes' examinations play a very vital role. Many aimed systems include online evaluation. But, these systems only include multiple -choice questions, which are efficient in checking the student's aptitude skills, in contrast, they fail to determine the theoretical knowledge a student possesses. The proposed system attempts to calculate the subjective answers. The proposed system calculates the student's answer based on the keywords. By judging against the original answer sheet keywords and the student's answer sheet accuracy is assigned to the student. Thus the proposed system could be of great effectiveness to the educational institutes, as it saves time and the trouble of checking bundles of papers. In future, a system can be developed to evaluate subjective answers with diagrams and mathematical expressions.

REFERENCES

- [1] Meena.K, Lawrance.R "Semantic Similarity Based Assessment of Descriptive Type Answers"978-1-4673-8437-7/16/\$31.00 ©2016 IEEE
- [2] Shweta Patil, Sonali Patil "Evaluating Student Descriptive Answer By Natural Language Processing" International Journal of Information & Technology (IJERT) Vol.3, Issue 3 March-2014
- [3] Pranali Nikam, Mayuri Shinde, Rajashree Mahajan and Shashikala Kadam "Automatic Evaluation of Descriptive Answer Using Pattern Matching Algorithm" International Journal of Computer Sciences and Engineering Vol.-3(1), PP(69-70) Jan 2015
- [4] Mousumi Saha, Mainak Chakraborty, Tamasree BiswasSaha "A Novel Approach for Descriptive Answer Script Evaluation" International Journal of Advanced Research in Computer Science and Software Engineering 6(3), March - 2016
- [5] Pooja Kudi, Amitkumar Manekar, Kavita Daware, Tejaswini Dhattrak "Online Examination with Short Text Matching"2014 IEEE Global Conference on Wireless Computing and Networking (GCWCN)