Scholars Journal of Engineering and Technology

OPEN ACCESS

Abbreviated Key Title: Sch J Eng Tech
ISSN 2347-9523 (Print) | ISSN 2321-435X (Online)
Journal homepage: https://saspublishers.com

An Efficient Approach for Automated Scoring of Short Answers in Punjabi Language

Dr. Pawan Kumar[1*], Dr. Amanpreet Kaur[2], Ms. Jaspreet Kaur[2]

[1]Associate Professors IT, Bhai Gurdas Institute of Engineering & Technology, Sangrur, Punjab
[2]Assistant Professors IT, Bhai Gurdas Institute of Engineering & Technology, Sangrur, Punjab

DOI: 10.36347/sjet.2022.v10i11.001 | **Received:** 17.10.2022| **Accepted:** 26.11.2022 | **Published:** 30.11.2022

***Corresponding author:** Dr. Pawan Kumar

Abstract **Original Research Article**

With the emergence new technology, the process of question answering is the main field of research of text mining. The system will assist classroom assessment and help to overcome time, cost, reliability, and generalizability issues in writing (typed) assessment. Reacting manually to student papers is a burden for teachers. Particularly if they have number of students and if they assign frequent writing assessment, providing individual feedback the student essays might be time consuming. Automated system can be very useful because they can provide the student with a score as well as feedback within seconds. Automated scoring of descriptive answers can find applications in educational assessment and is one of the applications of Natural Language Processing.

Keywords: NLP, Punjabi, Question, Answer, AES.

I. INTRODUCTION

As the requirement of information is essential part of our life [1]. There are numerous sources of information, but the major one is database. Database helps us to store, access and retrieve information. No organization or industry is possible without the use of database. Each and every computer-based application need to access information from database that requires knowledge of formal query language like SQL. But it is not possible for everyone to learn or write SQL queries. To overturn this problem many researchers have brought out to use Natural Language (NL) i.e. Punjabi, Arabic, English, Bengali etc. in place of formal query language which can be a perfect interface between an application of computer and nontechnical user. This idea of using NL has induced the development of new sort of processing method in database systems.

AES AND NLP

NLP have major tasks such as discourse analysis, morphological segmentation, parsing, word sense disambiguation and information extraction etc [2]. Automated Scoring can choose some tasks from NLP for scoring process. Automated Scoring systems (AES) are a combination of various techniques such as – NLP (Natural Language Processing) along with, Statistics, Artificial Intelligence (Machine Learning), Linguistics and Web Technologies, etc. Today, Automated Scoring is still a difficult, intricate and interesting issue for researchers in artificial intelligence and natural language processing though many English Automated Scoring systems have been proposed and developed but with little success. Automatic essay-scoring techniques are inappropriate for scoring the content of an essay because they either rely on grammatical measures of quality or machine learning techniques, neither of which identifies statements of meaning (propositions) in the text.

QUESTION ANSWERING

Research and development of systems capable of answering questions in natural language dates back to 1959 [3], but the notion of a question answering system was born in 1950, when Turing offered a solution to the question of whether or not machines can think. He proposed a task he called an Imitation Game," which has eventually become known as the famous Turing Test, in which a human communicates with a machine via a teletype interface and asks questions of it.

Mainly Question Answering (QA) is becoming an increasingly important research area in natural language processing. Since 1999, many international question answering contests have been held at conferences and workshops, such as TREC, CLEF, and NTCIR. Basically, QA is special task to search natural language answers from natural language questions. The

main task of QA is providing a short answer to a natural language query supported by a document in an underlying document collection. Question answering aims to develop techniques that can go beyond the retrieval of relevant documents in order to return exact answers to natural language questions, such as "How tall is the Eiffel Tower?", "Which cities have a subway system?", and "Who is Alberto Tomba?". Answering natural language questions requires more complex processing. From last many decades various researchers are working with this domain. Question Answering (QA) systems take in a natural language question, and return the answer from the set of documents. Answer retrieval, rather than document retrieval, will be integral to the next generation of search engines. Currently, there is a web site called AskJeeves that attempts to retrieve documents to answer a question. Question answering systems will handle query creation, and finding the exact entity that is the answer.

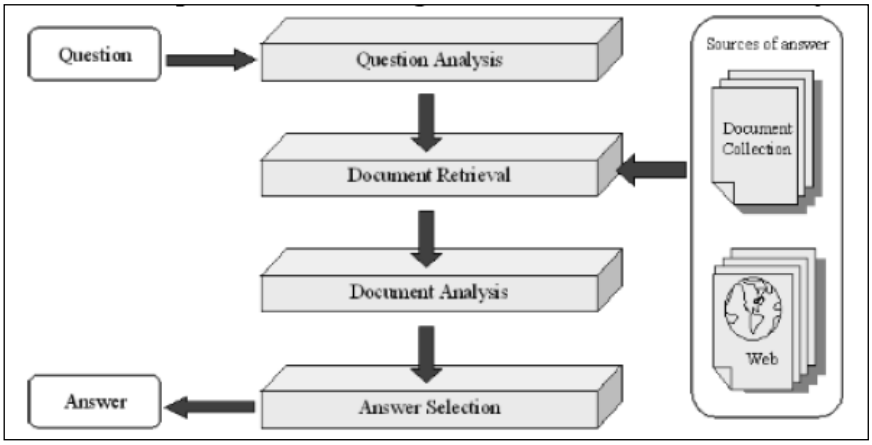

Figure 1: The General Approach to QA [3]

Figure 1 displays the main components, of such a general architecture, and the ways in which they interact. The prototypical system has four components: question analysis, document retrieval, document analysis, and answer selection.

AUTOMATED SCORING FOR ANSWERS

Assessment [4] is one of the most important things in studying. In this digital era, there are a lot of systems that have developed to handle assessment automatically. One of the system assessments that were developed by researcher is automatic scoring for essay. There are two types of essay; long and short answer essays. Our research is focused on the development of automatic short answer scoring. Some automatic scoring systems used on long answer have shown optimal results in giving a score on the students answer. Automatic long answer systems use the information retrieval method to measure similarity between students answer and references answer. Automatic short answer scoring does not give the best result yet. Short answer has a limited word in each answer. Each answer consists of one phrase to three sentences. Assessment of the short description that has limited number of words requires special handling, especially in the weighting process. With the limitations of the process of weighting the word, it cannot be done with frequency model, because the words occurrence is very rare.

The process of measuring student achievement through evaluation is important in a learning process. These measurements will determine the student's ability in absorbing information during the learning process. Type of assessment for measuring the student's ability can be objective or subjective test. The example of objective test is multiple choices, while short or long is the example of subjective test. The advantages of applying subjective test during the evaluation is being able to measure the ability of students in higher order thinking levels. The utilization of information technology in the implementation of evaluation enables the massive measurement of learning outcomes and provides fast and consistent results. Some researchers have specifically developed a system used to perform automatic assessment on the subjective test. There are two types of automatic scoring system of subjective test; AES (Automated Essay Scoring) and ASAS (Automated Short Answer Scoring). AES refers to the automatic correction system for the long description or explanation of open question, while ASAS refers to the automatic correction for the short description or explanation of close question. Basically, the difference between AES and ASAS is in the length and focus of assessment. The length of the sentence in ASAS is between one phrase to four sentences, and the maximum number of words is 100.

II.RELATED WORK

Madhvi Soni *et al*., [6] explained in the paper that Grammar checking is the task of detection and correction of grammatical errors in the text. English is the dominating language in the field of science and technology. Therefore, the non-native English speakers must be able to use correct English grammar while reading, writing or speaking. This generates the need of automatic grammar checking tools. So far many approaches have been proposed and implemented. But less effort have been made in surveying the literature in the past decade. The objective of this systematic review is to examine the existing literature, highlighting the current issues and suggesting the potential directions of future research. This systematic review is a result of analysis of 12 primary studies obtained after designing a search strategy for selecting papers found on the web. They present a possible scheme for the classification of grammar errors. Among the main observations, we found that there is a lack of efficient and robust grammar checking tools for real time applications. They present several useful illustrations most prominent are the schematic diagrams that they provide for each approach and a table that summarizes these approaches along different dimensions such as target error types, linguistic dataset used, strengths and limitations of the approach. This facilitates better understandability, comparison and evaluation of previous research.

V. V. Ramalingam *et al*., [7] presented in the paper that essays are paramount for of assessing the academic excellence along with linking the different ideas with the ability to recall but are notably time consuming when they are assessed manually. Manual grading takes significant amount of evaluator's time and hence it is an expensive process. Automated grading if proven effective will not only reduce the time for assessment but comparing it with human scores will also make the score realistic. The project aims to develop an automated essay assessment system by use of machine learning techniques by classifying a corpus of textual entities into small number of discrete categories, corresponding to possible grades. Linear regression technique will be utilized for training the model along with making the use of various other classifications and clustering techniques. They intend to train classifiers on the training set, make it go through the downloaded dataset, and then measure performance their dataset by comparing the obtained values with the data set values. They have implemented their model using java.

Ali Doğan *et al*., [8] explained in the paper "Automated Essay Scoring versus Human Scoring: A Reliability Check" that new materials have continuously been added to the assessment instruments in ELT day by day. The question of whether writing assessment in ELT can be done via E-Rater® was first addressed in 1996, and this system, which is commonly called "Automated Essay Scoring Systems" in especially America and Europe in recent years, has taken part in the field of assessment instruments of ELT with steady development. The purpose of this study is to find out whether AES can supersede the writing assessment system that is used at The School of Foreign Languages at Zirve University. It is performed at The School of Foreign Languages at Zirve University. The participants of the study were a group of 50 students in level C that is the equivalent of B1. The beginning of the quantitative study includes the assessment of essays written by C level students at The School of Foreign Languages at Zirve University by three human raters and E-Rater. After the study it was found that the writing assessment has been currently used at The School of Foreign Languages at Zirve University costs more energy, more time and it is more expensive. Thus, AES was suggested for use at The School of Foreign Languages at Zirve University, which has proven to be more practicable.

Shihui Song *et al*., [9] describe in the paper that they built an automated essay scoring system to score approximately 13,000 essays from an online Machine Learning competition Kaggle.com. There are 8 different essay topics and as such, the essays were divided into 8 sets which differed significantly in their responses to the features and evaluation. Their focus for this essay grading was the style of the essay, which is an extension on the studies, conducted determining the quality of scientific articles by adding maturity to the feature set (Louis and Nenkova, 2013). An aspect of this project was to recognize the difference between the advanced nature of scientific articles to the coherency of middle to high school test essays. They evaluated Linear Regression, Regression Tree, Linear Discriminate Analysis, and Support Vector Machines on our features and discovered that Regression Trees achieved the best results with k = 0:52.

Salvatore Valenti *et al*., [11] presented an overview of current approaches to the automated assessment of free text answers. Ten systems, currently available either as commercial systems or as the result of research in this field, are discussed: Project Essay Grade (PEG), Intelligent Essay Assessor (IEA), Educational Testing service I, Electronic Essay Rater (E-Rater), C-Rater, BETSY, Intelligent Essay Marking System, SEAR, Paperless School free text Marking Engine and Auto mark. For each system, the general structure and the performance claimed by the authors are described. In the last section of the paper an attempt is made to compare the performances of the described

systems. The most common problems encountered in the research on automated essay grading is the absence both of a good standard to calibrate human marks and of a clear set of rules for selecting master texts. A first conclusion obtained is that in order to really compare the performance of the systems some sort of unified measure should be defined. Furthermore, the lack of standard data collection is identified. Both these problems represent interesting issues for further research in this field.

III. METHODOLOGY

The methodology adopted for preparing this software can be classified through following steps:

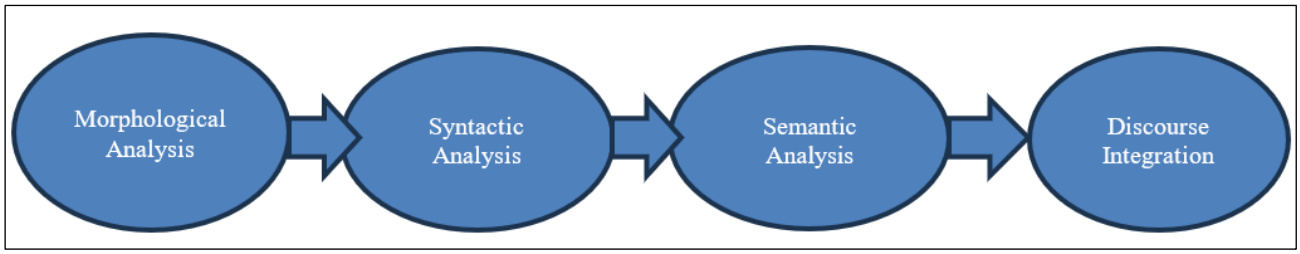

Figure 2: The Proposed Methodology [8]

MORPHOLOGICAL ANALYSIS

The lexicon of a language is its vocabulary that includes its words and expressions. Morphology depicts analyzing, identifying and description of structure of words. Individual words are analyzed into their components and non-word tokens such as punctuation are separated from the words.

SYNTACTIC ANALYSIS

Linear sequences of words are transformed into structures that show how the words relate to each other. The objective of syntactic analysis is to find the syntactic structure of the sentence. This step divides the sentence into simpler elements that are called tokens. Token Analyzing function is used to split the input string into a sequence of primitive units called tokens that is treated as a single logical unit. Then Spelling Checker function makes sure that each token is in the systems dictionary (lexicon) and if this is not the case then the spelling correction is performed or new words are added to the systems‟ vocabulary. Ambiguity reduction function reduces the ambiguity in a sentence and simplifies the task of the parser.

SEMANTIC ANALYSIS

In this structures created by the syntactic analyzer are assigned meanings. Semantics is associated with the meaning of language. Semantic analysis is concerned with creating representations for meaning of linguistic inputs. In this focus is on logical words, and no attention is paid to non-logical words. It deals with how the meaning of sentence is determined from the meanings of its parts. And thus, it generates a logical query which is fed as input to the database query generator.

DISCOURSE INTEGRATION

The meaning of an individual sentence may depend on the sentences that precede it and may influence the meanings of the sentences that follow it. The entities that structure those sentences may be related to entities that also were previous or may be introduced explicitly. The overall discourse must be coherent.

IV. RESULTS

For the implementation of algorithm we have used C#.net as Frontend and SQL server as backend. Table 1 shows the short descriptive question and answers in Punjabi with their respective score.

Table 1: Short descriptive question and answers in Punjabi

Question	Answer No.	Answer	Score
ਮਹਾਤਮਾ ਗਾਂਧੀ ਦਾ ਪੂਰਾ ਨਾਮ ਕੀ ਸੀ	1.	ਮਹਾਤਮਾ ਗਾਂਧੀ ਦਾ ਪੂਰਾ ਨਾਮ ਮੋਹਨ ਦਾਸ ਕਰਮ ਚੰਦ ਗਾਂਧੀ ਸੀ	5
	2.	ਪੂਰਾ ਨਾਮ ਮੋਹਨ ਦਾਸ ਕਰਮ ਚੰਦ ਗਾਂਧੀ ਸੀ	4
	3.	ਮੋਹਨ ਦਾਸ ਕਰਮ ਚੰਦ ਗਾਂਧੀ	3
	4.	ਕਰਮ ਚੰਦ ਗਾਂਧੀ	2
	5.	ਗਾਂਧੀਮ	0

In table 1, Answer No. 1 shows that as the student filled the complete and correct answer, he got the total score as 5.Answer No. 2 shows that as the student did not fill the complete answer(one word is missing in the answer), he got the total score as 4.Answer No. 3 shows that as the student did not fill the complete answer(two words are missing in the answer), he got the total score as 3.Answer No. 4 shows that as the student did not fill the complete answer(three words are missing in the answer), he got the total score as 2.Answer No. 5 shows that as the student given the incorrect answer, he got the total score as 0.

V. CONCLUSION & FUTURE SCOPE

This paper shows efficient approach that worked on checking of short answers as shown in Table 1. It makes a relationship between a word and the sentence through their meanings. The proposed method can detect the inequality between the students' answers. So this method can automatically evaluate the student scores in the subjective question answers. The proposed method works only in Punjabi Language. In future it may be extended for multiple languages.

VI. REFERENCE

1. Arora, P., & Goswami, P. (2014). An Efficient Hindi Language Interface using Relational Databases. *International Journal of Research in Computer and Communication Technology*, *3*(6).
2. Tarandeep, Singh, Walia., Gurpreet, Singh, Josan. & Amarpal, Singh, (2017). "Semantic Features for Automated Answer Scoring", *International Journal of Advanced Research in Science and Engineering*, Volume No. 6, Issue No. 10.
3. BGIET, S., & Jethi, A. (2015). Improved greedy routing protocol for VANET. *International Journal*, *2*(7), 15-18.
4. Khillare, S. A., Shelke, B. A., & Mahender, C. N. (2014). Comparative study on question answering systems and techniques. *International Journal of Advanced Research in Computer Science and Software Engineering*, *4*(11), 775-778.
5. Pribadi, F. S., Adji, T. B., Permanasari, A. E., Mulwinda, A., & Utomo, A. B. (2017, March). Automatic short answer scoring using words overlapping methods. In *AIP Conference Proceedings* (Vol. 1818, No. 1). AIP Publishing.
6. Soni, M., & Thakur, J. S. (2018). A systematic review of automated grammar checking in English language. *arXiv preprint arXiv:1804.00540*.
7. Ramalingam, V. V., Pandian, A., Chetry, P., & Nigam, H. (2018, April). Automated essay grading using machine learning algorithm. In *Journal of Physics: Conference Series* (Vol. 1000, p. 012030). IOP Publishing.
8. Gönen, K., & EnesTuncdemir, P. D. Automated Essay Scoring versus Human Scoring: A Reliability Check.
9. Song, S., & Zhao, J. (2013). Automated essay scoring using machine learning. *Stanford University*.
10. Kaur, H., Jethi, A., & BGIET, S. (2016). Efficient Routing Protocol in Mobile Ad-Hoc Networks.
11. Valenti, S., Neri, F., & Cucchiarelli, A. (2003). An overview of current research on automated essay grading. *Journal of Information Technology Education: Research*, *2*(1), 319-330.