

Online Evaluation System for Descriptive Type Answers

Prof. N. A. Deshmukh¹ Ms. Shreya Pawar² Ms. Snehal Augad³ Ms. Manjiri Deshmukh⁴
Ms. Charmi Kamdar⁵

^{1,2,3,4,5}Department of Computer Science & Engineering

^{1,2,3,4,5}Prof. Ram Meghe Institute of Technology & Research, Badnera, India

Abstract— Online evaluation of Descriptive test is rarely used. Now-a-days in online examination, objective test exams are already available but the descriptive exams are in need in terms of understanding the knowledge of examiner and to reduce manual efforts of evaluations by the examinees. This paper presents effective techniques for the evaluation of descriptive type answers. In this, the answers given by the examiners will be evaluated. The answers given by the examiners and the answers provided by the examinees will be evaluated using semantic similarity. Different techniques such as keyword extraction, pre-processing of text, and keyword density have been used. This system results in good accuracy, improved results of reliability, and reduces the time and efforts taken by the examinees.

Key words: Descriptive Test, Semantic Similarity, Pre-Processing, Keyword Density

I. INTRODUCTION

In today's Education System the entire entrance exam in all different fields are objective tests. Objective tests are all not that sufficient in order to test the knowledge of students. Students are only judged by the answers that they have marked. In this case, there can be two situations either the answer that is marked is surely known by the student or it can be an assumed answer. So in such situation we cannot completely judge whether the student is really intelligent or whether it is his/her luck. The student may have some knowledge about the topic but not complete, in such cases in order to actually test the knowledge of student, descriptive answers play an important role. But the evaluations of descriptive answers are mostly manual and that becomes too hectic for faculty. To solve this problem of manual checking of subjective answers we have proposed an online evaluation system of descriptive type answers.

Descriptive answers vary from student to student, so in the proposed paper to extract the meaning from the various answers the concept of semantic similarity is used. Faculty needs to give the answer. The answer will be stemmed and stop words will be removed [19]. The student's answers will be matched with the faculty's answers using brute force algorithm. For this the required text will be extracted from the database. The marks will be assigned using the concept of keyword density. The result analysis will then be messaged to the students.

A. Objectives

- To develop an online exam system for descriptive type questions.
- To implement techniques and algorithms that uses semantic similarity for evaluation of detailed type answers.

B. Motivation

The existing system of manual evaluation requires abundance of time and physical efforts. The current online checking of the answer sheets involves the scanning process which is quite expensive. So considering these points we will propose an online evaluation system of descriptive answers which will save time, price and human efforts.

II. LITERATURE REVIEW

There are various works proposed for short answers evaluation and objective answers. Even though, there are various work proposed for short-answer evaluation, the works related to the descriptive type answer evaluations are very limited. Some work related to descriptive answers evaluation is mentioned below.

Menaka Sand Radha N, have classified the text using keyword extraction. The keywords are extracted using TF-IDF and WordNet[1]. TF-IDF algorithm is used to select the words and WordNet is the lexical database of English used to find the similarity among the words. In this proposed work, the word which have the highest similarity are selected as keywords. Sungjick Lee and Han-joon Kim proposed conventional TFIDF model for keyword extraction. It involves cross domain filtering and table term frequency(TTF) for extraction[2]. Ari Aulia Hakim, Alva Erwin, Kho I Eng, Maulahikmah Galinium, and Wahyu Muliady works on the TF-IDF algorithm which create a classifier that can classify the online articles[3]. Stephen Robertson, explains the understanding concepts of IDF[4]. Professor Teuvo Kohonen, along with a group of researchers at the Neural Networks Research Center in Helsinki University of Technology, developed a few optimization techniques for SOM training.[5]. KristaLagus received the M.Sc. degree in Computer Science from Helsinki University of Technology, Espoo, Finland, in 1996. She has been a Research Associate at the Neural Networks Research Centre, Helsinki University of Technology, since 1995. Her main research interests are related to neural networks, especially self-organizing maps, and their application to natural language processing and data mining. Jarkko Salojärvi received the M.Sc. degree in technical physics from Helsinki University of Technology, Espoo, Finland, in 1998. His main research interests are related to neural networks, the emphasis being on self-organizing maps and their application to data mining[6]. Stop words Vangie BealWords that are filtered out by Web search engines and other enterprise searching and indexing platforms. Stop words are natural language words which have very little meaning, such as "and", "the", "a", "an", and similar words.[8]. In the University of Adelaide, study has been performed for compare LSA vector with word and n-gram feature vectors. When compared with word and n-gram, LSA gives better performance because according to an

entropy measure LSA vectors are weighted[7]. P.Y.Hui, and H.Y.Meng, used LSA for semantic explanation of a multimodal language with speech and gestures [9].

III. PROPOSED SYSTEM

The proposed system is an educational based system. In this system the online exam will be descriptive unlike other objective online exams. The system will be administered by the exam system admin. The teaching staff will conduct exams and students will be involved in solving the tests.

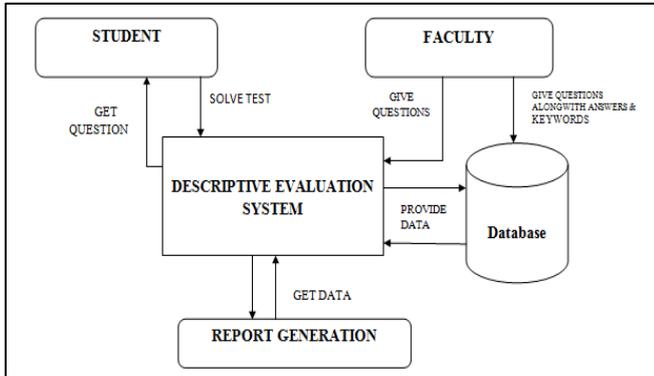


Fig. 1: DFD for online evaluation system of descriptive answers

The DFD diagram explains that the faculty will provide questions along with the answers that will be stored in the database. The questions will be sent to the online evaluation system. The system will provide the questions to the student which the staff has set. Student's answers will be given to the evaluation system where it will be compared with the data in the database and the system will check the answers and it will be passed to the report generation where the marks will be calculated and the result will be sent back to the system which will then will messaged to the students.

IV. TECHNIQUES USED

A. Stemming:

In information retrieval, the process of reducing modulated words to their word stem, base or root form—generally a written word form is known as stemming. The stem need not be identical to the morphological root of the word; it is usually sufficient that related words map to the same stem, even if this stem is not in itself a valid root[17].

B. Stop word Removal:

Stop words are those words which are extremely common words which would appear to be of least importance in helping select documents matching a users need are excluded from the vocabulary entirely. These words are known as stop words.

C. Brute-Force Algorithm:

The brute force algorithm is used to check the first occurrence of the pattern at all positions in the text between 0 and n-m. Then, after each attempt, it shifts the pattern by exactly one position to the right.

D. Keyword density:

Keyword Density is the percentage of occurrence of keywords from two texts used for comparison (one text is

answers provided by the teacher and the other is the student's answers).

V. CONCLUSION

The existing online exam systems are mostly objective exams because online evaluation for multiple choice questions is a very simple task. The proposed system aims on evaluating descriptive answers. From the proposed system it is clear that descriptive answers too can be evaluated automatically. This will reduce the work of manual evaluation of number of answer sheets. Various algorithms and techniques like stop words removal, stemming, Brute-force algorithm, keyword density, has effectively contributed in evaluating the student's answers though each student's answer is different. These algorithms will assign appropriate marks to the answer. The proposed system will surely help the educational system in getting the accuracy for marks allocation. The result will be mailed to the students in very short time.

REFERENCES

- [1] Menaka S and Radha N, —Text Classification using Keyword Extraction Technique, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 12, December 2013.
- [2] Sungjick Lee and Han-joon Kim, —News Keyword Extraction for Topic Tracking, Fourth International Conference on Networked Computing and Advanced Information Management, 2008.
- [3] Ari Aulia Hakim, Alva Erwin, Kho I Eng, Maulahikmah Galinium, Wahyu Muliady, —Automated Document Classification for News Article in Bahasa Indonesia based on Term Frequency Inverse Document Frequency (TF-IDF) Approach, 6th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 2014.
- [4] Stephen Robertson, —Understanding inverse document frequency: on theoretical arguments for IDF, Journal of Documentation, Vol. 60, No.5, pp 503-520, 2004.
- [5] Teuvo Kohonen, Fellow, IEEE, Samuel Kaski, Member, "Self Organization of a Massive Document Collection" IEEE, Krista Lagus, Jarkko Salojärvi, Jukka Honkela, Vesa Paatero, and Antti Saarela IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 11, NO. 3, MAY 2000
- [6] H.Ahonen, O.Heinonen, M.Klemettinen and A.I.Veramo, "Applying data mining techniques for descriptive phrase extraction in digital document collections"; Proc. IEEE International Forum, Digital Libraries, 1998, pp.2-11.
- [7] <https://www.techopedia.com/definition/13698/tokenization>
- [8] P.Y.Hui and H.Y.Meng, "Latent Semantic Analysis for Multimodal User Input With Speech and Gestures. Audio, Speech, And Language Processing IEEE/ACM Transactions", Volume 22, Issue 2,2014, pp.417-429.
- [9] L.D.Lee, H.Chuang and K.Seamons, "Document ranking and the vector-space model. Software", IEEE, Volume 14, Issue 2, 1997, pp.67-75.

- [10] Y.C.Liu, C.Wu and M.Liu, "Research of fast SOM clustering for text information", *Expert Systems with Applications*, Volume 38, Issue 8, 2011, pp.9325–9333.
- [11] D.S.McNamara, "Computational methods to extract meaning from text and advance theories of human cognition", *Topics in Cognitive Science*, Volume 3, Issue 1, 2011, pp.3-17.
- [12] K.Meena and R.Lawrance, "Evaluation of the Descriptive type answers using Hyperspace Analog to Language and Self-organizing Map", *Proc. IEEE International Conference on Computational Intelligence and Computing Research*, 2014, pp.558-562.
- [13] J.Vesanto and E.Alhoniemi, "Clustering of the Self-Organizing map", *Neural Networks, IEEE Transactions*, Volume 11, Issue 3, 2000, pp.586-600.
- [14] P.Willett, "The Porter stemming algorithm: then and now", *Program*, Volume 40, Issue 3, 2006, pp.219-223
- [15] K.P.N.V.Satya sree , Dr.J V R Murthy," CLUSTERING BASED ON sCOSINE SIMILARITY MEASURE" [IJESAT] INTERNATIONAL JOURNAL OF ENGINEERING SCIENCE & ADVANCED TECHNOLOGY Volume-2, Issue-3, 508 – 512
- [16] Ankita Patil and Prof. Achamma Thomas," A Survey of Effective Techniques for Subjective Test Assessment" *International Journal on Recent and Innovation Trends in Computing and Communication* ISSN: 2321-8169 Volume: 4 Issue: 3 400 – 403
- [17] Meena,K and Lawrance,R ," Semantic Similarity Based Assessment of Descriptive Type Answerstment "2016 IEEE International Conference
- [18] Sonal N. Deshmukh and Ratnadeep R. Deshmukh," Cosine Similarity for Substituted Text Detection " Volume 4, Issue 1, January 2014 ISSN: 2277 128X *International Journal of Advanced Research in Computer Science and Software Engineering*