



Statistical Technique for Automated Answer Scoring: An Overview

Tarandeep Singh Walia¹, Gurpreet Singh Josan², Amarpal Singh³

¹Research Scholar, IKG PTU, Kapurthala, Punjab, India,

²Punjabi University, Patiala, Punjab, India.

³Beant College of Engg. &Tech. , Gurdaspur, Punjab, India.

ABSTRACT

The Automated Scoring System has its importance in providing the student with a score as well as feedback within seconds. This paper describes an Automated Scoring system in which scores are assigned to answers automatically based upon predefined algorithms. It discusses various techniques for Automated answer scoring and explains how the rule-based technique followed by statistical techniques used to obtain the final answer score. For optimizing the scoring, some rules are set and used together with the statistical approach. This study reviews the various methods used in the previous Automated Answer Scoring system, and attempts to develop a new semantic features for Automated Scoring System.

Keywords: Automated Scoring (AS), Automated Essay Scoring (AES), Natural Language Processing (NLP).

1. INTRODUCTION

The automated scoring is defined as the act of assigning scores to answers automatically based upon predefined algorithms. The automated scoring receives an answer in the typed text form as an input and presents a score as output based upon various features of the text. While generating the score, the input text passes through various modules like pre-processing, extract features and classify (Fig. 1).

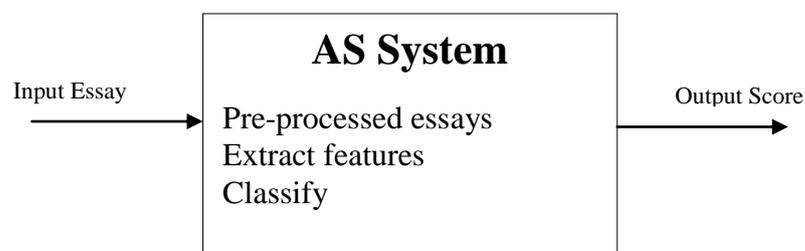


Fig. 1 Automated Scoring (AS) System



Current 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 (Mark et al, 2014).

The system to be developed is planned to follow two step process. The first step involves analysing the input answer so as to identify possible errors viz spelling errors and syntactic errors. The second step compares the input answer with given answer essay to detect semantics and differences as errors. There after the output generated by the system is compared with the result given by human rater and in this way the performance of the system is evaluated. To estimate the accurate score generated from syntactic similarity to semantic similarity.

2. AES and NLP

NLP have major tasks such as discourse analysis, morphological segmentation, parsing, word sense disambiguation and information extraction etc. Automated Scoring can choose some tasks from NLP for scoring process. Automated Scoring systems 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.

3. Related Work

This section deals with an statistical approach that includes VSM representing its importance in Automated answer scoring:

Vector Space Model(VSM) Approach

Li Bin, Lu Jun, Yao Jian(2008), used the K-Nearest Neighbour (KNN) algorithm for AES. Each vector is defined by the term frequency and inversed document frequency (TF-IDF) weight. The information gain (IG) and TF methods are used to select features based on predefined features. They calculated the similarity of essays with cosine in the KNN algorithm. Regarding the various methods of feature selection, they are able to achieve 76% accuracy.

Yali Li, Yonghong Yan (2012) demonstrates AES system using **vector regression**. Each essay is represented by the Vector Space Model (VSM). To implement the system, it gives score on several features, including the surface features and complex features such as grammar checking, sentences. Both the words and part-of-speech tag are taken into consideration. They get 86% accuracy given the two scores deviation compared to human raters.

Alzahrani, Alzahrani, Arfaj, Almohammadi (2015), developed and compared number of NLP techniques that accomplish the task of automating scoring. Their study is based on a vector space model (VSM) in which after normalization, the baseline-system represents each answer by a vector, and subsequently calculates its



score using the cosine similarity between it and the vector of the model answer. They represented the multi vector model which is closer to human judgement and gives more accurate and reliable indication. They also plan to apply their methodology in different languages.

Badr Hssina, Belaid Bouikhalene, Abdelkrim Merbuha (2016) in their study, have create a semantic similarity calculation system between text documents to add to their semantic clustering.

4. Vector Space Model

Vector Space Model (VSM) is a statistical technique in NLP. In this model, words are called terms and represented as vectors. There is a formula that calculates similarities of different pairs of words in the vector space (Kaja et al.,2017).

$$\text{sim}(d_j, q) = \frac{d_j \cdot q}{\|d_j\| \|q\|}$$

The cosine similarities are measured by comparing the query q (student) vector with the document d_j (reference) vector. Consider Fig. 2 given below:

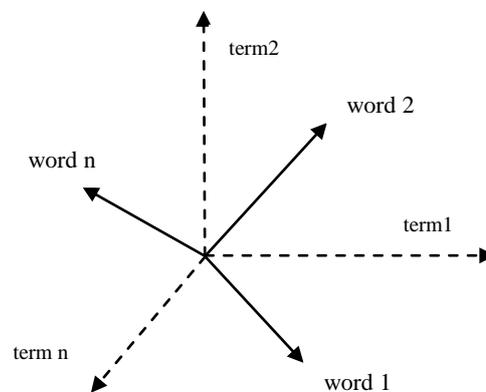


Fig. 2 Vector Space Model

Ensuring the Quality of Automated Scoring Implementation

The following are the measurement of quality of Automated Scoring that must be implemented through VSM:

- Automated scores are reliable and fair.
- The influence of automated scoring on reported scores is understandable
- Automated scores are produced is meaningful and understandable.
- Automated scores are consistent with the scores from expert human graders.
- Automated scores have valid scores, when checked against external measures in the same way done with human grading.



5. Proposed Technique

Following section discuss in detail the rule-based technique followed by statistical techniques used to obtain the final answer score Fig. 3.

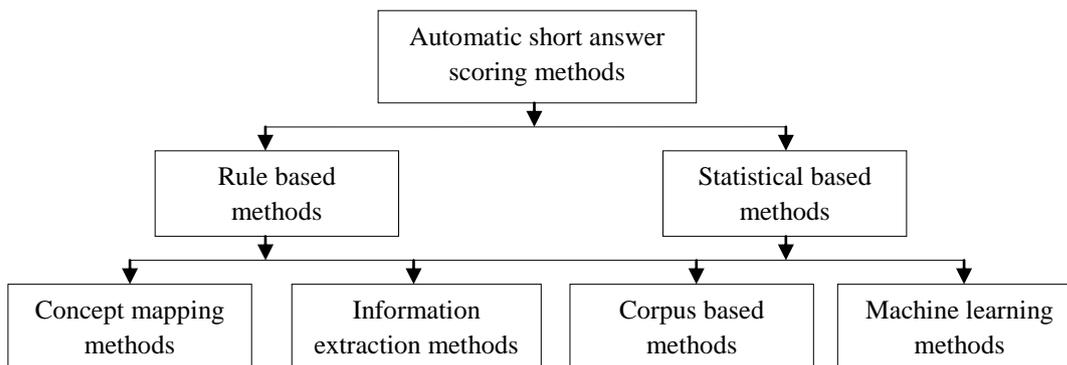


Fig 3: Methods of Automatic short answer

5.1 Rule-based technique

Rule-based systems are used as a way to store and manipulate knowledge to interpret information in a useful way. The rule-based technique helps in scoring more accurately. For optimizing the scoring, some rules are set and used together with the statistical approach.

5.2 Statistical based technique

The typical flowchart to statistical technique are shown in Fig. 4. It has three parts viz.: In the pre-processing part, input the student answer and reference answers and then information is extracted based on features. In the second part, the vector space model is built by assigning weight to the student vector and reference vectors. In the post-processing part, similarities are found and scoring is done.

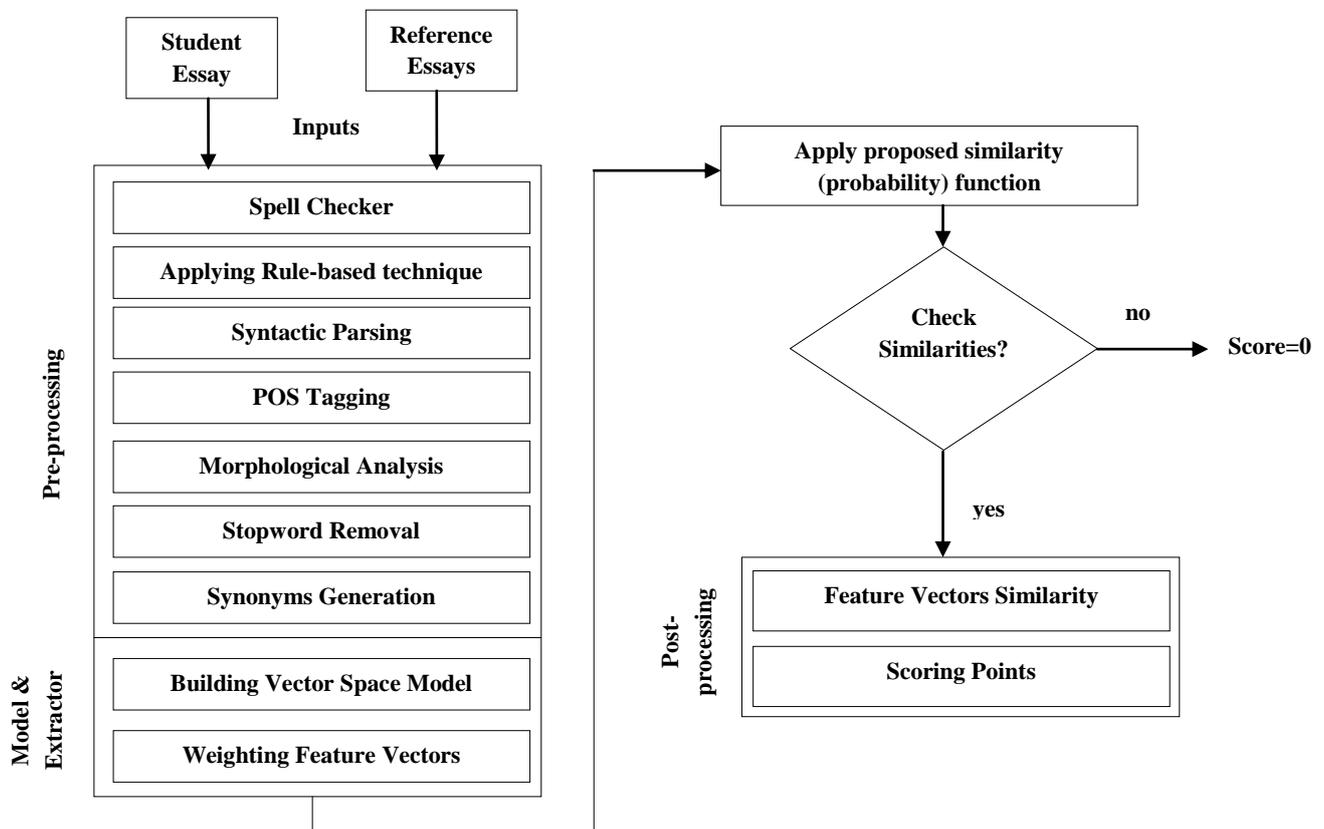


Fig. 4: A typical flowchart to Statistical technique

CONCLUSION

Automated system can be very useful because they can provide the student with a score as well as feedback within seconds. The Automated Scoring system is more objective and consistent than human raters. This system has vastly improved, incorporating many more syntactic and semantic features by integration and collaboration with statistical method (VSM) to fulfill the objective of the study. Automated Scoring will perform fair scoring, can be repeated again and again with consistency. The semantic features are essential while scoring because it holds whole content (meaning) of student answer. Therefore, a system for automated grading would be consistent in the way it grades answers, and enormous cost and time savings could be achieved. So there will be need of development of the new system for Indian languages.



REFERENCES

1. Ahmed Alzahrani, Abdulkareem Alzahrani, Fawaz K. Al Arfaj, Khalid Almohammadi, Malek Alrashid (2015), “ AutoScor: An Automated System for Essay Questions Scoring”, published by International Journal of Humanities Social Sciences and Education (IJHSSE), Volume 2, Issue 5, pp 182-187.
2. Ali Muftah Ben, Mohd Juzaidin(2013), “Automated Essay Grading System for Short Answers in English Language”, published by Journal of Computer Science, Volume:9 Number:10 ISSN:1549-3636.
3. Badr HSSINA, Belaid Bouikhalene, Abdelkrim Mebouha(2016),“Evaluation of semantic similarity using vector space model based on textual corpus”, published by 13thInternational Conference Computer Graphics, Imaging and Visualization, IEEE, pp 295-300.
4. Chen, Yen-Yu, Chien-Liang Liu, Tao-Hsing Chang, and Chia-Hoang Lee(2010) "An Unsupervised Automated Essay Scoring System", published by IEEE Intelligent Systems, vol 25, Issue 5. Pp 61-67.
5. Jovita, Linda, Andrei Hartawan, Derwin Suhartono(2015), “Using Vector Space Model in Question Answering System” published by ICCSCI 2015, Procedia Computer Science 59, Elsevier, pp 305 – 311.
6. Kaja Zupan, Zoran Bosni(2017), “Automated essay evaluation with semantic analysis”, published by Knowledge-Based Systems 120, Elsevier, pp 118–132.
7. Kong Joo Lee, Yong-Seok Choi, Jee Eun Kim(2011), “Building an automated English sentence evaluation system for students learning English as a second language”, published by Elsevier, Computer Speech and Language 25, pp 246–260.
8. Leila Kosseim , Jamileh Yousefi(2008), “Improving the performance of question answering with semantically equivalent answer patterns”, published by Data & Knowledge Engineering, Elsevier, 66, pp 53–67.
9. Li Bin, Lu Jun, Yao Jian, Zhu Qiao-Ming (2008) “Automated essay scoring using the KNN Algorithm”, published by International Conference on Computer Science and Software Engineering, IEEE, Volume: 1, pp 735-738.
10. Lihao Ge, Teng-Sheng Moh(2017), “Improving Text Classification with Word Embedding”, published by IEEE International Conference on Big Data (BIGDATA). Pp 1796-1805.
11. Mark D. Shermis & Jill Burstein(2014), “ Handbook of Automated Essay Evaluation: Current Applications and New Directions”, published by Language Learning & Technology <http://llt.msu.edu>, Volume 18, Number 2.



12. Matthias H. Heie, Edward W.D. Whittaker, Sadaoki Furui(2012), "Question answering using statistical language modelling", published by Elsevier , Computer Speech and Language 26, pp193–209.
13. Mohd Juzaidin Ab Aziz, Fatimah Dato' Ahmad, Abdul Azim Abdul Ghani, Ramlan Mahmud (2009), "Automated Marking System for Short Answer Examination (AMS-SAE)", Published by IEEE Symposium on Industrial Electronics and Applications (ISIEA 2009), Volume:1, pp 47-51.
14. Show-Jane Yen, Yu-Chieh Wu, Jie-Chi Yang, Yue-Shi Lee, Chung-Jung Lee, Jui-Jung Liu(2013), "A support vector machine-based context-ranking model for question answering" , published by Elsevier, Information Sciences, vol 224, pp 77–87.
15. Steven Burrows, Iryna Gurevych, Benno Stein (2015), "The Eras and Trends of Automatic Short Answer Grading", published by Int J Artif Intell Educ, Springer, Vol 25, pp 60–117.
16. Yali Li, Yonghong Yan(2012), "An effective Automated essay scoring system using support Vector regression", published in proceeding of International Conference of Intelligent Computation Technology and Automation(ICICTA), IEEE, pp 65-68.