

Automatic Scoring System for Evaluating Comprehensive Answers using key Features (ASCA)



N.Sivaranjani, Amudha S, Mary Linda I

Abstract: Automatic evaluation of learner's skills is the hot research area in the field of Machine learning. The key idea behind automatic scoring system for evaluating students answer is to reduce the time for instructors in evaluating the test paper and to make the scores consistent. This paper not only aims to evaluate the answer and also takes care of malpractice by implementing plagiarism and also redundant sentences. Redundant sentences may lead to erroneous grades. Experimental results show that our system shows better grading than the existing and also has higher match with the human judgment, especially for the answer that has higher similarity with answer key.

Keywords : Automatic scoring, similarity measure, sentence ordering

I. INTRODUCTION

Brief answers are very common in various learner's evaluation. The traditional way of evaluating the papers is time consuming and it is subjective. So automatic way of correcting the paper will largely reduce the time taken for correcting the paper and also human errors.

Evaluation is the most common and important part of any educational system. Currently all educational institutions have implemented automatic correction of multiple choice questions and True/False widely. Since MCQs and True/False type questions alone are not sufficient to test learner's knowledge on subject, as learners can easily guess or randomly choose the answer. Hence educational institution prefers brief answers also. Instructor can more effectively assess the learner's knowledge using brief-answer type questions. For brief answer assessment, it requires instructors for evaluation. Though many researchers have successful results for evaluating exam papers, automatic correction is not yet popular due to the challenges present in the evaluation and that are mostly invisible.

Natural language processing provides better way of carrying out this task without much deeper analysis. The most common approach for automatic evaluation is Text-to-Text similarity. The similarity value will be high, if the both answer key and the learners answer exactly matches. What if learner has just copied the content? In this case, other automatic scoring system gives the highest score but it is actually a malpractice. Our proposed system considers this as special case and awards zero marks[20],[22],[24].

Our system initially scores the learners answer considering three features like sentence ordering, missing points and extra points. Instructor prepares the answer for each question given for assessment which is then considered as answer key. Automatic scoring system calculates the similarity between the instructor's answer key and the learners answer and scores the learners answer based on the similarity value[25],[27],[29].

A model is proposed in this paper to monitor the garbage level of the respective garbage bins in real time and to detect the level when the threshold value is reached by combining Sensors and Raspberry pi. Using the WiFi module, this information will be sent to the control unit and updated on time, depending on which optimized path for Garbage Collecting Van (GCV) has to be found, depriving fuel consumption, cost, time and labor. The information will be given whether or not the waste is fully separated by a moist sensor and a humidity sensor that helps to recycle, eliminate and reuse waste. Qualitative analysis will be performed to produce reports using information mining. The primary goal of this scheme to be introduced is to replace the current tedious system that will help the town become a smart city[26],[28],[30].

Automating the paper evaluation is not as easy as, because no two learners answer the question in same way. The way one learner uses the word will be different from the other learners and also the from the answer key. For e.g. the answer key may contain the word Transmission control protocol whereas learner might have written as TCP. Though these two words are same, the similarity score will be low. Below are the some challenges related to this area[1].

1. Synonym Term: The words present in the answer key and in the learners answer are different but has the same meaning. This can be identified using Word Net
2. Numeric Term: The numeric value and the equivalent number in words are mapped automatically. (Eg. "7th" is aligned to "seventh").

Manuscript published on 30 August 2019.

* Correspondence Author (s)

N.Sivaranjani, Department of Computer science and Engineering, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India. Email: ranjibalas@gmail.com

Amudha S, Department of Computer science and Engineering, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India.. Email: amudha17s@gmail.com

Mary Linda I, Department of Computer science and Engineering, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India. Email: catchlin.18@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

3. Acronym Term: If the word is given in abbreviated form, it tries to correlate.

4. Stemmed Term: If the word derived from the root word. For e.g the term go, goes, going are all same and should considered as similar term.

II. LITERATURE SURVEY

P.Selvi, et.al [1]proposed ASAGS, an automatic short answer grading system. This system incorporate heuristic based rules to overcome the challenges associated with scoring the learners answer. The three modules (i) Mapping module: maps the unigram of learners answer with the unigram of Answer key based on the techniques such as exact, stemmed and heuristics rule. (ii) Feedback module: used for training the data in the later stage. This module gives the student's performance for the given question.(iii) Validation module: compares the system generated score with the human score and computes the correlation[19],[21],[23].

An unsupervised AES system proposed by Yen-Yu Chen et.al [7] mainly aims to reduce the training data that are within the domain and does not contain scoring information. This system implements a novel unsupervised learning algorithm voting algorithm that is based on initial score and similarities between the essays which is applied for scoring essays without training data. The essays are scored based on six-point scoring scale. AES also normalizes the score giving to each essay using Z-score value based on availability of historical information.

YigalAttali, et.al [8]proposed a well-known system to score the student essay by identifying the writing proficiency features. This system also identifies and gives feedback to the students about the area to be improved. It provides the essay with the holistic score and real-time diagnostic score.

Bara I A., Fung C J., and Dinh T., et al.[5] displayed a novel Twitter spam finding strategy, by dissecting the connection between records dependent on their tweeting design closeness. A spam score calculation is proposed to iteratively refresh the spam scores of clients and tweets dependent on their example likeness and their closeness to realized at first named spam tweets. The examination dependent on genuine information exhibit that a considerable measure of new spam tweets and spam records are found by their proposed strategy which are generally not distinguished.

Platt J., et al.[6] abridge that SMO is an improved preparing calculation for SVMs. Like other SVM preparing calculations, SMO separates a huge QP issue into a progression of littler QP issues. In contrast to different calculations, SMO uses the smallest conceivable QP issues, which are settled rapidly and scientifically, by and large improving its scaling and calculation time essentially. SMO was tried on both true issues and counterfeit issues. From these tests, the accompanying can be deduced:SMO can be utilized when a client does not have simple access to a quadratic programming bundle or potentially does not wish to adjust that QP package.SMO does very well on SVMs where a considerable lot of the Lagrange multipliers are at bound. SMO performs well for straight SVMs in light of the fact that SMO's calculation time is overwhelmed by SVM assessment, and the assessment of a direct SVM can be communicated as a solitary spot item, as opposed to an

aggregate of straight bits[14],[16],[18].

Zheng X., Zhang X., Yu Y., Kechad T., and Rong, C. et al.[7] presents an ELM-based spammer recognition technique for interpersonal organization stages. Utilizing information crept from SinaWeibo, a lot of substance and conduct highlights are extricated and connected into an ELM-based classification calculation. Through a lot of analyses and assessment work, their proposed arrangement is demonstrated to be possible, efficient, and significantly more steady than existing SVM-based models. In any case, any measure of marked information probably won't be sufficient in an interpersonal organization condition with an enormous amount of exceptionally assorted qualities. In this manner, further take a shot at the subject may incorporate the examination of a communitarian preparing based semi-administered learning model that is fit to prepare itself naturally dependent on a limited quantity of marked information. Then again, highlights separated in this proposed arrangement (and other existing methodologies) depend on measurable investigation and manual choice. In the time of huge information with tremendous information volumes and advantageous access, highlight extraction systems in our answer may be low in versatility and to some degree costive. Hence, taking into account how to import the idea of Machine Learning innovation (e.g., profound learning calculations [30–33]) into programmed highlight learning and extraction has turned into a significant inquiry[19],[21],[23].

III. PROPOSED SYSTEM

A.Pre-processing

One of the most common and important step of any NLP research is preprocessing, where the sentences are tokenized and stop words are removed. Learners answer and answer key both are preprocessed for further process.

B.Keyword Extraction

Once the documents are preprocessed, the important keywords are extracted from both the documents. These keywords are domain specific, since AGBA concentrates only on engineering related documents.

C.WordNet Mapping

Since there is a possibility that same words can be represented in different forms and word which has same meaning. The unigrams should be matched together for efficient scoring of answers. After the keyword extraction the root words are mapped using WordNet dictionary. As said before in the challenges, the unigrams, whose structure are different but has the same meaning can be mapped using WordNet dictionary.

D.Similarity Measure

The similarity between the learners answer and the answer key should be calculated. This should also include sentence ordering to calculate efficient grade.

The grade of learners answer is calculated by checking three things (i) sentence ordering (ii) number of points missing and (iii) extra points those belong to the same context.

E. Sentence ordering

The sentences in the learner’s answers are given with the position value. First the key words in the instructor answer key are extracted sentence by sentence. Similarly the keywords in the learners answer are also extracted. Now the sentence in the learners answer give the position value based on the position it occupies in the answer key. For example, the first sentence in the learner answer is the fifth sentence in the Instructor answer key, then the first sentence is given the value 5. Similarly all the sentences are given with the values based on its position in the answer key. The similarity between the answer key and the learners answer are calculated using semantic similarity. Sentence ordering score is given to each answer based on how much they differ from the answer key[20],[22],[24].

F. Missing points

The keywords those are present in answer key but missing in learners answer are considered as missing points and marks has to be reduced for those missing points. For eg., the keywords present in the first line of answer key does not match with any of the keywords in the learners answer, then it is treated as “one point missing”. Marks will be reduced accordingly. One mark will be reduced for 2 missing points.

G. Extra point

The keywords those are present in the learners answer but not present in answer key are treated as extra points. For those keywords which has no match with the answer key but are in context of the question given should be considered for adding marks. This feature is important when learners are asked to explain the question with an example because it is not necessary that learner should write the same example given in the answer key in the exam. Context based similarity measure [12][13] is used to identify whether the extra points written by the learner falls under the context of question asked.

All the above mentioned features are considered to score the learners answer. For eg., a learners answer which is to be evaluated for 10 marks is compared with instructor answer key and found that 3 sentences are jumbled, 2 points are missing and has one extra point. For this example AGBA reduces 2 marks for sentence ordering, 0.5 marks for missing points and adds one mark for extra point. Hence the score for this answer will be 8.5/10. Similarly a learner’s answer has no jumbling of sentences, no missing and extra points, then it is considered as special case because learner might have copied from the book or something similar. This special case is matched against text book, if the similarity value is high when compared to similarity value of answer key then it is considered as malpractice and ZERO marks will be awarded[25],[27],[29].

IV. EXPERIMENTAL RESULTS

The data set is created by involving 50 students. They were given a question paper of 100 questions related to computer science related subject. The question paper contains different types of questions like definition, questions asking for

example, list some points. The entire data set of 5000 answers is used as training data. Our system is also compared with human grading and also other automatic scoring system. Our system shows better performance than other automatic scoring system and also it is much better for the questions that ask for examples because our proposed system considers the extra points where other automatic scoring system ignores the same.

	Semantic based (%)	ASAGS (%)	ASCA (%)
Sentence ordering	14.3	10.2	2.1
Missing points	39.7	28.4	2.9
Extra Points	57.8	38.5	1.9

Table 1: Comparison of scores with existing system

Table 1 shows the percentage differences in marks given by the different automatic scoring system and instructor’s score. Our system ASCA shows the less difference with the human score and performs better than other automatic scoring system. For eg. For sentence scoring the marks given by ASCA differs from human scoring only by 2.1% whereas semantic based and ASAGS automatic systems differs by 14.3% and 10.2 percent respectively which is higher than ASCA.

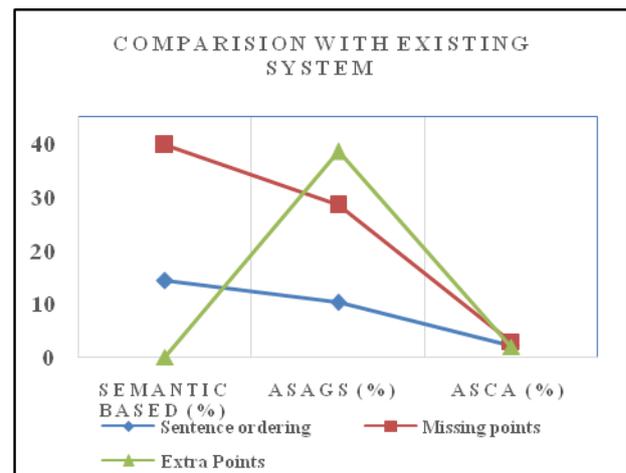


Fig 2: Comparison of proposed system with existing system

Fig.2 shows the performance analysis of different automatic scoring system with the proposed system. The graph clearly shows the percentage difference of each system with the human scoring[26],[28],[30]. The proposed system ASCA has the minimum variation with the human scoring when compared with the other automatic scoring system.

V. CONCLUSION

In this paper, automatic scoring of brief answers is implemented using three key features sentence ordering, missing points and extra points. The performance of our system is compared with the existing automatic scoring system and also with human scoring and ASCA outperforms the others by giving 72% closer value to human scoring.



REFERENCES

1. Gowri Sankaran, B., Karthik, B. & Vijayaragavan, S.P. 2019, "Weight ward change region plummeting change for square based image huffman coding", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 10, pp. 4313-4316.
2. Gowri Sankaran, B., Karthik, B. & Vijayaragavan, S.P. 2019, "Image compression utilizing wavelet transform", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 10, pp. 4305-4308.
3. Kandavel, N. & Kumaravel, A. 2019, "Offloading computation for efficient energy in mobile cloud computing", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 10, pp. 4317-4320.
4. Vinoth, V.V. & Kanniga, E. 2019, "Reversible data hiding in encrypting images-an system", International Journal of Engineering and Advanced Technology, vol. 8, no. 6, pp. 3051-3053.
5. Selvapriya, B. & Raghu, B. 2019, "Pseudocoloring of medical images: A research", International Journal of Engineering and Advanced Technology, vol. 8, no. 6, pp. 3712-3716.
6. Senthil Kumar, K. & Muthukumaravel, A. 2019, "Bi-objective constraint and hybrid optimizer for the test case prioritization", International Journal of Engineering and Advanced Technology, vol. 8, no. 6, pp. 3436-3448.
7. Kavitha, G., Priya, N., Anuradha, C. & Pothumani, S. 2019, "Read-write, peer-to-peer algorithms for the location-identity split", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 445-447.
8. Kaliyamurthie, K.P., Michael, G., Anuratha, C. & Sundaraj, B. 2019, "Certain improvements in alzheimer disease classification using novel fuzzy c means clustering for image segmentation", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 599-604.
9. Kaliyamurthie, K.P., Sundarraj, B., Geo, A.V.A. & Michael, G. 2019, "RIB: Analysis of I/O automata", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 1019-1022.
10. Velvizhi, R., Rajabhushanam, C. & Vidhya, S.R.S. 2019, "Opinion mining for travel route recommendation using Social Media Networks (Twitter)", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 508-512.
11. Kavitha, R., Sangeetha, S. & Varghese, A.G. 2019, "Human activity patterns in big data for healthcare applications", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 1101-1103.
12. Pothumani, S., Anandam, A.K., Sharma, N. & Franklin, S. 2019, "Extended VEOT framework - Implemented in a smart boutique", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 762-767.
13. Kaliyamurthie, K.P., Michael, G., Krishnan, R.M.V. & Sundarraj, B. 2019, "Pseudorandom techniques for the internet", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 915-918.
14. Aravindasamy, R., Jeffrin Rajan, M., Rama, A. & Kavitha, P. 2019, "Deep learning provisions in the matlab: Focus on CNN facility", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 990-994.
15. Theivasigamani, S., Linda, M. & Amudha, S. 2019, "Object sensing and its identification & motion sensing", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 545-549.
16. Mary Linda, I., Vimala, D. & Shanmuga Priya, K. 2019, "A methodology for the emulation of IPv4", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 848-852.
17. Velvizhi, R., Priya, D.J., Vimala, D. & Linda, I.M. 2019, "Increased routing algorithm for mobile adhoc networks", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 1606-1608.
18. Sangeetha, S., Anuradha, C. & Priya, N. 2019, "DNS in real world", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 937-940.
19. Geetha, C., Vimala, D. & Priya, K.S. 2019, "Constructing multi-processors and spreadsheets with SKIVE", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 516-519.
20. Yugendhar, K., Sugumar, V. & Kavitha, P. 2019, "A novel method of univac using fuzzy logic", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 435-437.
21. Kaliyamurthie, K.P., Michael, G., Elankavi, R. & Jijo, S.A. 2019, "Implementing aggregate-key for sharing data in cloud environment using cryptographic encryption", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 957-959.
22. Jeffrin Rajan, M., Aravindasamy, R., Kavitha, P. & Rama, A. 2019, "A novel method of object orientation variation in C++ and java", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 708-710.
23. Nayak, R., Dinesh, S. & Thirunavukkarasu, S. 2019, "A novel method improvement of rapid miner for the data mining applications", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 457-460.
24. Sivaraman, K., Krishnan, R.M.V., Sundarraj, B. & Sri Gowthem, S. 2019, "Network failure detection and diagnosis by analyzing syslog and SNS data: Applying big data analysis to network operations", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 883-887.
25. Vimala, D., Linda, I.M. & Priya, K.S. 2019, "Decoupling online algorithms from erasure coding in DNS", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 950-953.
26. Rama, A., Kumaravel, A. & Nalini, C. 2019, "Preprocessing medical images for classification using deep learning techniques", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 711-716.
27. Sangeetha, S., Srividhya, S.R., Anita Davamani, K. & Amudha, S. 2019, "A procedure for avoid overrun error in universal synchronous asynchronous receiver transmitter (usart) by utilizing dummy join and interrupt latency method", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 657-660.
28. Aravindasamy, R., Jeyapriya, D., Sundarajan, B. & Sangeetha, S. 2019, "Data duplication in cloud for optimal performance and security", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 1156-1158.
29. Aravindasamy, R., Jeffrin Rajan, M., Sugumar, V. & Kavitha, P. 2019, "A novel method on developing superblocks and the transistor using apodyral", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 982-985.
30. Sasikumar, C.S. & Kumaravel, A. 2019, "E-learning attributes selection through rough set theory and data mining", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 10, pp. 3920-3924.

AUTHORS PROFILE



N.Sivaranjani Assistant Professor, Department of Computer science and Engineering, Bharath Institute of Higher Education and Research, Chennai, India



Amudha S Assistant Professor, Department of Computer science and Engineering, Bharath Institute of Higher Education and Research, Chennai, India



Mary Linda I Assistant Professor, Department of Computer science and Engineering, Bharath Institute of Higher Education and Research, Chennai, India

