

Machine Learning based Automatic Answer Checker Imitating Human Way of Answer Checking

Vishwas Tanwar
Manipal University Jaipur

Abstract- In today's scenario, examinations can be classified into 2 types, one is objective and the other is subjective. Competitive exams are usually of mcq types and due to this they need to be conducted on computer screens as well as evaluated on them. Currently, almost every competitive exam is conducted in online mode due to the large number of students appearing in them. But apart from competitive exams, computers cannot be used to carry out subjective exams like boards exam. This brings in the need of Artificial Intelligence in our online exam systems. If artificial intelligence gets implemented in online exam conduction systems, then it will be a great help in checking subjective answers as well. Another advantage of this would be the speed and accuracy with which the results of the exams would be produced. Our proposed system would be designed in such a way that it will give marks in a similar way as of a human. This system will hence be of great use to educational institutions.

Index Term- Automatic Answer Checker, Answer Checker, Subjective Answer Checker, Answer Matching,

1. INTRODUCTION

In today's world, currently there are many exam conduction ways, be it online exams or OMR sheet exams or MCQ type exams. Various examinations are conducted every day around the world. The most important aspect of any examination is the checking of the answer sheet of the student. Usually it is done by the teacher manually, thus making it a very tedious job if the number of students is very large. In such a case automating the answer checking process would definitely prove to be of great use.

Automating the answer checking process would not only relieve the exam checker but the checking process would also get way more transparent and fair as there would not be any chances of biasedness from the teacher side. Nowadays various online tools are available for checking multiple choice questions but there are very few tools to check subjective answer type examinations.

This project aims to carry out the checking of subjective answer type examinations by implementing machine learning. This application can be used in various educational institutes for checking subjective answer type examinations. Further, on

improving the application, it can even be extended for conducting online subjective answer type examinations application, it can even be extended for conducting online subjective answer type examinations. On running the application, the main window of the application will give two options to the user, whether to login as an admin or as a student. After selecting one of the options, the user will get to see a login window where he will be asked to login using his/her credentials. The admin will have the options like uploading the question paper and seeing the responses of the students. The student will have the option to upload the answer sheet and see the marks allotted to them there and then.

2. BACKGROUND

Why did I choose Automatic Answer Checker ?

There is a lot of reasons why introduction of Artificial intelligence into the online exam systems would prove to be of great use. Firstly, as currently exams are marked by examiners and so this leads to fatigue and boredom as they have to check large number of answer sheets, but with the online system, this problem automatically gets solved. Moreover, the accuracy and speed with which a computer can generate results, it is something which a human would take hours to do. The proposed system would also produce unbiased results which would further make everything more transparent.

The proposed System would be having following features:

Balanced Load : Here the system would only be accessed by the admin and this would result in lower load on the server.

Ease of Use: The proposed system would be very easy and efficient to use. A student would be able to easily interact with the interface without any confusion.

Friendly Environment: The proposed system would be very friendly to use by any user who would work on it. No prior knowledge or set of instructions would be required by the user to operate the interface. Every part of the applications would be very easy to use. The application has been specially designed in a way that anyone can make the best use out of it without facing any major problem. The user would find the application extremely easy to use as well as unique in its own way.

Ease in Accessibility: The responses of the students would be easily accessible and their maintenance would also be easy. The teacher would be able to see the responses of each and every student who turned in his / her answer sheet , thus making it very easy for the teacher to keep the track of student's submission of his answer sheet . Moreover , the teacher would be easily able to view the answer sheet of the students who would have submitted their answer sheet quite easily . With the facility of viewing the responses of the students of their answer sheet , it would be quite easy and helpful for the teacher to complete the exam conduction process in a very smooth way .

Efficiency and Reliability of the System: The system would be having a high efficiency as it would be made with negligible errors , thus , making it even more reliable .

The marks that would be given to the student would be almost same as if the answer sheet has been checked by an examiner . It has been made sure that the accuracy in allotting the marks to students is quite high so that it hardly makes any difference that the answer sheet has been checked by a machine or it has been checked by a human being . The marks calculation has been done in a fair way so that there are no chances of complaints from the students regarding the marks given to them .

Each and every student would definitely be satisfied with their marks as the marks given to them would be as fair as possible without the chances of any error . Moreover , after the application would have given it's marks to the student , the marks would then be reviewed by the teacher as well and so if any student complains about his / her marks , the teacher would be able to review the answer sheet submitted by the student and increase his / her marks if the doubt of the student was genuine .

This further helps in increasing the accuracy of the application as well as making it more and more reliable .

Maintainence of the system: The automatic answer checker system has been designed in such a way that it's maintainence is very easy.

This system would definitely prove to be very much useful and productive for any educational institution as it would give examiners some relaxation from their tedious job and most importantly , the marks allotted would be almost the same . As the students would be getting to know their marks there and then , the whole exam checking process would get so much easy . The teacher would not have to carry answer sheets of students and then check that whole pile of sheets , rather , the automatic answer checker would give marks to the students making it such a time efficient process . A huge amount of time would get saved by introducing a machine to check the answer sheets of the students . Moreover , examiner would easily be able to see which student has submitted the answer sheet and who has not . Thus , this system would prove to be very much useful and productive .

3. LITERATURE SURVEY

A. SCENARIO

In today's world , competition among people has increased substantially . With growing population of the world , competition among people can be seen everywhere as everyone wants to live life of their dream . Everyone wants to be better than everyone else . Another big reason for this increased competition is limited resources , particularly jobs if we limit our focus of study to professional world . This competition begins in one's life from schools and colleges . The criteria to decide who is better than other academically is decided by exams in schools and colleges . The person who scores the highest marks is considered to be the most intelligent student , as simple as that . There are a number of types of examinations that are conducted all around the world . Some types are online examinations , mcq types examinations , omr based examinations .

The next part of examination and it can also be called the most important part of the examination is it's evaluation process . All the above listed examinations are evaluated either manually or in automated form . Another important type of examination is subjective examination . Subjective examinations are the ones that consist mostly of theory . Evaluating such examinations can be tiring and boring for the examiner especially when the number of students is quite large . The presented application intends to solves this problem by automatically checking the answer sheet of the student .

B. DATA COLLECTION

Data collection can be described as the process of firstly collecting and then measuring the information against the changes which are targeted in a well established system, which helps an individual to evaluate the situation and find answers to particular relevant questions. Data collection is such a part of research which exists everywhere in different domains of studies be it physical or social sciences, business, humanities etc. The main purpose of this phenomenon of Data Collection is that it helps in gathering quality material and evidence that would ultimately lead in the formation of concrete answers to the questions presented . The data that has been used in this project has been created from scratch .

4. METHODOLOGY

A. DEVELOPMENT ENVIRONMENT

The portion of the application which consists of Machine Learning to perform the analysis of the student's answer sheet has been conducted in Google Colab Notebook. This Notebook is widely used for performing and conducting projects and experimentations in the field of data science and visualisation of data. It is an open-source tool based on web. Most of the

machine learning applications out there make use of this notebook.

Visual Studio Code has been used for developing the Graphical User Interface through Flask . Visual Studio Code is an Integrated Development Environment developed by Microsoft. It provides us with a vast array of features which causes it to be one of the most preferred alternatives for development of applications in frameworks like Django and Flask which are based on Python . It also supports many other frameworks and is usually the go to option for any code designer for code editing.

B. ANALYSIS OF DATA

- The data set is collected in the very first step which consists of answers to the questions in the question paper.
- Upon collecting the data , all the text in the data is converted to lowercase .
- After the conversion to lower case , word tokenization is performed on the text . Word tokenization is the process of splitting a large sample of text into words. This is a requirement in natural language processing tasks where each word needs to be captured and subjected to further analysis like classifying and counting them for a particular sentiment etc. The Natural Language Tool kit (NLTK) is a library used to achieve this.
- Moving forward , next important step that is performed is removal of stopwords and punctuations . A stop word is a commonly used word (such as “the”, “a”, “an”, “in”) that a search engine has been programmed to ignore, both when indexing entries for searching and when retrieving them as the result of a search query. We would not want these words to take up space in our database, or taking up valuable processing time.
- At the end , stemming is applied to the dataset leading to a separate set of word . Stemming is the process of producing morphological variants of a root/base word. Stemming programs are commonly referred to as stemming algorithms or stemmers. A stemming algorithm reduces the words “chocolates”, “chocolatey”, “choco” to the root word, “chocolate” and “retrieval”, “retrieved”, “retrieves” reduce to the stem “retrieve”. Stemming is an important part of the pipelining process in Natural language processing. The input to the stemmer is tokenized words.

C. FORMATION OF WEBSITE

- The website has been designed using technologies like flask which is a python based framework .
- The website has a main page which provides the user to login as an admin or as a student .
- The admin window provides the option to upload the question paper

- The student window has got the facilities to upload his/her answer sheet .

D. WORKING

An automatic answer checker is an application that helps in checking the answer sheets submitted by the student in a similar manner as a human being .

This application has been built with an aim to check the subjective and long answer type questions and then allot marks to the students after performing the verification of the answers .

To carry out the whole operation , it is required by the user to store the answers of the questions so that the application can cross verify the answers from the answer sheet .

In this system , the admin has been provided the options to upload the question paper and also see which student has submitted the answer sheet .

When the student successfully logs into the system , he/she would be able to view the question paper and download it as well.

Upon completing all the answers , the student would be required to upload the answer sheet in pdf format for the system to evaluate it.

The answers of the questions may not be word to word same as in the answers given by the admin . Variations in the answers could be seen and that would be easily handled by the system and so the student need not worry about the incorrect checking of the answer sheet .

The system has been built with the help of various machine learning algorithms that at the end calculate marks of the students and return the same to them.

The application would be consisting of the following components :

Logging Facility:

Upon opening the main window of the application , the user would be greeted with two options namely – Admin and Student . User can choose any one option and proceed further .

Student log-in: For the student log-in the system uses the name of the student as ID and the registration number of the student as password to successfully log-in . If the student fails to log-in , then he would be required to re-enter the ID and password .

After successfully logging-in , the student would be able to see the question paper uploaded by the examiner as well as he would be able to download it.

The student would then be required to upload the answer sheet that he/she wants to be evaluated . Upon uploading the answer sheet , the student would be required to click the see marks button and then the marks of the student would be displayed along with the grade .

Admin log-in: The admin is a teacher and so this log-in has been configured for them .

The admin would be required to input his/her name as ID and a specific password assigned to the as password to be able to log-in into the system.

The admin will have the option to upload the question paper and see which student has submitted his/her answer sheet and which student hasn't.

Answer Checking Process: From the training input that would be provided to the machine learning algorithms , the system would perform word tokenization , remove stopwords and punctuation and also perform stemming which at the end would yeild a set of separate words . From these set of words , the system would then decide a set of keywords that must be in the answer . Depending on the number of keywords in the answer , appropriate marks would be given to the student .

5. ALGORITHM

- Step 1: Start
- Step 2: Main window opens
- Step 3: Login as Admin or a Student
- If user logs in as a Student , goto step 4
- If user logs in as an Admin , goto step 8
- Step 4: Student window opens
- Step 5: View / download question paper
- Step 6: Upload answer sheet.
- Step 7: Click see marks button
- Step 8: Admin window opens
- Step 9: Upload the question paper
- Step 10: See responses of students

6. HELPFUL HINTS AND FIGURES

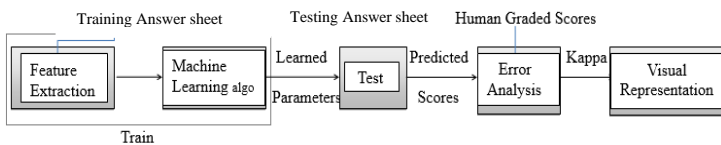


Figure 2. System Architecture

FIGURE

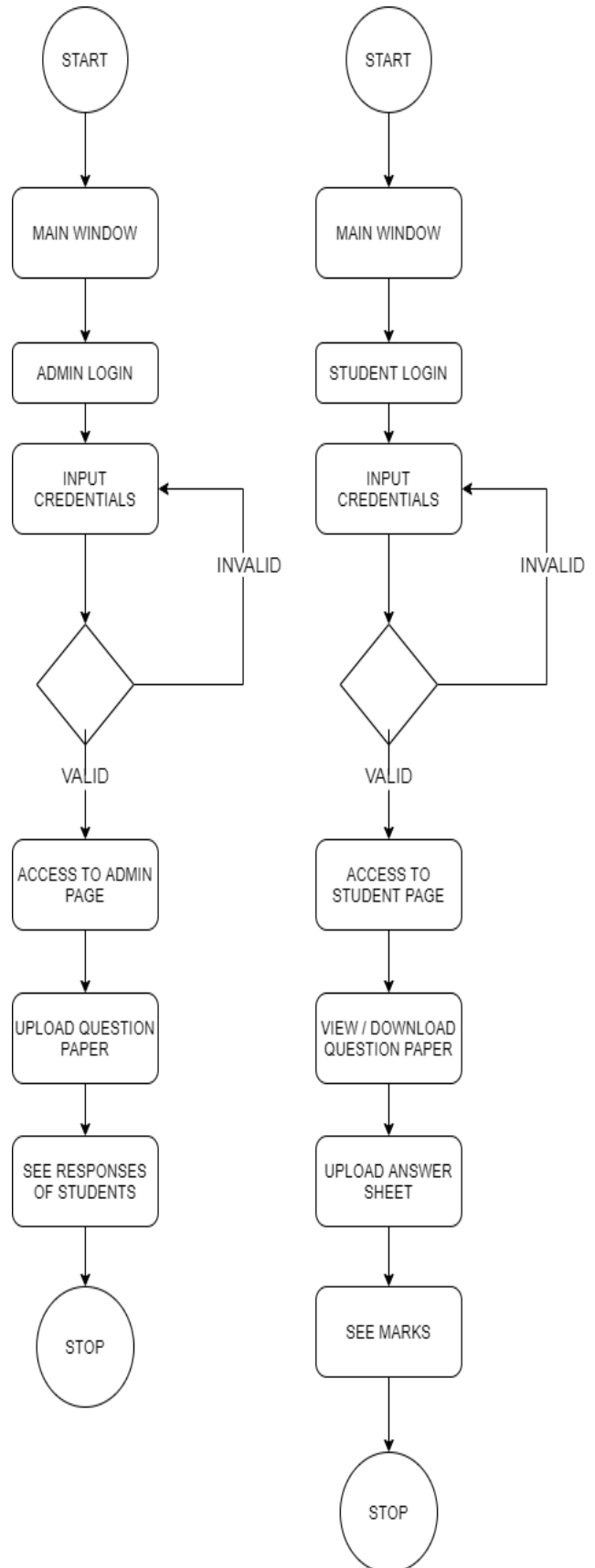


Figure 1. Functionality Flow Chart

7. EVALUATION

The project proposed has been evaluated on terms of various aspects to get a deep insight into the accuracy of our proposed method . The first aspect is on the basis of Quality . The Automatic Answer Checker was evaluated on the basis of quality to measure the accuracy of our system. The second aspect on which evaluation was conducted is performance which was done to get a deep insight into the comparison of automated answer checking method with traditional answer checking method .

7.1. QUALITY EVALUATION

A survey was conducted whose sole purpose was to check the quality of the automatic answer checker system . This survey was conducted on a bunch of university students as well as on the faculties of various departments in the university . Choosing such a sample was obvious as our automatic answer checker is related to their domain . There was almost an equal ratio of male and female among the sample chosen for survey . When all the results of the survey were analysed , we got to know that 85% of the survey participants strongly agreed and 15% of them agreed to the fact that our designed system was giving precise results when they used it . 77% of the survey participants strongly agreed and 23% of them agreed to the fact that the efficiency of our designed system was fast enough to fulfill all of their tasks. 97% of the survey participants strongly agreed and 3% of them agreed to the fact that designed system was quite user friendly and easy to use even when they had not used any similar type of thing before. 95% of the survey participants strongly agreed and 5% of the participants agreed to the fact that all the operations of user were fulfilled by the designed system . Thus , to sum up , each one of the person who participated in the survey was very much satisfied by the designed system. Therefore , our designed system passed the quality evaluation test pretty well .

7.2. AUTOMATIC SYSTEM V/S TRADITIONAL SYSTEM

This automatic answer checker has been experimented on a couple of students of the university . This study was done with the sole purpose of finding out the measures of evaluation such as precision , accuracy , recall , F-measure which would further help us in comparing the effectiveness of automated answer checking with the traditional approach of checking the answer sheets.

The following table shows the evaluation measure namely recall , accuracy , precision , F-measure . The True Positive represents the total number of questions that were answered accurately and which were actually the correct answers and so they are then termed as positive . The True Negative represents the total number of questions that were answered as incorrect by our system and in reality as well , they were incorrect and so they are then termed as negative. The False Positive represents the number of questions that were declared by our system as correct but in reality they were incorrect . False Negative

represents the total number of questions that were declared by our system as incorrect but in reality they were correct .

Evaluation Measure	Expression
Recall	$TP / (TP+FN)$
Precision	$TP / (TP+FP)$
Accuracy	$(TP+TN) / (TP+TN+FP+FN)$
F-measure	$(2*Recall*Precision) / (Recall + Precision)$

Table 1. Evaluation Measures

Figure depicts the difference between the automated approach and the traditional approach on the basis of evaluation measures .The Recall is also known as True Positive Rate and was observed to be 0.9712 for the automated approach where as it was observed to be 0.9443 in the case of traditional approach . The Precision is also called Positive Predictive Value and it was observed to be 0.9062 in the automated approach where as it was observed to be 0.9771 in the case of traditional method . The Accuracy is also called True Results and it was observed to be 0.8800 in the automated approach where as it was observed to be 0.9500 in the case of traditional method . F-measure can also be stated as the geometric mean of Recall and Precision and it was observed to be 0.9311 in the automated approach where as it was observed to be 0.9656 in the case of traditional method .

From this study , we can easily make out that the system that we have developed is quite near to the traditional method for checking the answer sheets . On comparing both the methods on the basis of evaluation measure , we can conclude that the Recall , Precision , Accuracy , F-measure are higher in the case of traditional approach than the automated approach .

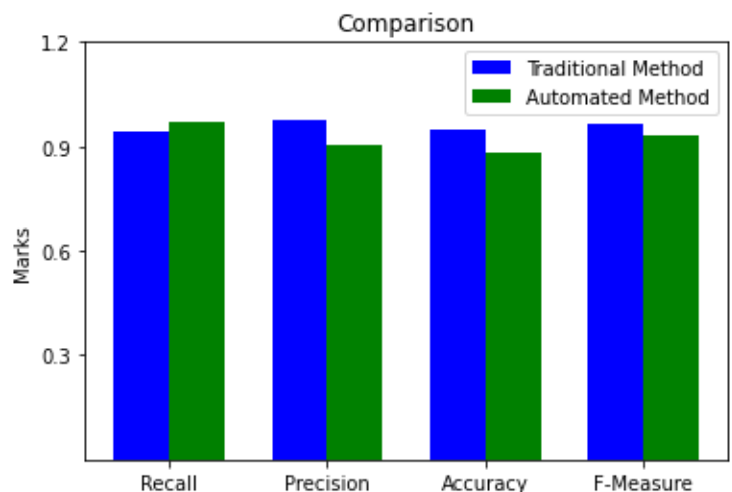


Figure 3. This is the performance analysis between the automated approach and the traditional approach.

The Below bar graph clearly depicts the difference in the marks given by the system and the marks given to the student following the traditional approach . As it can be clearly seen that the difference between the marks given by the system and the marks given in the case of traditional method is marginal , it clearly shows that our system has been designed with high accuracy and precision . This further goes on in proving the reliability of our designed project in correctly predicting the marks of the student .

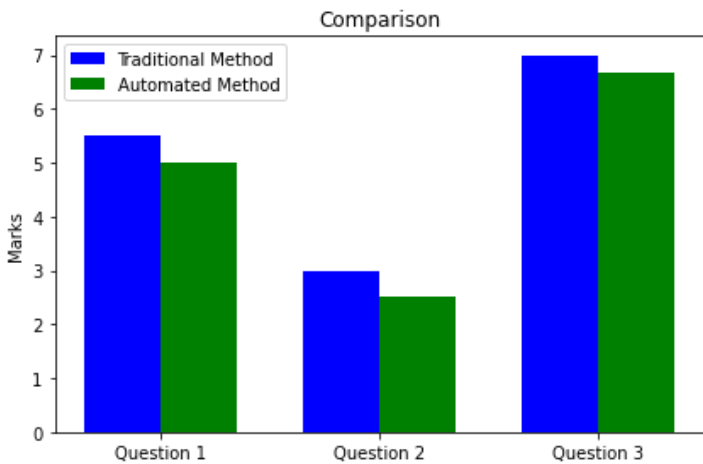


Figure 4. Comparison between the marks given by system with the marks given in traditional method .

Upon seeing the marks predicted by our designed project and when we compare them with the marks that would have been allotted to the student had the checking been done following the traditional method , we found that the Mean Square Error in correctly predicting marks comes out to be 0.205 . Observing the value of the mean square error , we can no doubtly conclude that our system has been designed with quite high precision and can be considered to be highly reliable when it comes to predicting the marks of the students correctly .

8. RESULTS



Figure 5. Website Snapshot

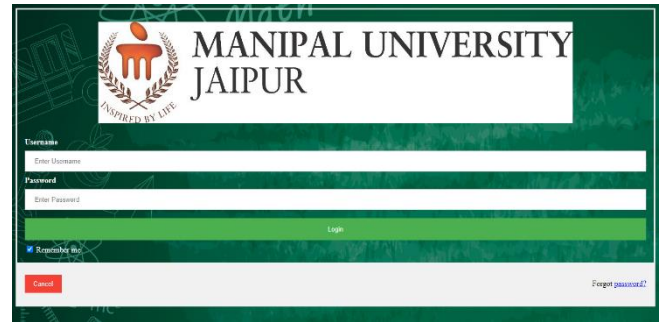


Figure 6. Website Snapshot

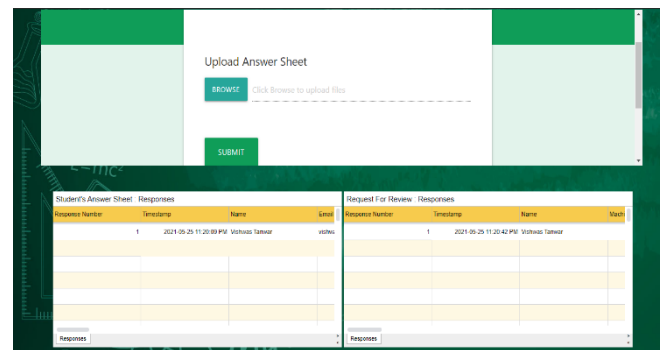


Figure 7. Website Snapshot



Figure 8. Website Snapshot

9. CONCLUSION

The project report whose title is Automatic Answer Checker has now reached it's last stage . The application has been made keeping every possible chance of error in mind and so the system is quite efficient and reliable. The application has a very unique property of being robust in nature due to which there are many ways of implementing improvisations in the application in the near future . The

application would soon be approved and authenticated and then implemented . Future work would be consisting of creating an algorithm for the assessment whose purpose would be to find all the syntax errors in our keywords and then we would be investigating it for high performance and high equality for addressing them .

10. ACKNOWLEDGEMENT

With a very deep essence of gratitude , I would like to acknowledge my project guide Dr. Rishi Gupta for his regular help in the project as well as the encouragement that he gave to me while pursuing this project . His regular assistance has been an important factor in the completion of the project.

11. REFERENCES

[1] Manvi Mahana, Mishel Johns, Ashwin Apte. (2012). Automated Essay Grading using Machine Learning, Stanford University.

[2] Likic, A., & Acuna, V.(n.d.). Automated Essay Scoring, Rice University.

[3] Lakshmi Ramachandran, Jian Cheng & Peter Foltz "Identifying Patterns For Short Answer Scoring Using Graph-based Lexico-Semantic Text Matching"

[4] Judy McKimm, Carol Jollie, Peter Cantillon, "ABC of learn-ing and teaching Web based learning" <http://www.bmj.com/>

[5] Yuan, Zhenming, et al. A Web-Based Examination and Evaluation System for Computer Education. Washington, DC: IEEE Computer Society, 2006.

[6] Effie Lai-Chong Law, et al. Mixed-Method Validation of Pedagogical Concepts for an Intercultural Online Learning Environment. New York: Association for Computer Machinery, 2007.

[7] Lan, Glover, et al. Online Annotation- Research and Practices. Oxford UK: Elsevier Science Ltd, 2007.

[8] Sophal Chao and Dr. Y.B Reddy Online examination Fifth International Conference on Information Technology: New Generations, 2008

[9] Hanumant R. Gite, C.Namrata Mahender "Representation of Model Answer: Online Subjective Examination System" National conference NC3IT2012 Sinhgad Institute of Computer Sciences Pandharpur.

[10] J. Dreier, R. Giustolisi, A. Kassem, P. Lafourcade, G. Lenzini, and P. Y. A. Ryan. Formal analysis of electronic exams. In SECYPT'14. SciTePress, 2014.

[11] P. Kudi, A. Manekar, K. Daware, and T. Dhattrak, "Online Examination with short text matching," in Wireless Computing and Networking (GCWCN), 2014 IEEE Global Conference on, 2014, pp. 56–60.

[12] K. Woodford and P. Bancroft, "Multiple choice questions not considered harmful," in Proceedings of the 7th Australasian conference on Computing education-Volume 42, 2005, pp. 109–116.

[13] X. Hu, and H. Xia," Automated Assessment System for Subjective Questions Based on LSI," Third International Symposium on Intelligent Information Technology and

Security Informatics, Jinggangshan, China, pp. 250-254, April 2010.

[14] M.F. Al-Jouie, A.M. Azmi, "Automated Evaluation of School Children Essays in Arabic," 3rd International Conference on Arabic Computational Linguistics, 2017, vol.117, pp.19-22.

[15] A.Kashi, S.Shastri and A. R.Deshpande, "A Score Recommendation System Towards Automating Assessment In Professional Courses," 2016 IEEE Eighth International Conference on Technology for Education,2016,pp.140-143.

[16] T. Ishioka and M. Kameda, "Automated Japanese essay scoring system: Jess," in Proc. of the 15th Int'l Workshop on Database and Expert Systems Applications, 2004, pp. 4-8. [17]

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.

[18] Vkumaran and A Sankar," Towards an automated system for short-answer assessment using ontology mapping," International Arab Journal of e-Technology, Vol. 4, No. 1, January 2015.

[19] R. Siddiqi and C. J. Harrison, "A systematic approach to the automated marking of short-answer questions," Proceedings of the 12th IEEE International Multi topic Conference (IEEE INMIC 2008), Karachi, Pakistan, pp. 329-332, 2008.

[20] M. J. A. Aziz, F. D. Ahmad, A. A. A. Ghani, and R. Mahmud, "Automated Marking System for Short Answer examination (AMSSAE)," in Industrial Electronics & Applications, 2009. ISIEA 2009. IEEE Symposium on, 2009, pp. 47-51.

[21] R. Li, Y.Zhu and Z.Wu,"A new algorithm to the automated assessment of the Chinese subjective answer," IEEE International Conference on Information Technology and Applications, pp.228 – 231, Chengdu, China, 16-17 Nov. 2013.

[22] X.Yaowen, L.Zhiping, L.Saidong and T.Guohua," The Design and Implementation of Subjective Questions Automatic Scoring Algorithm in Intelligent Tutoring System," 2nd International Symposium on Computer, Communication, Control and Automation, Vols. 347-350, pp. 2647-2650, 2013.

[23] Y. Zhenming, Z. Liang, and Z. Guohua, "A novel web-based online examination system for computer science education," in 2013 IEEE Frontiers in Education Conference (FIE), 2003, vol. 3, pp. S3F7– 10.

[24] M. S.Devi and H.Mittal,"Machine Learning Techniques With Ontology for Subjective Answer Evaluation," International Journal on Natural Language Computing, Vol. 5, No.2, April 2016.