IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

EVALUATING DESCRIPTIVE ANSWER ASSESSMENT SYSYEM WITH MACHINE LEARNING

[1]Miss Ashlesha S. Phalke, [2]Dr.G.R.Bamnote, [3]Prof. Ms. S.W. Ahmad

[1]Master of Engineering Student, [2]Head of Department, [3]Assistant Professor

[1]Computer Science & Engineering,

[1]Prof. Ram Meghe Institute of Technology & Research, Badnera, Amravati, India

Abstract: Online subjective tests are very rarely proposed .all the online subjective exams set for the students are pen and paper-based evaluated by the teachers manually. This paper presents a survey on the effective way of online subjective tests. In this, the answers are unstructured data that have to be evaluated. The calculation is based on the semantic similarity between the faculty's answers and student's answers. Different techniques i.e. TF-IDF (Term Frequency-Inverse Document Frequency), keyword extraction are compared and new techniques like LSA (Latent Semantic Analysis), SOM(Self Organizing Map Clustering) are used to develop an evaluate subjective test assessment of the text.

Index Terms - Descriptive type answer, Keyword Extraction Algorithm, Latent Semantic Analysis, Self Organizing Map.

I. INTRODUCTION

In the day to day Education System, the complete entrance exam in all different fields is objective tests. Objective tests are all not that sufficient 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 a 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 to test the knowledge of the student, descriptive answers play an important role. But the valuations of descriptive answers are mostly manual and that becomes too hectic for faculty. To solve this problem of manual checking of subjective answers I have proposed an online evaluation system of descriptive type answers.

Descriptive answers vary from student to student, so in this paper to extract the meaning from the various answers the concept of semantic similarity is used. The Faculty needs to answer along with some compulsory expected keywords in it. The answer will be pruned, stemmed which will reduce the size of the answer, and then transformed into vectors and matrix form. Depending upon the keywords used in the answer marking will be done. For this, the mandatory text will be extracted from the database by using various methods such as Term Frequency/ Inverse Document Frequency (TF/IDF). To extract the meaning from the text, techniques like LSA (Latent Semantic Analysis), SOM(Self Organizing Map) are used. The TF/IDF method with LSA semantic work is suitable for information recovery, text classification, etc. The marks will be assigned using the Cosine Similarity depending upon the value of records. In this paper, the length of the answers will also be taken into consideration while allotting the marks. Then the result will be mailed to the students.

OBJECTIVE:

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

LITERATURE REVIEW:

In any study, the review of previous studies is considered as important for getting understanding the problem, the methodology followed to identify the unexplored part of the field of study under consideration. In this regard, some of the relevant studies have been reviewed in the present study. So following is the list of research paper which research in the field of the answer evaluation system.

Birpal Singh J. Kapoor and Shubham M. Nagpure in june 2020 researched "An Analysis of Automated Answer Evaluation

System Based on Machine Learning", have analysed the automated answer using keywords and automated grading also Declared in this paper using content-based similarity, natural language processing [1].

Ankita Patil and Prof. Achamma Thomas in 2017 researched "A Survey of Effective Technique for Subjective Test Assessment". In this paper, the answer is unstructured data which has to be evaluated. The evaluation is based on the semantic similarity between the model answer and the user answer. Keyword extraction methods such as TF-IDF, CRF model & Query focused method are used. Ontology method and Context-based method are used author conclude that CRF model & LSA is an effective method for extraction of keywords & assessment of the subjective test [2].

Meena K. and Lawrance R in 2016 researched "Semantic Similarity Based Assessment of Descriptive Type Answer". In this paper assessment algorithm uses for semantic similarity for the evaluation of detailed type answer. Stemmed words are converted to vectors using a semantic method, latent semantic analysis (LSA). Vector obtained from the sematic method are clustered using the self-Organizing Map. Cosine Similarity is used to measure the similarity between two vectors based on that value returned by the similarity measures and marks will be awarded [3].

P.Willett in 2006 researched "The Porter stemming algorithm: then and now", Program, .In this paper author presented a simple algorithm for Stemming English Language Word [4].

Vimala Balakrishnan and Ethel Lloyd-Yemoh in Aug 2014 researched "Stemming and Lemmatization: A Comparison of Retrieval Performances", In this paper They compare document retrieval precision performances based on language modeling techniques, particularly stemming and lemmatization. Stemming is a procedure to reduce all words with the same branch to a common form whereas lemmatization removes inflectional endings and returns the base or dictionary form of a word. Comparisons were also made between these two techniques with a baseline ranking algorithm (i.e. with no language processing). A search engine was developed and the algorithms were tested based on a test collection. Both mean average precisions and histograms indicate stemming and lemmatization to outperform the baseline algorithm. Stemming techniques including the Paice/Husk stemmer, Porter's stemmer, and Lovin's stemmer. As for the language modeling techniques, lemmatization produced better precision compared to stemming, however the differences are insignificant. Overall the findings suggest that language modeling techniques improves document retrieval, with the lemmatization technique producing the best result [5].

Shweta M. Patil and Prof. Ms. Sonal Patil in 2014 researched "Evaluating the student descriptive answer using Natural Language Processing". In this, the system can perform grading as well as provide feedback for students to improve in their performance. The techniques for automatic marking of free-text responses are categories into three main kinds, Statistical, Information Extraction and Full Natural Language Processing [6].

Ari Aulia Hakim, Alva Erwin, Kho I Eng, Maulahikmah Galinium, and Wahyu Mulia in 2014 researched "Automated Document Classification for News Article in Bahasa Indonesia based on Term Frequency Inverse Document Frequency (TF-IDF) Approach" In this paper they works on the TF-IDF algorithm which create a classifier that can classify the online Articles [7].

P.Y. Hui and H.Y.Meng in 2014 researched "Latent Semantic Analysis for Multimodal User Input with Speech and Gestures. Audio, Speech, And Language Processing IEEE/ACM Transactions", used LSA for semantic explanation of a multimodal language with speech and gestures [8].

Meena K and Lawrance R in 2014 researched "Evaluation of the Descriptive answers using the hyperspace Analog to language algorithm and self-organizing Map". These works focused on the online evaluation of the descriptive answer which will eliminate the discrepancy in that manual evaluation. The HAL algorithm is used to separate categories of the word [9].

Sonal N. Deshmukh and Ratandeep R. Deshmukh in Jan 2014 researched "Cosine Similarity for Substituted Text Detection". In this paper deals with Detection of Substituted text in a sentence using statistical techniques and cosine similarity [10].

Menaka S and Radha N in 2013 researched "Text Classification using Keyword Extraction Technique". They have classified the text using a keyword extraction algorithm. The keywords are extracted using TF-IDF and WordNet. 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 research work, the words which have the maximum similarity are selected as keywords [11].

Yuan-Chao Liu, Chong Wu, Ming Liu in 2011 researched "Research on fast SOM Clustering for text information" .In this

Paper fast SOM clustering technology for Text Information In the feature extraction phase, several methods were discussed to find the semantic similarity. For this proposed fast SOM clustering technology for text information [12].

Sungjick Lee and Han-joon Kim in 2008 researched "News Keyword Extraction for Topic Tracking", proposed a conventional TFIDF model for keyword extraction. It involves cross-domain filtering and table term frequency (TTF) for extraction [13].

Stephen Robertson in 2004 researched "Understanding inverse document frequency: on theoretical arguments for IDF" explains the understanding concepts of IDF [14].

Teuvo Kohonen in May 2000 researched "Self Organization of a Massive Document Collection". In this paper describes the implementation of a system that can organize vast document collections according to textual similarities. It is based on the self-organizing map (SOM) algorithm [15].

Helena Ahonen, Oskari heinonen in 1998 researched "Applying Data Mining Technique for Descriptive Phrase Extraction in Digital Document Collection". In this paper data mining method apply to text analysis task such as descriptive phrase extraction, they present a general framework for text mining, the framework follow General Knowledge Discovery (KDD) to preprocess texts based on the intended use of the discovered result & introduce a weighting scheme that helps in puring redundant or non – descriptive phrases [16].

III.PROPOSED SYSTEM ANALYSIS / DESIGN

The Dissertation work is an educational-based system. In this system, the online exam will be descriptive, unlike other objective online exams. The system will be managed by the exam system admin. The teaching staff will conduct exams and students will be involved in answering the tests.

3.1.1 Describe Main Objective

- To develop a descriptive type of answer checker filter.
- To develop auto checking system for descriptive types of answer
- To reduce the time consumption in results generation for university exams
- Provide an efficient and fast way for descriptive types of examination

3.1.2 Problem Definition

Nowadays the online examination is only getting conducted on objective types of question and answers. There is not any system existing to get the answering for the descriptive type of examination so that it is necessary to stop the traditional ways of taking an exam because it is time-consuming so it is best to get the solution which will automatically generate results by efficient pattern matching technique. Old-style exams taking a long time for results evaluation and will require large human resources to manipulate.

3.2 DFD

The DFD diagram explains that the faculty will provide questions along with the answers and keywords 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 they 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 be mailed to the students.

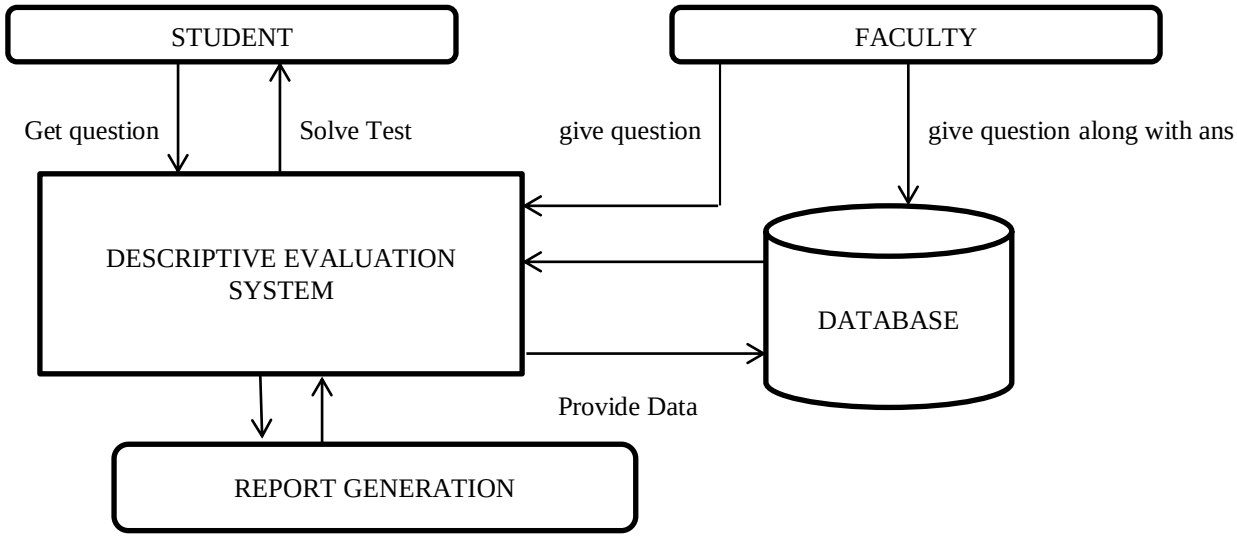


Fig. Dataflow Diagram

IV. PROPOSED SYSTEM

This 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.

4.1 Module

A module can encapsulate code and data to implement a particular functionality. It is an interface that agreements the client to access its functionality uniformly and it is easily pluggable with another module that expects its interface. So satisfying all the conditions of what a module is our system will have the following modules:

- Exam System Admin
- Student
- Online Exam Portal
- Evaluation Process

EXAM SYSTEM ADMIN: This is the module of an admin, the person or a team under whom the entire system will be controlled. The admin will monitor all the actions of the system. The subjects will be registered by the admin. He will register the student and create the user id and password for each student. He will supervise all the activities of the system and will have the authority to check all the details of the students too. He will allot each student to a particular subject to solve the papers for that subject

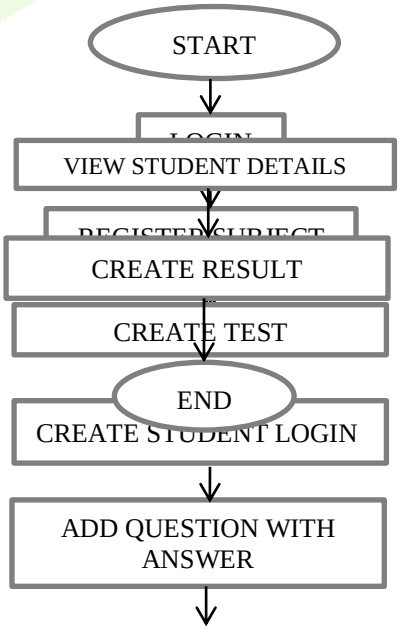


FIG. FLOWCHART OF EXAM SYSTEM ADMIN MODULE

Student: Students will be registered by the admin. They will be provided with username and password to login in the system. If the student is not registered he/she can contact to admin. Once the student will login he/she will be able to see the exam scheduled by the staff. According to the time the test is available the student will give the test. In the test he/she is expected to give detailed answers of the questions. All the questions will be descriptive type. After finishing the test he/she will submit the test. The test results will then be mailed to the student on their respective mail-id.

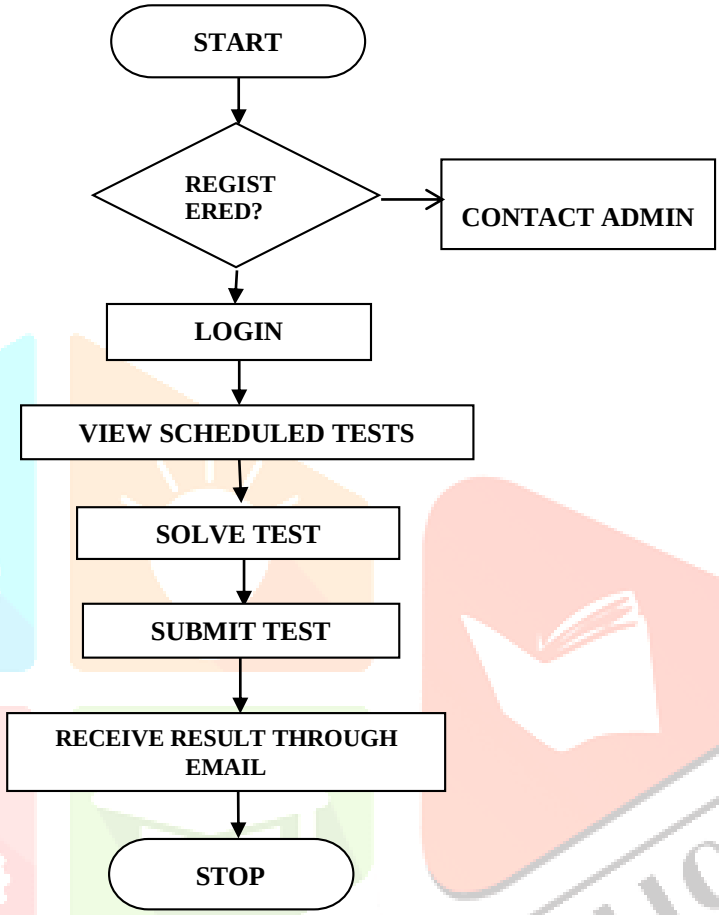


FIG. FLOWCHART OF STUDENT MODULE

Online Descriptive Exam Portal and Evaluation

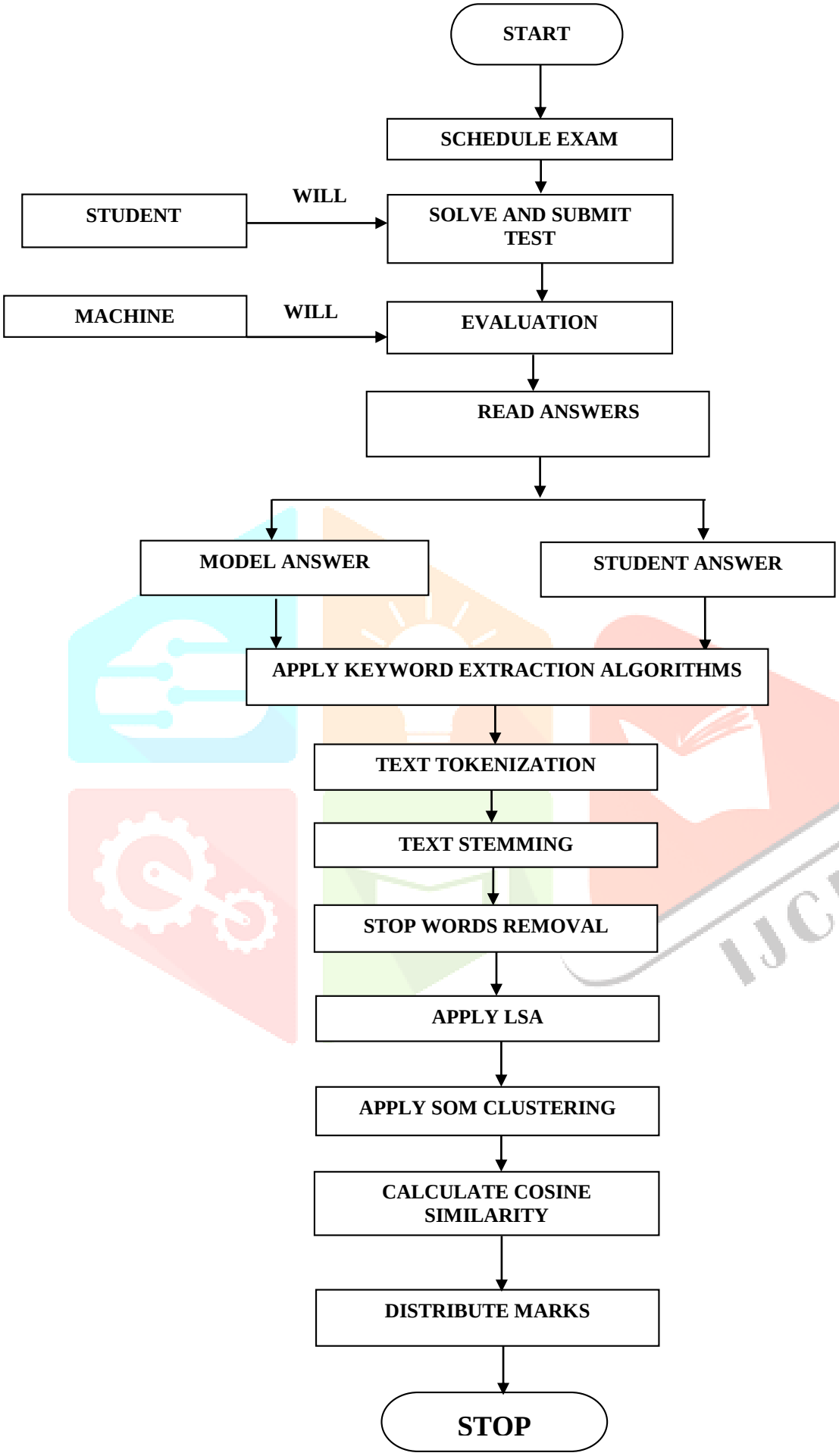


Figure: Online Exam Portal and Online Evaluation module

Exam Portal: The exam portal will be as such the admin will schedule exams. The students will solve the tests and submit them. Then the complete evaluation of the answers in the test papers will be done by the machine. The evaluation process will have both the teacher's answer i.e. model answer and students' answers. Every student's answer will be different. The ideal answer will be

stored in the database. When the student submits its test those answers written by the student will be compared with that of the model answer. The system will read both the model and student answer and then extract the keywords using a keyword extraction algorithm such as TF/IDF (text frequency/ inverse document frequency). Keyword extraction is information retrieval that automatically identifies the best terms in the given document. These terms can be key phrases, key terms, or just keywords. In this approach, keywords are extracted. Keywords are easy to define as they are widely used within information retrieval (IR).

V. TERM FREQUENCY- INVERSE DOCUMENT FREQUENCY (TF-IDF) ALGORITHM:

Term Frequency (TF):

TF measure the frequency of word (w) in a document (d).

TF is defined as the ratio of word's occurrence in a document to the total number of words in a document.

$$\textbf{TF (w, d)} = \frac{\textbf{Occurrence of w in document}}{\textbf{Total no. of words in document}}$$

Inverse Document Frequency (IDF)

It is measure of the importance of word.

IDF provides weightage to each word based on its frequency in the corpus D

$$\textbf{IDF (w, D)} = \frac{\textbf{Total no. of document (N) in corpus D}}{\textbf{Number of document containing w}}$$

Term Frequency/ Inverse Document Frequency

TF-IDF is a numerical measure that evaluates how relevant a word is to a document in a collection of documents. This is complete by multiplying two metrics, how many times a word appears in a document and the inverse document frequency of the words across a set of documents.

It is the product of TF & IDF.

$$\textbf{TFIDF (w, d, D) = TF (w, d) * IDF (w, D)}$$

It is most importantly in automated text analysis, and is very useful for scoring words in machine learning algorithms for Natural Language Processing (NLP).

Tokenization: Tokenization is the act of breaking up an order of strings into pieces such as words, keywords, phrases, symbols, and other elements called tokens. Tokens can be individual words, phrases, or even whole sentences. In the process of tokenization, several characters like punctuation marks are rejected. The tokens become the idea for another process like parsing and text mining.

Stemming: In information retrieval, stemming is the process of reducing inflected (or sometimes derived) words to their word stem, base, or root form generally a written word form. 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.

Stop Words Removal: Sometimes, some very common words which would appear to be of little value in helping select documents matching a user's need are excluded from the vocabulary completely. These words are called stop words. The common method for determining a stop list is to sort the terms by collection frequency (the total number of times each term appears in the document collection), and then to take the most common terms, often hand-filtered for their semantic.

Latent Semantic Analysis (LSA): Latent Semantic Analysis is a natural language Processing technique used for analyzing the relationship between the set of responses and the terms. LSA exactly means examining the documents to discover the core meaning of those documents. It is a completely mathematical method for mining and gathering associations of words in the documents and returns a matrix. In this matrix, rows represent the unique terms and columns represent eais the amount of answers. This matrix (m) contains a number in each cell which specifies the exact number of appearances of every word in all answers.

Self-Organizing Map Clustering (SOM): Self -Organizing Map is an unsupervised learning process which is normally used in pattern recognition, image processing, natural language processing and data mining. SOMs map multidimensional data onto lower dimensional subspaces where geometric relations between points identify their similarity. The main advantage of using a SOM is that the data is easily understood. The reduction of dimensionality and grid clustering makes it easy to detect similarities in the

data. SOMs factor in all the data in the input to generate these clusters and can be changed such that certain pieces of data have more/less of an effect on where an input is placed. SOMs are capable of handling several types of classification problems while providing a useful, cooperative, and intelligible summary of the data.

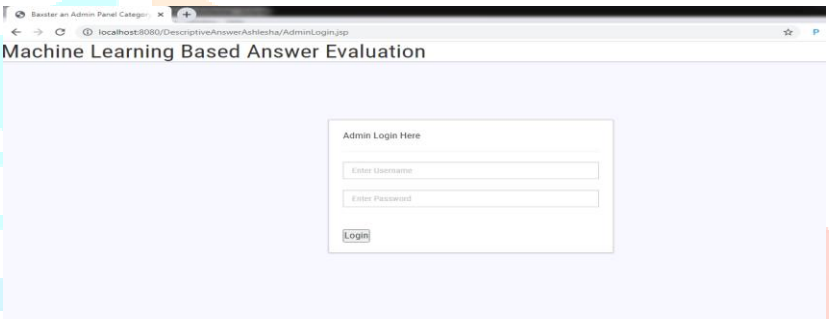
Cosine similarity: Cosine similarity is used to measure the comparison between two Vectors. It will produce a value that tells how two answers are related by considering at the angle. Cosine similarity measurement is used to match the key vectors with the answer vectors of the students. Cosine similarity function is given by

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}.$$

Where 'A' will be the answer by a student and 'B' will be the corresponding model answer. This function is used to calculate the angle between an answer and the corresponding keyword. If the value returned by cosine function is 1, both answers are similar. Based on the value returned by the similarity measure, marks will be awarded.

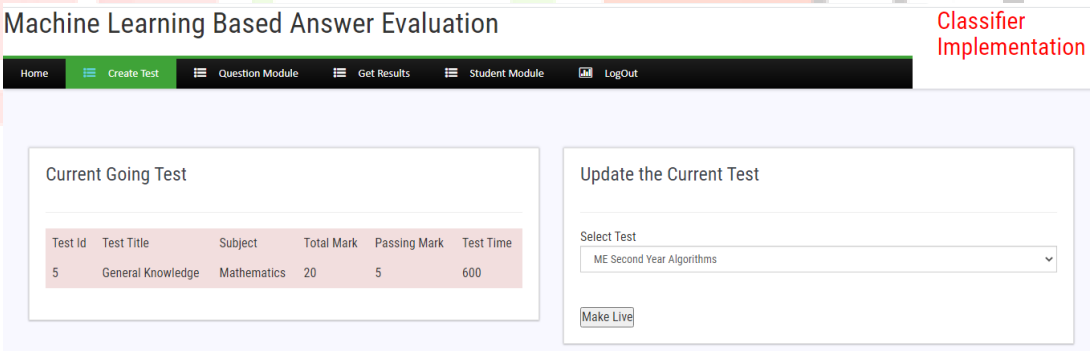
VI. SYSTEM OUTPUT SCREENSHOT

a) Admin Page



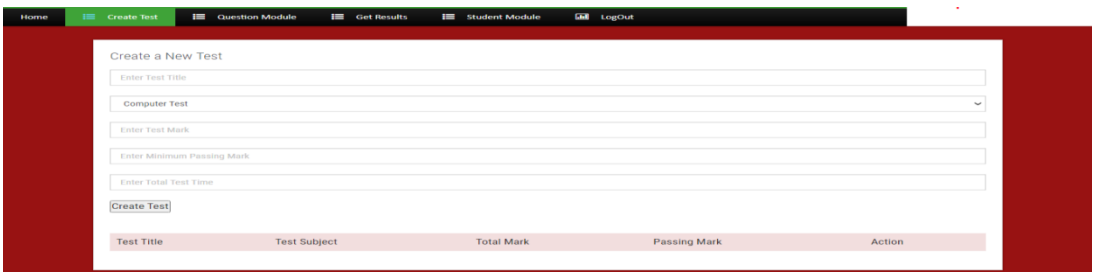
Screenshot a: Admin Login Page

b) Home Page



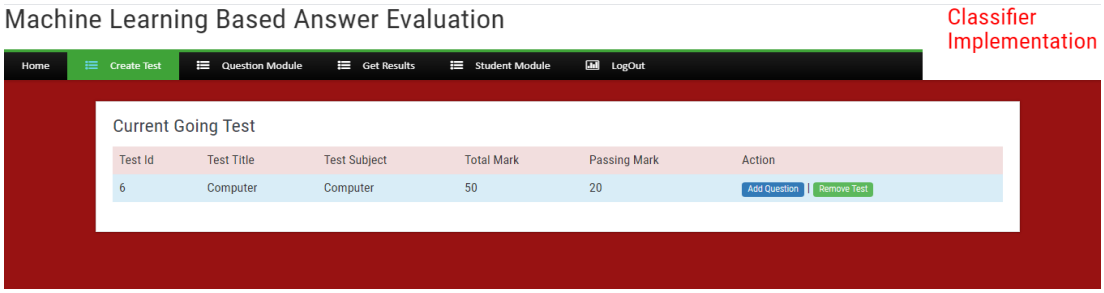
Screenshot b: Home Page

c) Create Test



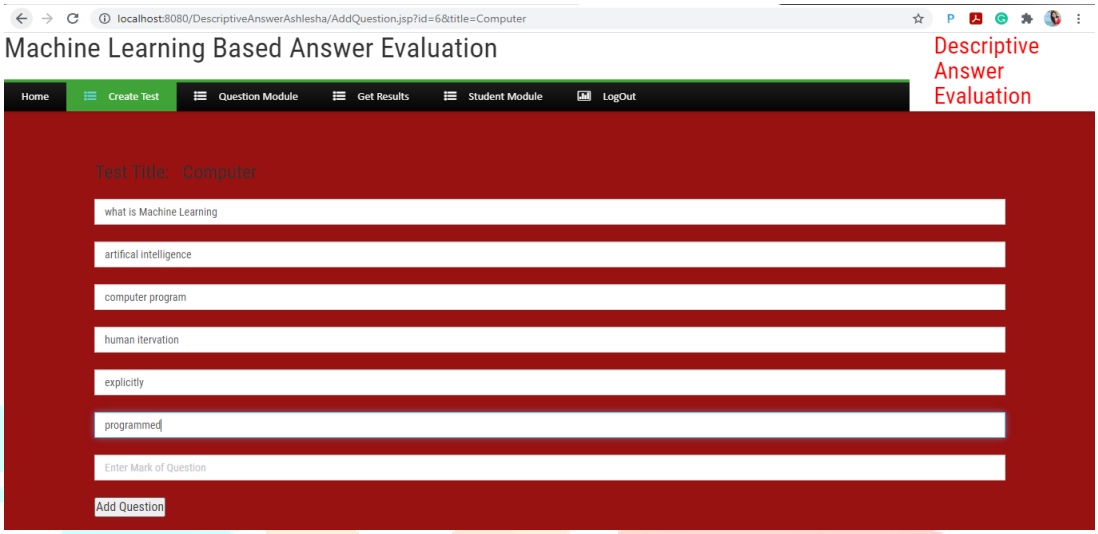
Screenshot c: Create Test

d) Show Test



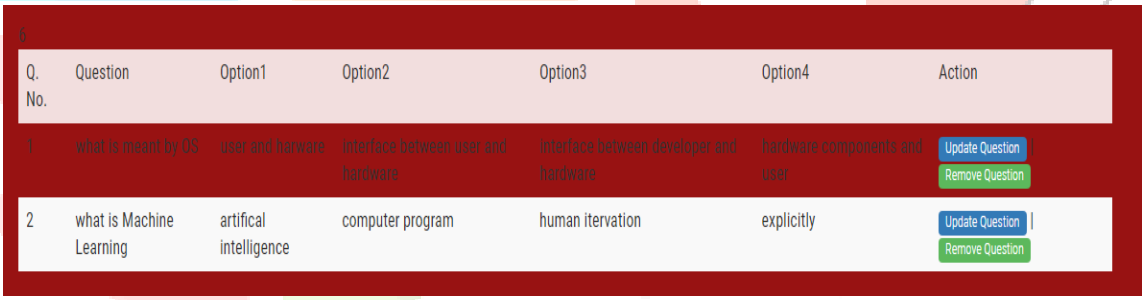
Screenshot d: current ongoing tes

e) Add Question



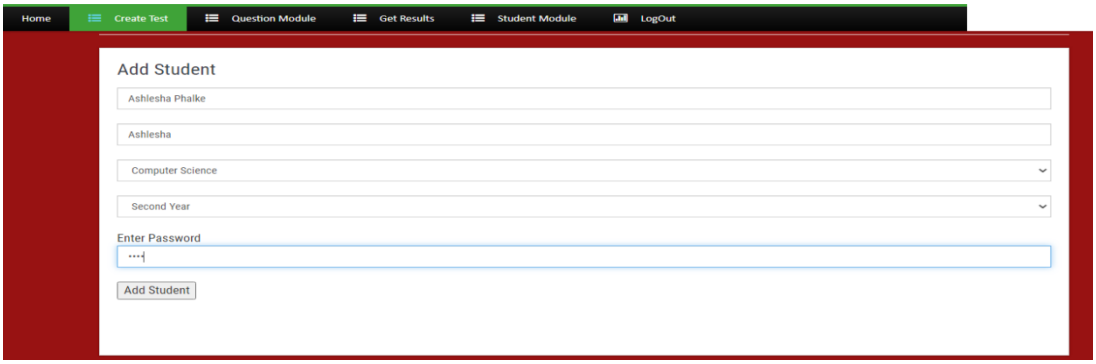
Screenshot e: add question

f) Update & Remove Question



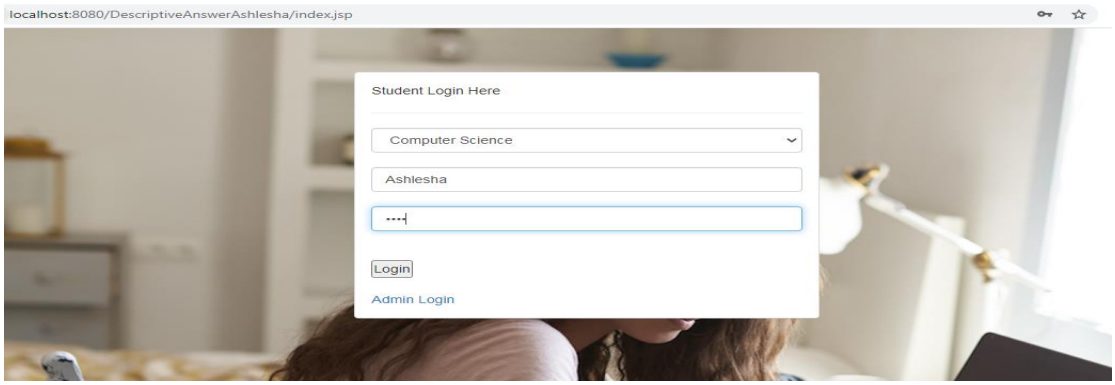
Screenshot f: update and remove question

g) Add student



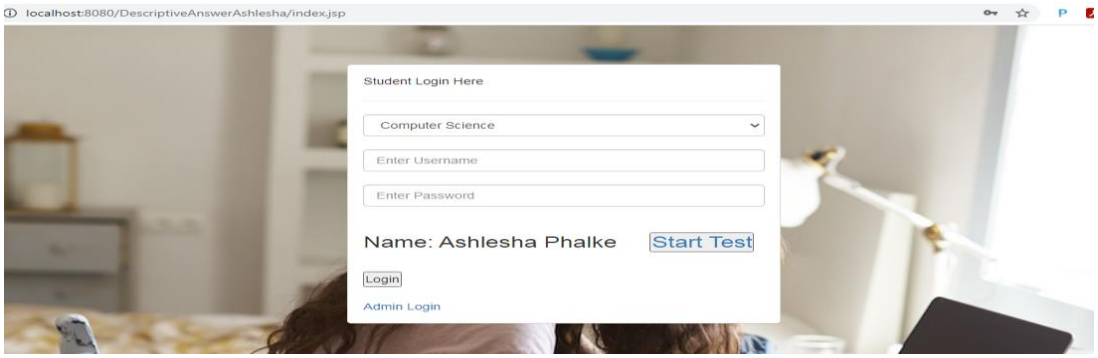
Screenshot g: Add Student Details

h) Student Login



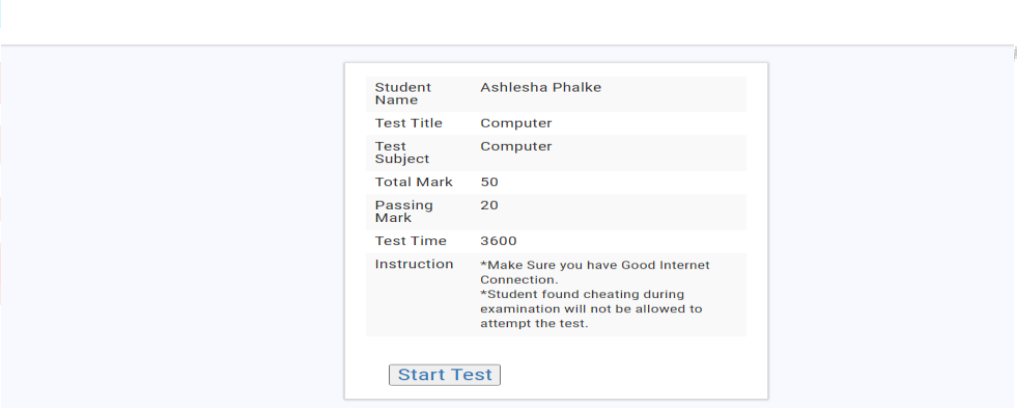
Screenshot h: Student Login page

i) Start Test



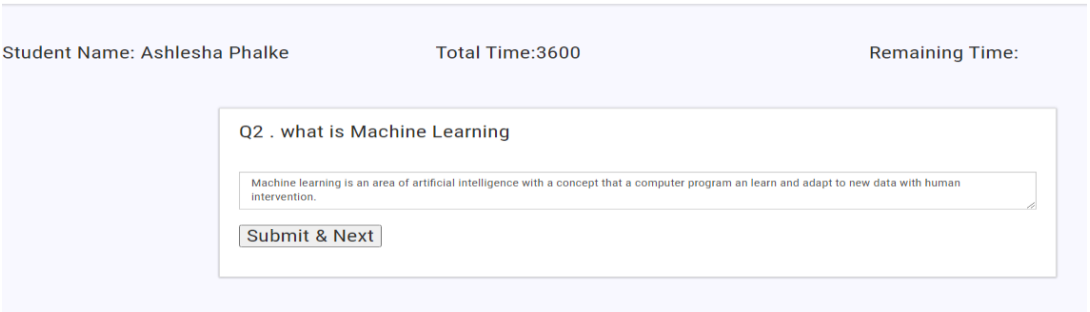
Screenshot i: Start Test Page

j) Guidelines for Test



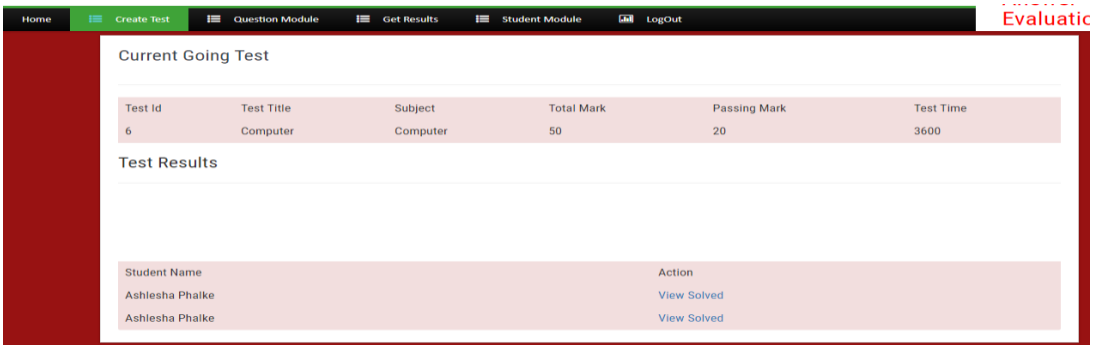
Screenshot j: Guidelines before Start the Exam

k) Test Screen



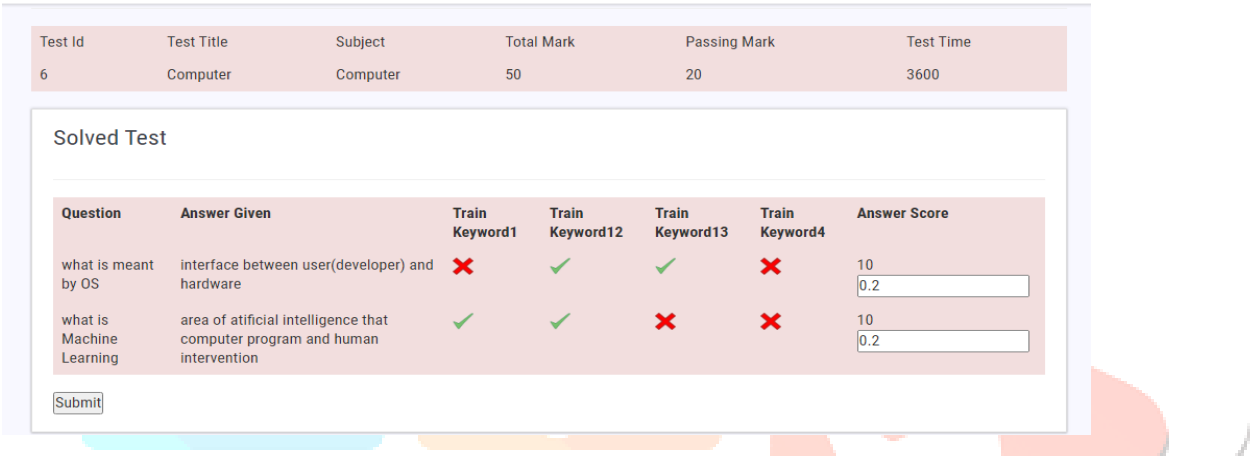
Screenshot k: Test Screen

l) Display the test details



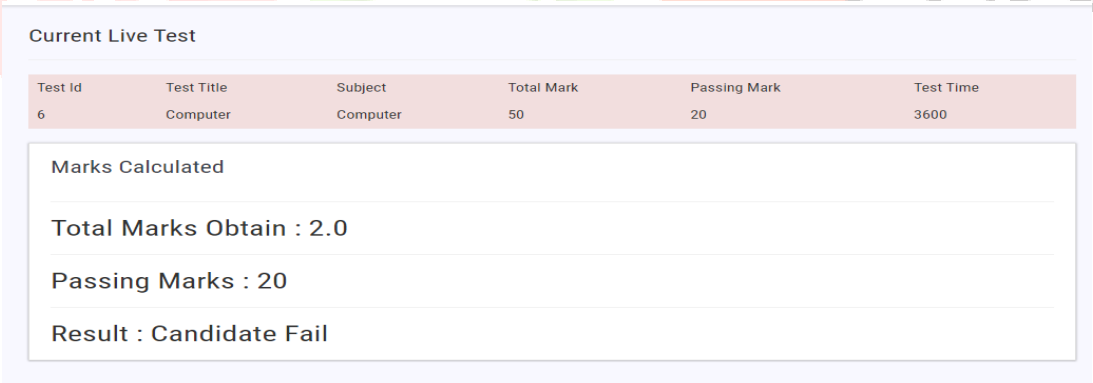
Screenshot l: Record of No of Student solving test

m) Answer will be Check



Screenshot m: Answer checking

n) Result will be display



Screenshot n: Result

ADVANTAGES

- Security and confidentiality: Prepared exams need to be securely kept. Any leaked information will definitely compromise the standard of the exam and may result in a cancelation or a retake. All these structures are well addressed using an online system because not only is the content of the exam safely locked away in a database; access to the database is only possible with authorized personnel.
- Time Management: Online evaluation systems make use of computers that helps in saving time. Manual evaluation of descriptive answers requires more effort and more time so automatic evaluation makes it easier.
- Cost Saving: The cost of paper, copying, and distribution expenses are all reduced or eliminated. Even the cost of scanning the papers and then distributing them among the staff for evaluation is eliminated as the students will directly give online tests and they will be checked automatically by the machine.

DISADVANTAGES

- In the proposed system, the dictionary will be provided for identifying the similar words of the given keywords. The words written by the student other than the words mentioned in the dictionary will not be evaluated.

CONCLUSION

The existing online exam systems are mostly objective exams because online evaluation for multiple-choice questions is a very simple task. The system aims on evaluating descriptive answers. From the system, it is clear that descriptive answers too can be evaluated automatically. Which will reduce the work of manual evaluation of the number of answer sheets. Algorithms like TF-IDF, tokenization, stop words removal, stemming, LSA, SOM methods are used in the system using machine learning. The Cosine similarity technique has been effectively used in evaluating student answers as the different students have different answers to the same question. This system will be surely helpful in getting the accuracy for marks allocation as the marks are mailed to the student individually.

REFRENCES

[1] Birpal Singh J. Kapoor and Shubham M. Nagpure and Sushil S. Kolhatkar, "An Analysis of Automated Answer Evaluation System Based on Machine Learning."doi:10.1109/icict48043.2020.9112429 June 2020.

[2] 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 2017.

[3]Meena K and Lawrance R, "Semantic Similarity Based Assessment of Descriptive Type Answer", 2016 IEEE International Conference on Computing Technologies and intelligent Data Engineering, doi:10.1109/icctide.2016.7725366

[4] Vimala Balakrishnan and Ethel Lloyd-Yemoh, "Stemming and Lemmatization: A Comparison of Retrieval Performances," Lecture Notes on Software Engineering vol. 2, no. 3, pp. 262-267, 2014.

[5] Ms. Shweta M. Patil and Prof. Ms. Sonal Patil, "Evaluating the student descriptive answer using Natural Language Processing" International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Vol. 3 Issue 3, March – 2014.

[6] 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.

[7] 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.

[8] 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.

[9] Sonal N. Deshmukh and Ratnadeep R. Deshmukh, "Cosine Similarity for Substituted Text Detection", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 1, ISSN: 2277 128X January 2014.

[10]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.

[11] Y.C.Liu, C.Wu and M.Liu, "Research of fast SOM clustering for text information", Expert Systems with Applications, Volume 38, Issue 8, pp.9325–9333, 2011.

[12]Sungjick Lee and Han-joon Kim, "News Keyword Extraction for Topic Tracking", Fourth International Conference on Networked Computing and Advanced Information Management, 2008.

[13] P.Willett, "The Porter stemming algorithm: then and now", Program, Volume 40, Issue 3, 2006, pp.219-223

[14] Stephen Robertson, "Understanding inverse document frequency: on theoretical arguments for IDF‖, Journal of Documentation", Vol. 60, No.5, pp 503-520, 2004.

[15] Teuvo Kohonen, Samuel Kaski, Member," Self Organization of a Massive Document Collection" Krista Lagus, Jarkko Salojärvi, Jukka Honkela, Vesa Paatero, and Antti Saarela IEEE transactions on neural networks, vol. 11, no. 3, May 2000.

[16] H.Ahonen, O.Heinonen, M.Klemettinen and A.I.Verkamo, "Applying data mining techniques for descriptive phrase extraction in digital document collections"; Proc. IEEE International Forum, Digital Libraries, 1998, pp.2-11.

[17] 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.

[18] K.P.N.V.Satyasree, Dr. J V R Murthy, "clustering based on cosine similarity measure" [ijesat] international Journal of engineering science & advanced technology Volume-2, Issue-3, 508 – 512

[19] Self Organizing Map https://www.cs.hmc.edu/~kpang/nn/som.html

[20] https://www.techopedia.com/definition/13698/tokenizatio