

Automatic Online Subjective Text Evaluation using Text Mining



Shweta Deotare, Rubeena A. Khan

Abstract: *Semantic based text mining is essential in high dimensional data processing in today's environment. In educational sector Question Answer (QA) evaluation has done using semantic as well as semantic analysis in many approaches. Numerous existing mechanisms have implemented using different machine learning algorithms. Semantic and semantic both works for evaluate the text data, but semantic approach should done same work with low time complexity. In this work system carried out automated text evaluation for online examination system with semi structured dataset. The system has categorized into two phases, NLP and Features base evaluation. Natural Language Processing (NLP) has used for preprocessing of data using tokenization, stop word removal, porter stemmer etc. Similarity technique has used for generate similarity score between test answer and train answer data. Artificial Neural Network (ANN) has used to generate the similarity score between two features vectors. Experimental analysis shows the how proposed system is better than some traditional approaches for semantic text evaluation*

Index Terms: *Artificial neural networks (ANN), Semantic similarity text mining, Natural language Processing (NLP), Text mining..*

I. INTRODUCTION

The computer assisted assessment evaluation system was developed to evaluate the one word answer such as of multiple choice questions. It can also evaluate the paragraph answer such as descriptive answer based on the keyword matching. The great drawback of this system is the student cannot know their mistakes and they won't try to improve them. Hence to make them to improve their English knowledge and grammar knowledge the proposed new method called evaluating the student descriptive answer using the Natural Language Processing algorithm and Artificial Neural Networks algorithm will be used. Many Researchers at this field only try to provide the marks and by this method the student will not know their mistakes and again at other exams they will make the same mistakes. It cannot improve the student knowledge on study. Hence this method of Evaluation of student answer using natural language processing and artificial neural networks is used

Revised Manuscript Received on 30 July 2019.

* Correspondence Author

Shweta Deotare*, P.G. Student, Department of Computer Engineering Modern Education Society College of Engineering, Pune (M.H), India

Rubeena A. Khan, Professor, Department of Computer Engineering, Modern Education Society College of Engineering, Pune (M.H), India

© 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/>

mapping, succession, linear sequence matching and quantitative identification and semantic investigation. The motivation behind Text Mining is to process unstructured (textual) data, remove significant numeric files from the content, and, along these lines, make the data contained in the content open to the different information mining (statistical and machine learning) calculations. Data can be extricated to determine outlines for the words contained in the archives or to word documents for the records dependent on the words contained in them. Content mining is utilized to separate vital data or information or example or learning from the test proprietors and applicants answer which are in the unstructured frame. The basic reason behind of of text mining is to find helpful information from natural language text. After the text mining is applied the words like am, is, are, was etc. are like a stop words has eliminates by NLP algorithms and we get keywords from the answer. After identification of the keywords system shows total number of keywords to the exam owner. According to the quantity of catchphrases test administrator can choose checking plan for that specific answer.

The students answer evaluation system has developed in many existing approaches like [5] [6] [9] etc. Each system proposed the own specific evaluation techniques, cosine similarity based techniques are very effective when evaluate system based on text matching, but it sometimes fail when system works with polarity as well as symbolic dataset. Now Teachers can feed the servers with the eligible terms in the categories to be present for student evaluation. The main objectives of this study is to evaluate the student descriptive type answers using the NLP and ANN algorithm and to design a tool for evaluation of the student descriptive type answer using the NLP algorithm for Grammatical checking and produce marks and ANN algorithm for normal answer comparison and produce marks.

II. LITERATURE SURVEY

According to system [1] introduces an Automatic evaluation of an essay (AEE) system which is written in Arabic. The system presents a It is a hybrid approach which incorporates the LSA and Rhetorical Structure Theory (RST) calculation. LSA technique underpins the semantic investigation of the paper, and the RST to assess the composition strategy and the attachment of the article. The LSA technique finds the semantic proportion among two vectors regardless of whether they do exclude comparative words.

The framework forms input paper into two phases is a training phase and testing phase. The training stage is comprised of three sections: identification the normal of words per exposition, identify the most ten obvious words on a given point and applying LSA calculation. The testing stage goes through various procedures: 1) computing LSA distance. 2) ascertaining the quantity of a vernacular. 3) finding various rehashed sentences. 4) ascertaining the length of the essay. 5) ascertaining number of spelling botches. 6) Applying RST algorithm. 7) checking union of exposition identified with the theme. At that point applying two stages, the framework registers the last score dependent on the cosine separation of LSA between the info papers and the preparation articles. The framework reviewed school children essay articles dependent on three criteria which are 40% of the complete score for composing technique, half for the attachment of the exposition and 10% for spelling and grammar mistakes.

Kashi et.al. [2] Proposed a system Score Recommendation System proposed an Automating Assessment in Professional Courses. It illustrates dynamic examination system for expressive English answers that contains numerous sentences. This approaches basically evaluates the students answer with an answer-key which is generated by various students for questions of professional courses. System also depends on a group of algorithms for natural language processing (NLP) which are Wu and Palmer, Longest Common Substring (LCS), LSA.

Alikaniotis D. et. Al. [3] system proposed a automated text evaluation using Artificial Intelligence (AI). System carried out the ANN algorithm for evaluate the two string vectors and generate the runtime similarity score using sigmoid function. LSTM based called Long-Short Term Memory networks to describe the semantic based meaning of texts, System also demonstrate that a fully automated framework is able to achieve excellent results over similar approaches. The experiment evaluation of system generated most successive accuracy rate and encourage by ongoing advances in imagining neural systems, we present a novel technique for recognizing the areas of the content that the model has found more discriminative.

Meena, K. et. Al. [4] proposes an approach of evaluation of online descriptive type student's answers using Hyperspace Analog to Language (HAL) procedure and Self-Organizing Map (SOM) method. To assess understudies reply, the understudy composes the appropriate response and sent as contribution to HAL. HAL develops a high dimensional semantic network from a gathering of a n-word vocabulary. Strategy for build network through inspiration a window of length "1" by the corpus through single word increase. HAL disregards sentence limits, accentuation and changes over each word to numeric vectors communicating data on its implications for words. Inside window registers the separation between two words is "d", at that point processes "(1-d+1)" which means the heaviness of an affiliation among two words.

V. kumaran et. Al. [5] towards a automated framework for short-answer assess utilizing philosophy based framework. Framework comprise the three stages for evaluating the short answers are RDF sentence manufacturer, cosmology

development, and metaphysics mapping. In the main stage, the framework builds the RDF sentence for each sentence in understudy answer and model answer in the wake of perusing the model answer and understudy reply as contribution to plaintext shape. The framework parses each sentence and constructs the linguistic connections for each sentence. It utilizes Stanford composed reliance parser to speak to reliance connections. In the second stage, the RDF sentences are as contribution to metaphysics constructor to build cosmology for them. The creators utilize consecutive and organize connections to develop RDF chart for the RDF sentences.

R. Siddiqi et. Al. [6] proposed a system that provides a novel approach for automated marking of short answer questions. System first evaluate the grade for the desired answer, system also introduce the architecture for the system that is composed of three different modules to evaluate the student's answer. Three modules are given below

1) Checking the spelling and identify the grammatical mistakes using open source tools.

2) Apply NLP for parsing the current text vector like Stanford parser parsing the student's answer using the Stanford Parser. This statistical parser can be creating parses with high accuracy, during the proposed system used various machine learning algorithms like chunked, parser, POS tagger etc.

3) In the final phases system evaluate the trained answer with tagged text data. This process also generates similarity between two desired vectors and predicts the system results.

M. J. A. Aziz et. Al. [7] proposes a design which parses each sentence and builds the grammatical relationships for each sentence. This system basically focus on evaluate the short text according textual as well as grammatical features. The 200 words is the basic limit for given short text. System work like supervised learning using train and text phases.

R. Li, Y. Zhu and Z. Wu et. Al. [8] describes automated evaluation system for descriptive English answers that contains multiple sentences. This system basically used for linguistic parser for Chinese data evaluation. They have made own parser to convert Chinese to English and then evaluate the whole text. The system accuracy was around 85%, which is more efficient than other linguistic parser.

X. Yaowen, L. Zhiping, L. Saidong and T. Guohua et. Al. [9] system proposed ANN base text evaluation method, using feed forward approach. The evaluate given text data according to text similarity algorithm, and generate the similarity distance between two answers. Two quality threshold has used for classification of each answer, they system provide also best accuracy than other existing machine learning algorithms.

Y. Zhenming, Z. Liang, and Z. Guohua et. Al. [10] proposes the online examination system design using machine learning approach. System consist the various machine learning algorithm with semantic analysis approach. It is too much useful to classify the short text data in defined clusters.

III. PROPOSED METHODOLOGY

In the proposed research, design and implement a system for subjective answer analysis approach using semantic analysis. This work also carried out the similarity based supervised learning technique to evaluate long text.

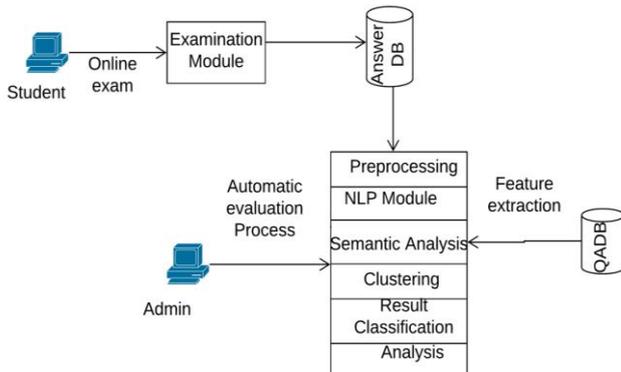


Figure 1: Proposed System Architecture

dataset, and generates runtime weight. Following are the modules system work through:

Examination: In this module students register to the system, login and apply for the examination. System Conducts examination and accepts student’s answers and store it in the Answer DB which is a training dataset for the proposed system.

Preprocessing: In the preprocessing, sentence detection and tokenization applied on answers accepted from students as Answer DB.

NLP: In this module, first of all stop word removal process is applied and all the helping words are removed. Then porter stemming is applied and all the suffix and prefix are removed and root words are extracted from the dataset.

Semantic analysis: In the module, the weightage of the each answer is calculated using ANN approach and answers are evaluated.

Clustering: system has some predefined clusters such as grade a, b, c, d etc.

Result Classification: in the result classification, final result is evaluated and classifies it in the grade.

Analysis: in the analysis module final analysis is made using graph Answer DB and provides system Question Answering is a specialized form of information retrieval. This is a system for automatically optimizing the evaluation of text-based answers and removing issues that have not been addressed. Given a collection of documents, a Question Answering system attempts to retrieve correct answers to questions posed in natural language. Open-domain question answering requires question answering frameworks to probably answer inquiries regarding any possible point. Such frameworks can't, in this way, depend available created space explicit learning to discover and extricate the right answers. System enhancement to provide a platform to educational institutions processes in Assignment Assessment.

IV. SYSTEM ANALYSIS

Dataset Used

The dataset has used for this work as Question Answer (QA) dataset which is generated from our own web source. System

consists the two different modules admin as well as user, when module communicates with system he will upload the train data for question with answer set. Moreover when user login to system he will solve all those questions and those answers as test data for execution of system.

Pseudo code for algorithm

Pseudo code 1: Stop word Removal Approach

```

Initialize the data string S[]
Initialize a=0, k=0
For each (read a to L)
    If (a. equals (L[i]))
        Then Remove S[k]
    End for
    add S to D.
End Procedure
    
```

Pseudo code 2: Stemming Algorithm.

```

Initialize w
Initialize all steps of Porter stemmer
For each (Char ch from w)
    If (ch.count==w.length ()) && (ch.equals (e))
        Remove ch from (w)
    if (ch.endswith(ed))
        Remove 'ed' from (w)
    k=w.length()
    If (k (char) to k-3 .equals (tion))
        Replace w with te.
    end procedure
    
```

Pseudo code 3: Artificial Neural Network

```

Feature of BK rules TrainF [], features if test record TestF []
for all (T in TrainF [] !=null) do
    items [] split(T)
    items1 [] split(TestF)
    w = classifyToAll (Train,TestF[], Label)
Return w;
    
```

Mathematical Model

A System has represented by a 5-different phases, each phase works with own dependency System $S = (Q, \sum, \delta, q_0, F)$. The system will return us weight score between 0 to 1 according to the assign marks.

$Q = QA$ set which is generated by admin.
 $\sum = \{\text{Stopword removal, Tokenized, Porter stemmer, ANN}\}$
 $\Delta = \{\text{input} \rightarrow \text{Hidden} \rightarrow \text{Feedback}\}$ it's a recursive process
 $q_0 = \{\text{student given answer for } \{Q[i]A[i], \dots, Q[n]A[n]\}\}$
 $F = \{\text{clusterset}[1], \text{clusterset}[2], \dots, \text{clusterset}[n]\}$
 $\text{Agg} = \{\text{CS}[i1] + \text{CS}[i2] + \text{CS}[in]\}$
 $\{\text{Weight} \leftarrow \text{Agg}\}$

V. RESULTS AND DISCUSSION

For the system performance evaluation, calculate the matrices for accuracy. The system is executed on java 3-tier architecture framework with INTEL 2.8 GHz i3 processor and 4 GB RAM.

The system carried out the both semantic approach with works with ANN base feed forward approach, its also work like term base neurons verification with hidden layer. This system similar like deep learning approach and it having a ability to work with structures as well as semi structure dataset. The various quality threshold in activation functions will provides drastic supervision filtration which definitely improve the system result than classical classification based machine learning algorithms.

In second experimentation system show the user verification time with different approaches. In current system we consider as four different authorities for runtime verification. The below Fig. 2 shows the performance measures using different parameters with some existing approaches. The below Table 1 shows the proposed system accuracy of various existing machine learning algorithms. The below table also illustrates how proposed system accuracy is higher than existing algorithms.

Table 1: System performance with proposed vs existing algorithms

Approach	Accuracy	Error Rate
Text evaluation[1]	89	2.4
Assessment [2]	87	2.7
ANN [3]	84	2.5
Text Mining [5]	86	2.8
Semantic [10]	91	1.95
Proposed	97	1.7

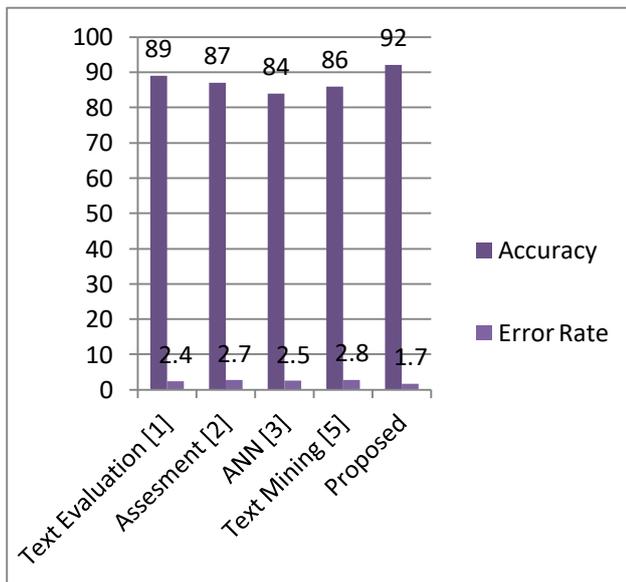


Figure 2: System Performance Measures proposed vs Existing approaches

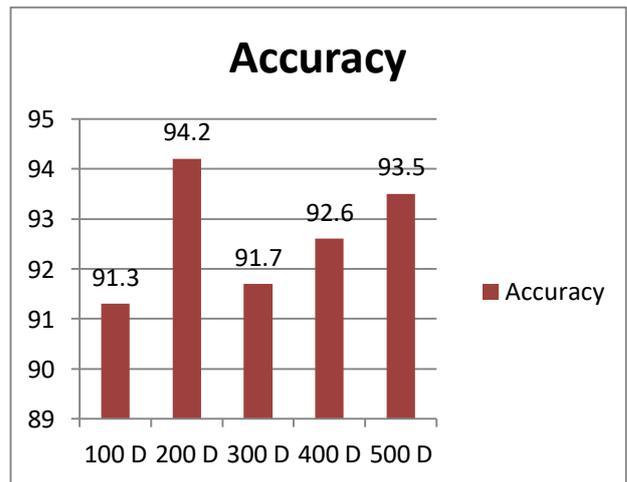


Fig. 4: Data Accuracy with different fold validation (size of documents)

According to above experiment in table 1 as well as figure 2 and 3 illustrates the propose system performance evaluation with entire data set. The system reflects around 92% average accuracy with different experimental setup. Basically the system gives the impact due to initial preprocessing using natural language processing. In the classification phase we used artificial neural network with feed-forward approach.

VI. CONCLUSION

The architecture described and presented in this project will have a high accuracy (up to 97%) with manual human performance. The project drives with the similar facts which a human being thought of while evaluating such as Answer length, keywords presence, and context of keywords. Use of Natural Language Processing coupled with classification techniques, checks for keywords and question specific things as well. Students would have good amount of freedom while writing the answer as the system evaluates for the occurrence of keywords, synonyms, right word context and coverage of all concepts. It can be concluded, using ML techniques it gives satisfactory results due to robust evaluation system. Evaluation accuracy can be raised by feeding it a significantly large and accurate training dataset. Different classifiers can be employed as the technicality of the subject matter changes. By taking feedback from all the stakeholders such as students and teachers can improve the system dramatically as further improvement.

- The system carried out the semantic analysis approach for weight calculation.
- Training and testing phase work during the execution.
- The ANN approach and cosine base weight similarity algorithm was proposed.
- The system overcomes the negation handling as well as false detection issues of semantic analysis approach.

REFERENCES

1. 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.



2. 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.
3. Alikaniotis D, Yannakoudakis H, Rei M. Automatic text scoring using neural networks. arXiv preprint arXiv:1606.04289. 2016 Jun 14.
4. 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.
5. V. Kumaran 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.
6. R. Siddiqi and C. J. Harrison, A systematic approach to the automated marking of shortanswer questions, Proceedings of the 12th IEEE International Multi topic Conference (IEEE INMIC 2008), Karachi, Pakistan, pp. 329-332, 2008.
7. M. J. A. Aziz, F. D. Ahmad, A. A. A. Ghani, and R. Mahmood, "Automated Marking System for Short Answer examination (AMSSAE)," in Industrial Electronics & Applications, 2009. ISIEA 2009. IEEE Symposium on, 2009, pp. 47-51.
8. 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.
9. 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.
10. 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. S3F710.

AUTHORS PROFILE



Shweta Deotare, P.G. Student, Department of Computer Engineering Modern Education Society College of Engg, Pune, MH, India



Rubeena A. Khan, Professor, Department of Computer Engineering, Modern Education Society College of Engg, Pune, MH, India