VOL. 13, NO. 24, DECEMBER 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

NOVICE RETROACTION REPORT

R Shiva Shankar, L V Srinivas, D Ravibabu and Ch Raminaidu
Department of Computer Science and Engineering, Sagi Ramakrishnam Raju Engineering College, Bhimavaram, West Godavari, India
E-Mail: raminaidu.chitri@gmail.com

ABSTRACT

The Online Student Feedback Analysis System is a web based system which collects the feedback from every individual student and provides an automatic generation of a collective feedback which has been used to collect feedback from the students on the main aspects of course such as preparations, contents, delivery methods, punctual, skills, appreciation, and learning experience. We have developed a feedback system to provide the feedback in an easy and quick manner to any particular department in a college or an educational institute. The feedback is collected in terms of qualitative scores. Recent approaches for feedback mining use manual methods and its focus mostly on the qualitative comments. So the evaluation cannot be made through deeper analysis. Student feedback mining system (SFMS) which applies text analytics and sentiment analysis approach to provide instructors a quantified and deeper analysis of the qualitative feedback from students that will improve the students learning experience. We have collected feedback from the students and then text processing is done to clean the data. Feedback comments about each topic are collected and made as a cluster. Classify the comments using sentiment classifier and apply the visualization techniques to represent the views of students. This proposed system is an efficient approach for providing qualitative feedback for the instructor that enriches the students learning. With the help of this application, we can give feedback through the online system much faster than the existing paper feedback system. The existing system takes more time to get the feedback from the students, thus the online feedback system is implemented. Students will fill online feedback using a standard form provided online. Special care has been taken to provide the security in the proposed system as only the authentic user is able to see the collective feedback of a batch of the students and can also get to know about the collective opinion. The application of giving the feedback is not only objectively (i.e. rating out of a fixed constant value) but also in subjective manner by leaving their comments and reviews about any particular field or subject. The main objectives of feedback analysis system are the conventional objective analysis as well as Subjective analysis with the help of Sentiment Analysis.

Keywords: qualitative scores, feedback mining system, text analytics, sentiment analysis.

1. INTRODUCTION

Students provide feedback in qualitative comments related to preparation, contents, delivery methods, punctual, skills, appreciation, and learning experience. The delivery methods and preparation component refers to instructor's interaction, delivery style, ability to motivate students, out of class support, etc. The content refers to course details such as concepts, lecture notes, labs, exams, projects, etc. The preparation refers to student's learning experience such as understanding concepts, developing skills, applying acquired skills, etc. The paper correction refers to correction of mistakes and providing solutions to overcome it. The punctual refers to the class timing and assignment or record submission [8]. The appreciation refers to the comments given when something is done perfectly. Analyzing and evaluating this qualitative data helps us to make better sense of student feedback on instruction and curriculum.

This system provides facilities for selecting any particular subject for feedback and generates a report automatically; build collective opinion of the students, student's needs and requirements in the college. The Online Student Feedback Analysis System (OSFAS) is an automatic feedback generation system that provides the proper feedback about the teachers by using comments and categories like good, interesting, late, interactive, etc.

Recent methods for analyzing student course evaluations are manual and it mainly focuses on the quantitative feedback. It does not support for deeper analysis. This paper focuses on providing qualitative feedback to analyze and provide better teaching to improve the student's performance [7].

Questionnaires are of huge importance in the dialogue with the students, since these questionnaires are the best tool currently available for collecting objective, detailed, and reasonably systematic information on a wide range of questions, which informs the teacher about student's perception about the course - its strengths and weaknesses.

Responses are gathered and analyzed on behalf of the department by this system and will be used for the purpose of the quality enhancement. The aim of this system is to save time and also to decrease human load and efforts.

In order to gain maximum advantage:

a) Students must be instructed well on how to fill and submit the feedback online, when and how the results will be published, and that their contribution and opinion is important and taken seriously.

b) The head of the department must discuss a complete summary of the feedback on each course with the concerned faculty.

c) The comment section for the students in the online feedback form should be filled with direct statements by the students and must not contain any ironical statements or phrases.

2. LITERATURE SURVEY

Jill Burstein [1] has worked on "Automated Evaluation of Essays and Short Answers" using E-rater software which is first used to analyze the essays submitted in GMAT exam in 1999. This typically use: organization, sentence structure, and content. The conclusion derived through this is, E-rater scores are comparable to human reader scores, and automated scoring procedures can reduce the time and costs involved with manual essay scoring. Automated essay scoring would appear to be a favorable solution toward the introduction of more writing assessments on high-stakes standardized tests, and in a lower stakes environment.

Daniel Marcu [6] researched towards "Automatic Classification of Discourse Elements in Essays" about writing features that can facilitate the essay revision process. Using a relatively small corpus of manually annotated data, Bayesian classification is used to identify thesis statements. This method yields results that are much closer to human performance than the results produced by two baseline systems. The results obtained are comparably favourable with results reported by Teufel and Moens (1999) who also use Bayes classification techniques to identify rhetorical arguments with good accuracy such as aim and background in scientific texts, although the texts we are working with are extremely noisy.

Karen Kukich [2] has made his work on "Automated Scoring Using a Hybrid Feature Identification Technique" which is E-rater but shows the advanced feature of it, including syntactic structure analysis, rhetorical structure analysis, and topical analysis, to score essay responses from test-takers of the Test of Written English (TWE) and so on. Score prediction for cross-validation sets is calculated from the set of predictive features. Exact or adjacent agreement between the Electronic Essay Rater (e-rater) score predictions and human rater scores ranged from 87% to 94% across the 15 test questions. Clauses and sentences annotated by APA as "the beginning of a new argument" might be used to identify main points of an essay (Marcu (1997)).

Yuji Nakamura [3], a Tokyo researcher has contributed on "A comparison of holistic and analytic scoring methods in the assessment of writing" examines the strengths and weaknesses of holistic and analytic scoring methods, using the Weigle adaptation of Bachman and Palmer's framework, which has six original categories of test usefulness, and explores how we can use holistic or analytic scales to better assess student compositions. For practical and economical reasons, holistic (one item evaluation) assessment can be used, but to avoid risky idiosyncratic ratings, analytic assessment (with several evaluation items) is suitable. The best practice is to have multiple raters and multiple rating items. The next best practice is to have one overall evaluation item and multiple raters. The third choice would be to have one rater and multiple items. The least recommended solution would be to have one rater and one item. Even worse than this, however, would be to have one rater and an impressionistic scale.

Hunter M. Breland [4] has worked on "The Direct Assessment of Writing Skill: A Measurement Review": Reliability is examined as it is influenced by reader inconsistency, domain sampling, and other sources of error. Validity evidence is presented, which shows reported relationships between direct assessment scores and criteria such as class rank, English course grades, and instructors' ratings of writing ability. New automated methods of textual analysis and new kinds of direct assessment in which more than a single score is produced are suggested as two approaches to better direct assessment. The test is printed with the entire essay on one page of the test booklet. A comparative validity examination of task types would be valuable.

Francis F. Balahadia; Ma. Corazon G. Fernando; Irish C. Juanatas, has developed "Teacher's performance evaluation tool using opinion mining with sentiment analysis". They collected the feedback from the students and identified the strength and weakness of the particular teacher. They evaluated the qualitative and quantitative data and provided sentiment score of the teacher in a school.

3. PROPOSED SYSTEM

Background

Text mining approach is useful in the sentiment analysis process. In this section, we provide a brief description about the methods that we adopted to extract the keywords from the students feedback document.

3.1 Tokenization

Tokenization is the act of breaking up a sequence 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, some characters like punctuation marks are discarded.

3.2 Lowercasing

Lowercasing is the process of turning the upper case letters to lowercase. The words may not differ from uppercase to lowercase.

3.3 Numerical feature

Numerical feature evaluates the number of sentences written by student, no of words in each sentence and average word length used in the sentence. These features help in evaluating the efficiency of feedback as valid words are only considered.

3.4 Stemming

Stop words are words which are filtered out before or after processing of natural language data. These words are removed to extract only the meaningful

VOL. 13, NO. 24, DECEMBER 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

information. The list of stop words may be ‘the, is, at, which, on, who, where, how, hi, before, after’ etc.

3.5 Clustering

Clustering is the process of making a group of abstract objects into classes of similar objects. A cluster of data objects can be treated as one group. While doing cluster analysis, we first partition the set of data into groups based on data similarity and then assign the labels to the groups. The main advantage of clustering over classification is that, it is adaptable to changes and helps single out useful features that distinguish different groups.

3.6 Classification

Data classification is the process of organizing data into categories for its most effective and efficient use. A well-planned data classification system makes essential data easy to find and retrieve. This can be of particular importance for risk management, legal discovery, and compliance.

4. SYSTEM ARCHITECTURE

In the proposed a system Figure-1 is developed to mine the feedback given by the students and obtain knowledge from that and present that information in qualitative way. Feedback was collected for a course; those feedbacks were pre-processed using text processing techniques. In pre processing, the feedback files are generated as a flat file. The flat file is tokenized into sentences and the keywords are listed after removing the stop words. We have identified the frequency of each word and extract the topic which has the highest frequency count. Similar comments in each topic are clustered and then the clustered words are classified into positive or negative comments. The classified comments are generated as a chart for easy visualization.

Token matching

Bag of words are pre defined words set for feedback evaluation. After the process of stemming tokens were formed for the opinion given by the students. The tokens were compared with the bag of words resulting in token matching process. If true, the feedback is evaluated otherwise the feedback is to be rewritten.

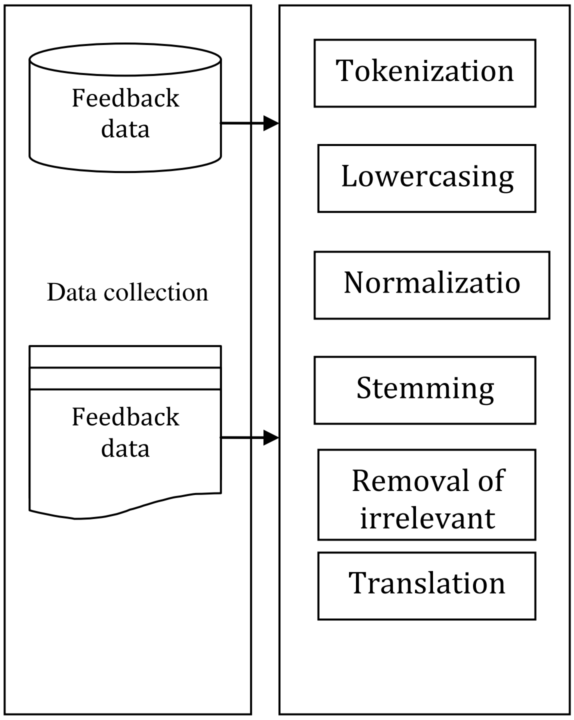


Figure-1. System architecture.

Sentiment analysis

Sentiment analysis refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials. Sentiment analysis is widely applied in review and social media for a variety of applications, ranging from marketing to customer service. Sentiment analysis aims in determining the attitude of a speaker or a writer with respect to some topic or the overall contextual polarity of a document. The attitude may be his or her judgment or evaluation, affective state, or the intended emotional communication is shown in Figure-2.

ARPN Journal of Engineering and Applied Sciences

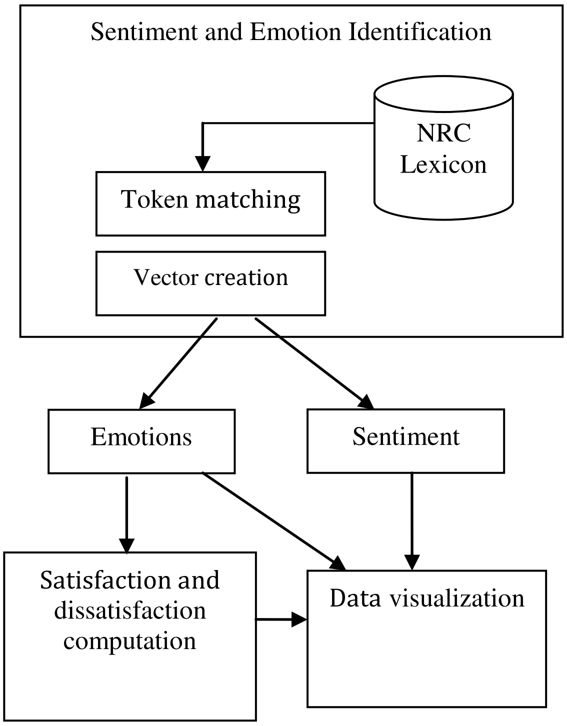


Figure-2. Sentiment and emotion identification.

5. IMPLEMENTATION

Feedback and growth analyser takes two parts of working: Student, Admin. The student page helps the students to register feedback using their name. Admin works for posting the type of feedback, mining data, producing score.

Initially the admin registers with a username and password. If already registered then he gets logged in to his account and proceeds to add topic phase which help to add the type of feedback required from a specific class. The topics and the feedback give to the faculty by students is stored in a data base. The admin applies sentimental analysis which utilises natural language processing to the saved feedback. The NLP technique includes breaking the paragraph in to sentences, removing stop words, speech evaluation, spell checker, bag of words which takes a part in evaluating the feedback. The scores generated are stored in a database. The evaluated scores help to analyse the data by plotting the graphs.

The algorithm for keyword extraction is as follows:

To clean the data, the collected feedback is subjected to tokenization and stop word removal. The following sequence of steps shows that how do we perform pre processing.

Input: feedback collection
Output: topics
Step 1: Read each student feedback document and append it into a document.
Step 2: Tokenize the document based on (separators). Or, to identify the sentence.
Step 3: On each sentence, remove the stop words.
Step 4: Update the document.
Step 5: Now the document is removed from stop words.

Topic extraction

From the pre processed document, the parts of a sentence like adjectives, verbs, adverbs, pronouns, nouns, proper noun etc., is removed to identify the topics available in the student feedback. The topic might be teaching, project, communication, interaction, punctuality etc.

We have used a threshold δ to limit the number of topics. The frequency of each word in D is counted by using the equation 1. If a word exceeds the δ then it is identified as topic.

$$\sum \quad \mu(\mathrm{w}) = \frac{N(W)}{T}$$

Where μ(w) is the frequency of the word w in document.
N(w) is the number of times the word w appeared in document.
T is the total number of words in document.

Algorithm for sentimental analysis

Input: Feedback given by the students related to a topic
Output: Score for the selected feedback comment based on different features for each student

ALGORITHM

Step-1: Admin will post the feedback topics for each class and add bag of words for each topic.
Step-2: Student then selects the class and writes the feedback related to that topic.
Step-3: The words in the feedback are then compared with the bag of words to get similarity ratio. It is calculated as follows:

Ratio = (words (feedback) ∩ words (bag)/words (feedback) ∪ words (bag)) *100

Step-4: If, ratio ≥ 20 then jump to step 5, else go to Step 12.
Step-5: The sentences are searched for „." to split and sentences are counted.
Step-6: Stop words are removed by comparing the each word with stop word list and replaced it with an empty space and continued further to compare with other words.
Step-7: The weight of the essay is calculated by incrementing the word occurrences which help in gaining the word count.
Step-8: The score for feedback is calculated using the ratio, no of sentences frequency of word occurrences, average word length.

Average word length = Length of all words ÷ Number of words

VOL. 13, NO. 24, DECEMBER 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

((sentences > 10) && (wordcount >= 60) && (wlength > 7.0D))
Score=8.0D

Step-9: To validate a sentence each word is tagged with its parts of speech using a POS tagger. Now get the link between one word to other. This sentence forms are now compared with certain rules. If they satisfy the rules they are considered to be valid sentences, else, not valid. The probability is calculated as:

prob = correctlines / maxlines * 100.0D
if (prob >= 75.0D)
prob=8.0D

Step-10: To check spellings of all the words we use an en-US dictionary. Each word from essay after stop word removal is compared with the words in dictwords.txt. The counts of misspelled words are used to find the probability.

prob = mis spellwords / total words * 100.0D;
if (prob >= 75.0D) {
score += 1.0D;

Step-11: A final score is given based on scores from step-6 to step-10.
Step-12: End.

Algorithm explanation

The novice retroaction system takes the feedback given by students as input and generates the score for each feedback as an output. The system runs on an algorithm which follows certain steps to generate the score. Initially the admin will post the topics on which the students need to give the feedback along with some bag of words. When a student gives a feedback, the words in the comment are compared with bag of words and produce a ratio. In which the comment is accepted only if the ratio is greater than 20. After the acceptance of the feedback sentences are separated by using". words by stop word and are counted , word length is calculated. The sentence count, parts of speech and spelling check are used to generate the score for each feedback.

6. EXPERIMENTAL ANALYSIS

The score for all the students from different departments are collected and stored in the database. A sample is considered for analysis of the system performance and graphs are generated for both system analysis and manual evaluation. The graphs help in evaluating the efficiency of the system.

The below information is system generated scores for three different departments:

Table-1. CSE department scores.

CSE	Punctuality	Assignment	Teaching
C01	37	27	11
C02	49	40	41
C03	26	35	8
C04	20	28	25

Table-2. Mechanical department scores.

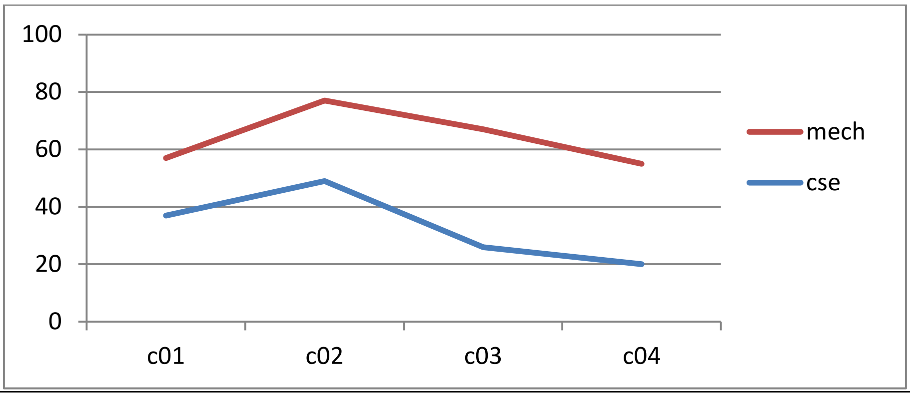
Mech	punctuality	Assignment	Teaching
M01	20	46	44
M02	28	15	25
M03	41	37	19
M04	35	23	40

Table-3. ECE department scores.

Ece	Punctuality	Assignment	Teaching
E01	21	36	24
E02	36	45	8
E03	38	24	31
E04	12	46	33

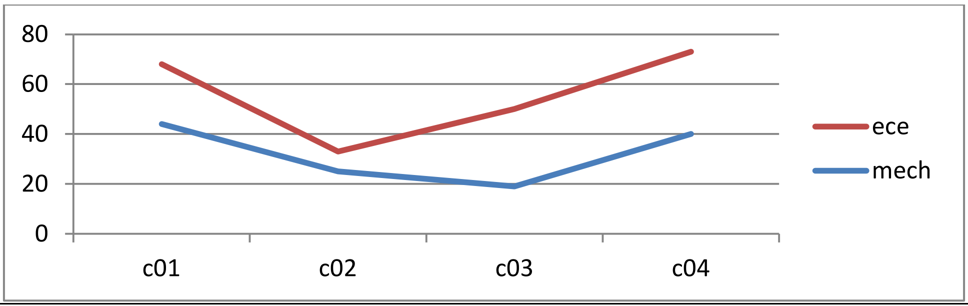
The comparison graphs are considered in order to compare the different departments in different aspects.
The graphs are shown below:

Punctuality



Graph-g1. Punctuality level.

Assignment



Graph-g2. Assignments.

The comparison of the system evaluated scores with manual evaluated scores gives the system efficiency. For comparison of all the departments the average score is considered and graphs are generated for both manual and system scores

Table-4. System generated average scores.

Average Score	CSE	MECH	ECE
S1	33	31	26.75
S2	32.5	30.25	37.75
S3	21.25	32	24

System Scores:

Table-5. Manual evaluated average scores.

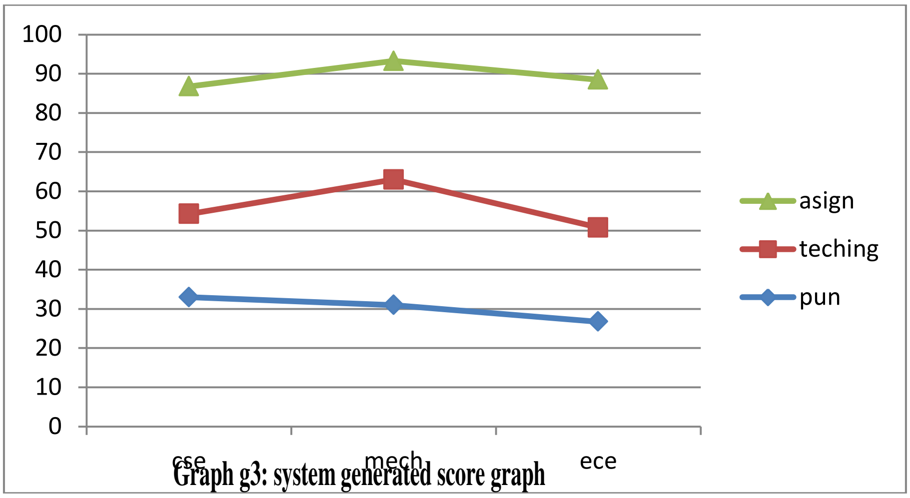
Average Score	CSE	MECH	ECE
S1	33.25	32.5	28
S2	31	29.75	36.25
S3	22.75	31.25	25.75

Manual scores:

Graph for system generated scores

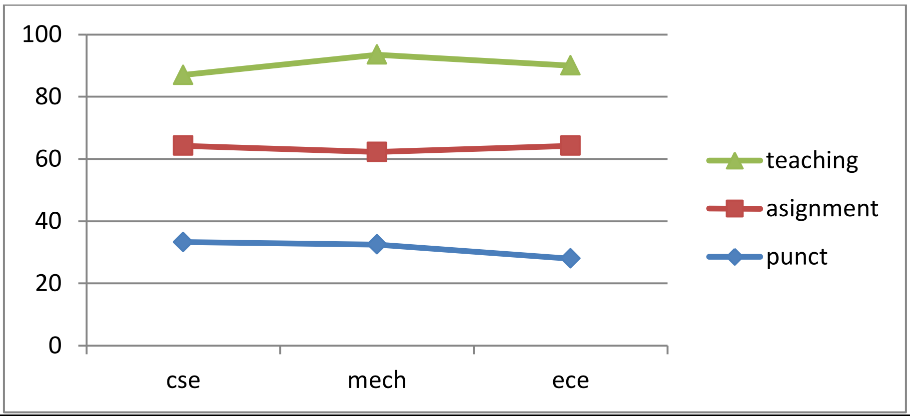
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com



Graph g3: system generated score graph

Graph-g3. For manual scores:



Graph-g4. Manual evaluation scores graph.

The graphs indicate the system performance. The error percentage is very low so the system is efficient for usage and can be used for evaluating the feedback as it may not affect the student and faculty in a negative way.

7. CONCLUSIONS

The proposed system is designed to reduce the time and save the efforts of the faculties from maintaining huge amount of records. At the time of the feedback generation it applies the proposed algorithms and methodologies for generating a feedback of a particular batch of students. After this a whole collective report is generated with pie-charts and graphs showing the collective rating of a particular field and a collective opinion is provided on the basis of sentiment analysis done on the subjective review/comment provided by the students about any particular field such as faculties, course structure, subject's topics etc. which can be provided to all the officials such as the principal, HOD etc. When compared with the existing feedback system, the proposed faculty feedback and growth analyser is much easier to implement and manage.

In comparison of the existing system, faculty feedback and growth analyser very easily save each and every record of the individual students in the database and avoid usage of paper and save time and effort of the workers in charge or faculties.

REFERENCES

[1] Jill Burstein, Claudia Leacock, Richard Swartz. 2001. Automated evaluation of Essays and short answers. Published by ETS Technologies, Inc. A Subsidiary of Educational Testing Service, Princeton.

[2] Burstein, J., Kukich, K. Wolff, S. Lu, C. Chodorow, M, Braden-Harder L. and Harris M.D. 1998. Automated Scoring Using A Hybrid Feature Identification Technique. Proceedings of ACL, 206-210.

[3] Yuji Nakamura. 2004. A comparison of holistic and analytic scoring methods in the assessment of writing. The Interface between Interlanguage, Pragmatics and Assessment: Proceedings of the 3rd Annual JALT Pan-SIG Conference. May 22-23, 2004. Tokyo, Japan: Tokyo Keizai University.

[4] Hunter M. Breland. 1983. The Direct Assessment of Writing Skill: A Measurement Review. proceedings of College Board Report No. 83-6 ETS RR No. 83-32

[5] Francis F. Balahadia; Ma. Corazon G. Fernando; Irish C. Juanatas. 2016. Teacher's performance evaluation tool using opinion mining with sentiment analysis.

[6] A Burstein J., Marcu D., Andreyev S. and Chodorow M. (submitted). Towards Automatic Classification of Discourse Elements in Essays. To appear in Proceedings of the Annual Meeting of the Association for Computational Linguistics, Toulouse, France, July.

[7] Duyu Tang. 2004. A joint segmentation and classification in conference on empirical analysis in conference on empirical methods in natural language processing. 5: 477-487.

[8] R. Shiva Shankar, D Ravibabu, KVSSR Murthy, VMNSSVKR Gupta. 2017. An approach for essay evaluation using system tools. Proceedings by IEEE, Innovative Research in Electrical Sciences (IICIRES), 2017 International Conference on 16-17 June 2017.