OPINION ASPECTS IN COGNIZING STUDENT FEELINGS VIA REVIEWS

[1]Ms. Ramya R Pai, [2]Ms. Shalini S, [3]Ms. Sirina V.A.P, [4]Ms. Sushma N,[5]Mr. Sanjoy Das,
Fourth Year B.E, Fourth Year B.E, Fourth Year B.E, Fourth Year B.E, [5] Asst. Prof
Department of Computer Science & Engineering
K. S. Institute of Technology, Bengaluru-62, India

Abstract: The Main aim here is finding the opinion of the lecturer on the teaching based on their reviews. Reviews can help them to improve the quality of teaching. However, it is a challenging issue to acquire the valuable reviews by recognizing the review sentiment.This project will dynamically identify the student opinion about the particular lecturer whether the student will have a positive or a negative opinion about the lecturer. Reviews have become an important source of information for organization to know the student opinion about lecturers.Identifying reviews is helpful to monitor and manage early promotion and also reviewers. The final graph is generated based on the positive and negative opinion of the students.

***Keywords*- Data pre-processing, Term frequency construction, classification**

I. INTRODUCTION

A system to extract lecturer aspects and corresponding opinions from lecturer reviews. Determining a consensus opinion on a lecturer is no longer easy, because of different reviews expressed by each students. To address this problem, researchers have used various approaches, such as looking for feelings expressed in the documents and exploring the appearance and syntax of reviews.

Aspect-based evaluation is the most important aspect of opinion mining, and researchers are becoming more interested in lecturer aspect extraction; however, more complex algorithms are needed to address this issue precisely with large data sets. It introduces a method to extract and summarize lecturer aspects and corresponding opinions from a large number of lecturer reviews in a specific domain.

Sentiment classification, also known as document-level sentiment analysis, is the most broadly researched topic. It classifies a review as conveying a positive or negative feeling. In this task, the whole document is considered as the elemental information unit. Sentiment analysis can be categorized into three subtasks: sentiment classification, subjective/objective identification, and aspect-based sentiment analysis. Sentiment classification, also known as document-level sentiment analysis, is the most broadly researched topic. It classifies a review as conveying a positive or negative feeling. In this task, the whole document is considered as the elemental information unit.

II. RELATED WORK

Previously they considered two components: binary classifiers trained using single layer feed forward network for aspect category classification (Slot 1), and sequential labelling classifiers for opinion target extraction (Slot 2). Besides extracting a variety of lexicon features, syntactic features, and cluster features, we explore the use of deep learning systems to provide additional neural net-work features. The goal is to correctly identify the aspects of entities and the polarity expressed for each aspect.

Due to huge, unstructured and scattered amount of data available on web, it is very tough for users to get relevant information in less time. Later achieved this by improvement in design of website, personalization of contents, prefetching and caching activities are done according to behaviour analysis.

III. BACKGROUND PROBLEM

Classifying students by reviews is more challenging than in the case of long documents associated with them as it is difficult to track students' reviews in streaming sparse data. The existing causing the ineffectiveness for sentiment classification. Loss of information during the classification, This lead to propose this model.

IV. METHODOLOGY

FLOW OF THE PROPOSED SYSTEM

The flow of the proposed system is as shown below. The proposed system is introduced to overcome all the disadvantages that arises in the existing system. It will effectively do sentiment analysis on the reviews that are helpful for the organization to easily identify opinion about the particular lecturer.

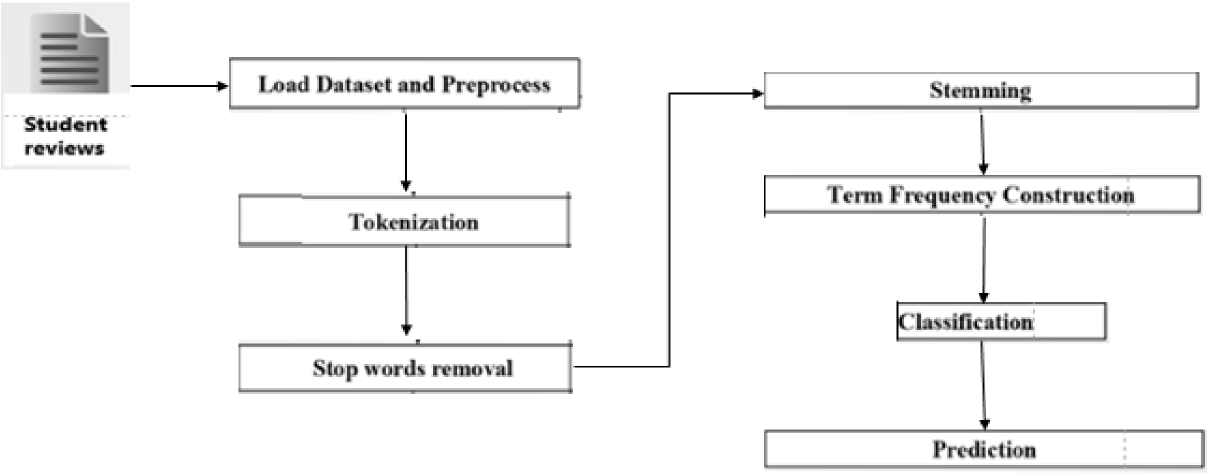

Figure 1: Flow of the proposed system

MODULE DESCRIPTION

1. **Data selection and loading**

Data selection is the process of selecting the appropriate data set for processing.The selected dataset is going to be used for identifying the student opinion about the lecturer.The dynamically distributed reviews are detected from the dataset.

2. **Data preprocessing**

The data is preprocessed to remove reviews from anonymous students, since we would like to associate each review with a unique student. We then remove duplicate reviews often caused by multiple versions.

3. **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.

4. **Stop word removal**

Stop words are natural language words which have very little meaning, such as "and", "the", "a", "an", and similarwords. The stop words are detected from the reviews and it’s removed.

5. **Stemming**

Stemming is the process of converting the words of a sentence to its non-changing portions.It works by identifying word suffixes and stripping them off, with some regularization of the endings.

6. **Term frequency construction**

Term frequency (TF) is used in connection with information retrieval and shows how frequently an expression (term, word) occurs in a document. Term frequency indicates the significance of a particular term within the overall document.

7. **Classification**

Classification is used to classify each item in a set of data into one of predefined set of classes or groups. Classification is a data mining function that assigns items in a collection to target categories or classes. The goal of classification is to accurately predict the target class for each case in the data.

8. **Result Generation**

The overall clustering report is generated based on the reviews in given by students. The level that describe the positive and negative opinion of the customer.

V. EXPERIMENTAL RESULTS

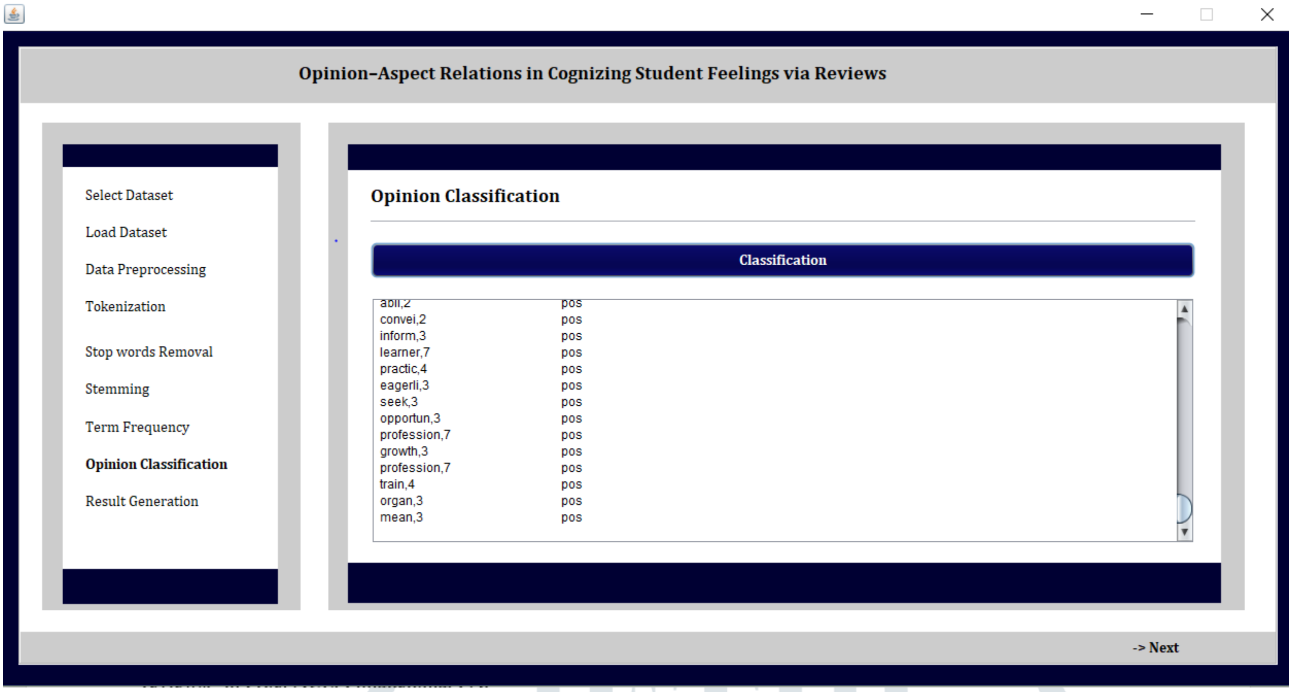

Figure 2: Classification

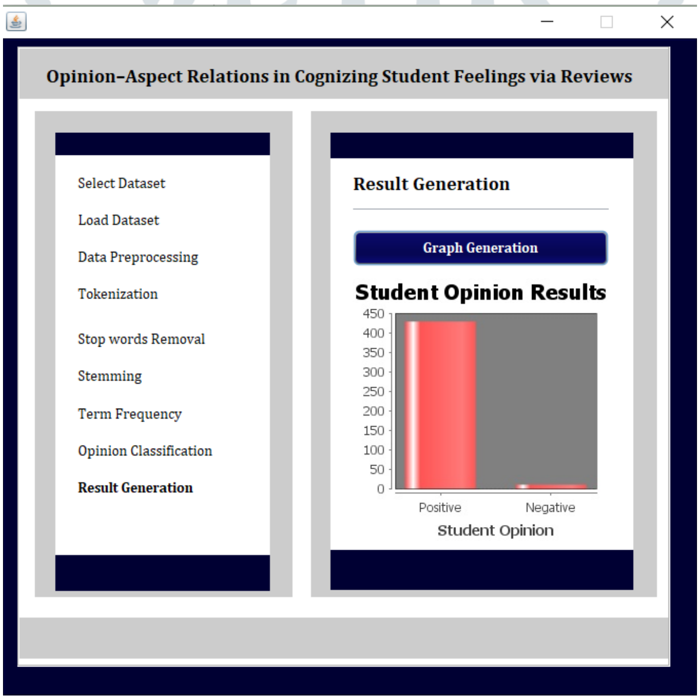

Figure 3: Graph generation

VI. CONCLUSION

This project enhances the performance of the overall classification and prediction.It finds the different groups of words and it summarizes effectively and quickly.It alleviate the sparsity problem and reduces the information loss.The accuracy of the result generation is highly increased.

VII. FUTURE SCOPE

We will investigate this task with real student portal cases in collaboration which helps to distinguish the opinion sentiments in the future. In future we can store all the data in a Hadoop storage for increasing the processing speed.

VIII. ACKNOWLEDGEMENT

The authors thank their management and institution for providing them the resources and platform for showcasing their idea. They would also like to thank their HOD and all the lecturers and professors for motivating them into framing this research.

IX. REFERENCES

[1] Dave, K., Lawrence, S., & Pennock, D. M. (2003, May). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In Proceedings of the 12th international conference on World Wide Web (pp. 519-528). ACM.
[2] Whitelaw, C., Garg, N., & Argamon, S. (2005, October). Using appraisal groups for sentiment analysis. In Proceedings of the 14th ACM international conference on Information and knowledge management (pp. 625-631). ACM.
[3] Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. Management science, 53(9), 1375-1388.
[4] Devitt, A., & Ahmad, K. (2007). Sentiment polarity identification in financial news: A cohesion-based approach. In Proceedings of the 45th annual meeting of the association of computational linguistics (pp. 984-991).
[5] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.
[6] Hatzivassiloglou, V., & Wiebe, J. M. (2000, July). Effects of adjective orientation and gradability on sentence subjectivity. In Proceedings of the 18th conference on Computational linguistics-Volume 1 (pp. 299-305). Association for Computational Linguistics.
[7] Wiebe, J. (2000). Learning subjective adjectives from corpora. Aaai/iaai, 20(0), 0.
[8] Wiebe, J., Wilson, T., & Bell, M. (2001, July). Identifying collocations for recognizing opinions. In Proceedings of the ACL-01 Workshop on Collocation: Computational Extraction, Analysis, and Exploitation (pp. 24-31).
[9] Turney, P. D. (2002, July). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In Proceedings of the 40th annual meeting on association for computational linguistics (pp. 417-424). Association for Computational Linguistics.