# A Methodological Framework to Identify the Students' Opinion using Aspect based Sentiment Analysis

<sup>1</sup>A. Jenifer Jothi Mary, <sup>2</sup>S. Santiago <sup>3</sup>Dr. L. Arockiam <sup>1</sup>Research Scholar, <sup>2</sup>Research Scholar, <sup>3</sup>Associate Professor, <sup>1,2,3</sup>Department of Computer Science, St. Joseph's College, Tiruchirappalli, Tamil Nadu, India.

Abstract- Now-a-days, students write feedback of the courses in every semester and discuss with their friends in chat rooms, mails and social media sites. This type of discussions on social media which is called as opinions, may be in the form of feedback. remarks, reviews, comments, clarifications, observations and explanations irrespective of their course. It is evident that these opinions may help the teaching community to identify students' feelings and attitudes towards teachinglearning. The feedback taken during the course will help to address the issues of teaching-learning rather than taking at the end of every semester or a course. So, the opinions collected from the social sites, need to be analyzed using any of the sentiment analysis techniques for better understanding of the students. Analyzing such information improves monitoring of students' performance and their learning outcome.To implement this idealogy, the reseachers propose a methodological framework for extracting the insights from students' opinion, that is posted on social media websites.

#### Keywords— Students' Opinion; Aspect Based Sentiment Analysis; Methodological Framework; Teaching-Learning; Social Media Sites.

#### I. INTRODUCTION

The significant advancement of the last two decades is the swift growth of information technology in many areas such as education, economy, health, agriculture, social life and entertainment. Today, Internet is one of the vital learning environment that delivers a new and rich style of learning. The technological penetration lives in both formal and informal education. It has twisted a need to study the numerous characteristics of this new way of learning and exploring with different learners' needs. Today, everybody is interested in knowing others' opinions, feelings. The opinion of the other people may be positive, negative or neutral. Students normally inhibited to express their opinions to the teachers directly regarding teaching methods and their level of understanding during class hours. They feel free to share their observations and views in social media sites such as Twitter, YouTube, Facebook, Forums or blogs etc., on any subjects or teachers with their friends. So, we can easily collect fair comments and feedback from social media sites. But, accumulation and examining of this kind of social media data is a challenging task.

## II. ANALYTICS IN EDUCATION

Today's society is driven by data, as evidenced by the popular use of the term Analytics. Education Analytics will help academicians to improve teaching and learning, cut costs and improve efficiencies by enhancing student achievement. Enhancing teaching and learning through educational data analytics is a value-added assessment and helpful to a teacher to improve her / his students' academic gains. The processes of data assessment and analysis will enable the academia to measure, improve and compare the performance of individuals, departments and institutions. This Academic analytics will help to predict students' academic performance, and advisors to customize learning paths. These mining data will support in teaching and learning customization, tutoring, or intervention within the learning environment. The proposed study uses the data and models to predict student progress and performance, and the ability to act on that information. This analysis is associated with student learning to observe and understand learning behaviors in order to enable appropriate intervention.

## III. SENTIMENT ANALYSIS

In this scenario, Sentiment Analysis (SA) or Opinion Mining (OM) is the computational study of opinions, emotions and sentiments that are shared by others in social networks [11]. Social networking sites, Online Reviews and Open Forums generate large amount of data that are shared by all who have the same idea or view. Today, education system has to be tuned to know the opinion/attitude/emotion of students to identify their mind-set towards the teachinglearning process adopted by the teachers. Therefore, the proposed conceptual data analytic framework aims to position Big Data Analytics within the academic domain. This determines the polarity of the students' opinion. The polarity measurement will be performed in three levels namely, Document Level (DL), Sentence Level (SL) or Entity Level (EL)/Aspect Level (AL) [13]. An Entity/Aspect Level provides a better result to analyze the comments, views and feedback that are posted in the social networking sites with DL and SL. Analytical tools like Hadoop, Map Reduce, Pig, Hive and machine learning algorithm such as Naïve Bayes, Support Vector Machines, Maximum Entropy and Lexicon Based approaches are used to determine the polarity [12].

## IV. REVIEW OF LITERATURE

Nabeela Altrabsheh et al. [1] introduced an architecture for Sentiment Analysis in Education (SA-E). Different methods were used to collect the feedback which helped to identify the student's attitudes/opinions towards the teachers. They compared with previous feedback systems which provided the limited amount of text as feedback such as clickers, SMS and Mobile Phones.

Miaomiao Wen et al. [2] proposed a methodology to find out the student drop-outs in Massive Open Online Course (MOOCs). The students' opinions were collected from open forums, blogs and posts in the social networks in a particular week. Through this, they extracted the positive and negative feedbacks about a topic/ course. With these, the MOOCs were helpful improve the lectures, lecturers, courses and its tools. This was done at course level and user-level.

Pallavi K. Pagare [3] developed a workflow for both qualitative inquiry and large-scale data mining. It focused on engineering students' Twitter posts to know the problems and difficulties in learning such as heavy learning load, lack of social meeting and sleep deficiency in educational practices. This study presented an insight on student's incident by applying multi-label classification and decision tree algorithm. To get accurate result, this was compared with other generic algorithms.

N. D. Valakunde et al. [4] presented a concept paper on multi-aspect and multi-class document classification for various levels of SA. They have verified faculty performance on aspect based sentiment scores, like, knowledge, presentation, communication and regularity. These aspects were taken from the NAAC accreditation criteria. They have experimented with the dataset of students' feedback foreight years, which were not about the particular students or subject and not in real situation.

M Ravichandran et al. [5] developed the Twitter Sentiment Mining (TSM) framework based on e-learner's emotional state, in order to find the frequently interacting learners in e-learning environment. E-learning propagates the knowledge with the help of computers in a web based classroom environment. It is too hard to identify the learner's emotional state in the online courses. Using Naïve Bayesian approach, the model has been constructed for the training data of e-learners. This was compared with the standard classification algorithms too.

Sheetal R. Patodiya et al. [6] proposed a multi-label Naïve Bayes algorithm to classify the various aspects by analyzing the students' opinion from social media. This has multiple languages and images to improve the probability of sentences of other aspects. These aspects confer the positive way of taking education and the benefits of education.

Geetika et al. [7] proposed a technique of machine learning with semantic analysis for classifying the sentences on product reviews of twitter data. They proposed a framework and experimented it. This increased the result from 88.2% to 89.9% with the improved Naïve bayes and semantic analysis WordNet. Nabeela Altrabshehet al. [8] considered learning capabilities of machine learning algorithms such as Naïve Bayes, Complement Naïve Bayes, Maximum Entropy and Support Vector Machines. This was used to experiment the students' textual feedback in a unigram feature selection method. The tests were about the learning performance of the four models and 10-fold cross-validation. The SVM (Support Vector Machine) produced highest accuracy as compared with other models.

P. Bharathisindhu et al. [9] proposed a system that determined the sentiment of the E-courses through e-learning portals. With an automated sentiment analysis, all the sentiments were classified using Bayes rule. The sentiments classified to identify frequently accessed pages, current trends and remove the untapped pages to enhance the Ecourses. It mainly helped the developers to modify the courses accordingly.

Zied Kechaou et al. [10] proposed and applied the combined HMM (Hidden Markov Model) and SVM to classify the users' opinion to extract the useful information from the users' feedback which is posted in the social networks. They applied only 20% of the data as testing data and the remaining as training data. They did not discuss preprocessing to remove the noisy text.

#### V. OBJECTIVES

After reviewing all the above literature, this research work is framed to identify the student's mind-set towards the teachers and their teaching. Using this system, the analysts can recommend modifications in course content and methodology of teaching.

### VI. METHODOLOGY

Methodological way of presenting this framework will give a clear view. Sentiment Analysis is helpful in extracting knowledge from this unstructured students' feedback for the improvement of teaching and learning process. It makes it easy to know the opinions of the learners through an automated system. The aim of this framework is to identify the student's mind-set towards the teachers and their teaching. By this system, the analyst can be used,

- To know the students' views, feelings, comments, mind-setup towards their course and subject.
- To evaluate the teachers' performance and teaching.
- To extract and classify the aspects based on the subject.
- To measure the polarity of each aspect based on the subject.



ACKNOWLEDGMENT Figure 1. A students' portal framework to analyze student's attitude

There are five major steps involved in the proposed model:

Step-1: Creating the students' web portal

Creating students' web portal can include of selecting the department, choosing the subject, selecting the teaching staff and posting feedback in the text-box. Students belonging to their college can post their opinion. There is no need to mention the students' name.

Step-2: Collecting the feedback from the students' web portal

This step is to collect the students' comments on the subject of the teaching faculty and their traits. Web crawler or designed web API issued to collect the comments from the web portal. This can automatically store in a Hive NoSQL database.

Step-3: Pre-processing the feedback

Pre-processing is done by removing the stopwords, extracting the sentiment/opinionated text, extracting the aspects of the comments posted by the students. Apache Hadoop components are mostly used for Big Data preprocessing. Pig and Hive of Hadoop components will be used to handle the students' large unstructured data for analyzing. It is a data warehouse system of Hadoop that facilitates stakeholders' data summarization, processing ad-hoc queries and analyze large datasets stored in Hadoop file systems. It queries the data using a SQL language called HiveQL. Hive allows the traditional MapReduce programmers to plug in their custom mappers and reducers when it is inconvenient or inefficient to express this logic in HiveQL.

Step-4: Selecting the features

In Aspect based sentiment analysis, machine learning and Natural Language Processing techniques are applied to provide results with better accuracy. All the pre-processed data will be analyzed using classification and pattern mining algorithm of machine learning to measure the polarity of the datasets. It provides a structural design to measure the polarity (positive, negative or neutral) of the students' opinion.

Step-5: Applying the NLP technique to develop a recommender system

In the higher education scenario, recommender system is an intelligent agent that suggests different options for student and educators. This recommender system in this work yields changes required in teaching methodology, different learning styles and curriculum designed to meet the industry requirement. The learning algorithms are used with NLP acquires rules for suggesting the recommender system. It looks as the good probability of success for recommendations.

# VII. CONCLUSION

The proposed methodological framework uses the data and models to predict students' progress and performance which will be helpful in understanding learning behaviors and design suitable intervention strategies. This also motivates the students to use internet to write feedbacks about the course which they study in every semester and discuss with their friends through chat rooms, mails and social media sites.

### REFERENCES

- [1] Nabeela Altrabsheh, Mohamed Medhat Gaber and MihalelaCocea, "SA-E: Sentiment Analysis for Education", http://eprints.port.ac.uk/11394/1/idt13-014-published.pdf, Internet source as on 25/7/2015.
- [2] Miaomiao Wen, Diyi Yang and Carolyn Penstein Rose, "Sentiment Analysis in MOOC Discussion Forums: What does it tell us?", http://www.cs.cmu.edu/~mwen/papers/edm2014camera-ready.pdf, retrivedon 25/7/2015.
- [3] PagarePallavi K. "Analyzing Social Media Data for Understanding Student's Problem." IJCA Proceedings on Innovations and Trends in Computer and Communication Engineering 3, December 2014, pp. 17-22.
- [4] N. D. Valakunde and M. S Patwardhan, "Multi-Aspect and Multi-Class Based Document Sentiment Analysis of Educational Data Catering Accreditation Process", International Conference on Cloud & Ubiquitous Computing & Emerging Technologies, IEEE 2013, pp. 188 – 192.
- [5] M. Ravichandran and G. Kulanthaivel, "Twitter Sentiment Mining(TSM) Framework based Learners Emotional State Classification and Visualization for E-Learning System", Journal of Theoretical and Applied Information Technology,November 2014, Volume 69 No 1.
- [6] SheetalR.Patodiya and Namrata Shroff, "Evaluate the Positive Aspects of Student Learning Experience in Social Media Using Naïve Bayers Algorithm", International Journal of Scientific Research, May 2015, Volume 4 Issue 5, pp.3-4.
- [7] Gautam, Geetika, and Divakar Yadav. "Sentiment analysis of twitter data using machine learning approaches and semantic analysis." In Contemporary Computing (IC3), 2014 Seventh International Conference on IEEE, 2014, pp. 437-442.
- [8] Nabeela Altrabsheh, MihaelaCocea, and SanazFallahkhair, "Learning Sentiment from Students' Feedbackfor Real-Time Interventions in Classrooms", 2014International Conference on Adaptive and Intelligent Systems, Springer International Publishing Switzerland 2014, LNAI 8779, pp. 40–49.
- [9] P. Bharathisindhu and S. SelvaBrunda, "Identifying E-Learner's Opinion using Automated Sentiment Analysis in E-Learning", International Journal of Research in Engineering and Technology, Jan 2014, Volume 03 Issue 01, pp.479 -482.
- [10] ZiedKechaou, Mohamed Ben Ammar and Adel.MAlimi,
  "Improving e-learning with sentiment analysis of users' opinions", 2011 IEEE Global Engineering Education Conference (EDUCON) Learning Environments and Ecosystems in Engineering Education, IEEE 2011, pp. 1032 1038.

- [11] Medhat, Walaa, Ahmed Hassan, and HodaKorashy, "Sentiment analysis algorithms and applications: A survey" Ain Shams Engineering Journal 2014, Volume 5 issue 4, pp. 1093-1113.
- [12] Bing Liu, "Sentiment Analysis and Opinion Mining", Morgan and Claypool Publishers, May 2012.