Enhancing Student Learning through Social Network Experiences

[1]Dinesh Zende,[2]Priyanka Lokhande
[1]Asst.Professor,[2] Post Graduate Student
Vidya Pratishthan's Kamalnayan Bajaj Institute of
Engineering and Technology, Baramati, Pune, India.

***Abstract*- In Social Networking Microblogging sites enables the user to share their everyday encounters (i.e., shortest updated message) with friends within a second. Engineering student shares their daily experiences with friends on Social Networking sites such as Twitter, Facebook, etc. A large amount of data is generated from the student experiences; this helps for educational Learners and Learning analytics to understand the student experiences on Social network. In our system, we have analyzed engineering student Twitter post, i.e., Tweets for identifying the positive and negative aspects of learning.**

***Keywords*- Education, computers and education, social networking, web text analysis.**

I. INTRODUCTION

In recent time social network use is spreading like a virus. The short nature of some of the social networks such as Twitter enables a user to share their every moment with their friends in quick time. Engineering students are the most frequent user of these social networks; they share their everyday encounter about learning the process.

In the old days, Educational researcher gathers student learning related comments and feelings through Focus interviews, survey and Classroom Activities. This data is not that much helpful for understanding student learning experiences as the student behavior in a classroom is more formal than on the social network. For getting deep insight into student learning, we have analyzed student posted content on social network, i.e., Twitter Tweets.

It 's hard to understand a big amount of student-generated data manually, and no any automatic algorithm gains deep knowledge about learning experiences. For grounding student-produced data we have developed a workflow that combines both the qualitative and large scale data mining techniques. In this study we mainly focus on Engineering Student Twitter posts[1]:

1) The engineering student is facing problems in recruitment and success. They are the essential part of Nation's future.

2) Understanding student worries and good things about learning the process, this will help educational researchers for improving education quality and success.

II. RELATED WORK

K Bahar et al. [2] proposed a framework that identifies student interest, their learning performance, etc. The framework considers both the teacher and student relationship into account. The teacher tweets hashtag related to key concepts. The comments of the students are retrieved using Twitter APIs and stored in a local database. The teacher analyzes the retrieved comment/post to modulate student views. The performance is calculated on the same hashtag before and after lecturing.

K.E. Arnold [3] Learning analytic and educational data mining analyze the student generated structured data such as classroom activities, their results and the data generated from the Course Management System (CSM).The researchers at the Purdue University developed the system, i.e., Signal, which analyzes the student performance data such as a number of hours spend on specific topic preparation, time spend in course discussion and their grades. Signal gives red, yellow and green depending upon their progress.

The Goffman's [4] has stated the theoretical base of social performance from the informal conversation of data on the web. This theory said the actual gap between front stage and backstage performance of the people on the social network. As compared to front stage backstage behavior is more relaxing and which promote the student to discuss their everyday encounter, feeling, their interest, etc.

A.Go et.al [5] developed a framework which analyzes the sentiment of Twitter shortest updated message quickly. This opinion is useful for distant understanding about the product before purchase. Companies can also monitor the public sentiment of their brands. The machine learning algorithms such as Naïve Bayes, Maximum Entropy or SVM are used for sentiment classification.

III. PROPOSED SYSTEM

We have developed a workflow that combines both manual analysis and large-scale data mining technique. The developed framework is organized into three phase, the collection of data, data cleaning, the training phase and classification phase respectively. The system architecture is as shown in Fig.1. The method can be useful to achieve deeper and finer understanding of student's experiences especially their learning groundings such as feelings, like and dislikes, etc.

***Gathering of Data*:**

In this phase, we have gathered learning-related data of engineering student from Twitter. We have used a different hashtag for collecting data related to engineering student for example:#engineeringProblems, #ladyEngineer, #engineeringPerks etc.

Manual Analysis of Data:
In this phase, we have manually analyzed the social media content and developed a set of categories as below. Analyzing social media content 's hard because the diversity of the languages, posting times, ambiguous learning experiences, etc.

Negative Categories: [1]

1) Heavy study load: The classes, homework, exams, and labs dominate the student's life. Libraries, lab, are the most frequently visited places. Students express a very stressful incident in engineering. They are not able to manage the heavy study load that results to such as lack of social engagement, lack of sleep, stress, depression, and some health problems.
For example: *"going to bed at 3 A.M. Still have about 8 hrs of homework and studying to do…"*

2) Lack of social engagement: Student needs to sacrifice the time for social engagement and enjoying holidays and special occasion with family and friends to do homework, and to prepare for classes and exams. Social engagement can provide support for releasing stress for students.
For example: *"I feel like I am hidden from the world life of an Engineering student."*

3) Negative emotion: As their lots of negative emotions flowing in tweets, so they have categories "negative emotion" when it explicitly expresses negative emotions such as hatred, anger, stress, sickness, depression, disappointment, and despair.
For example: *"I hate finals."*

4) Sleep problem: Because of lots of studies, sleep problem is widely common among engineering students. Student suffers from a lack of sleep because of heavy study load.
For example: *"If I don't schedule in sleep time, it doesn't happen."*

5) Diversity issues: There are different gender (male/female) students and male students in engineering are regarded as bad at talking with female students because they usually do not have many female students around in their class.
For example: *"I'm sorry. We're not used to having girls around."*
Another Diversity is, there are too many foreign professor and students around. Students say that they cannot understand the accents of foreign professors in class, and they do not like to work with international students in course projects.
For example: *"Avoiding doing group work with international students."*

Positive Categories:

1) Enjoying study: The classes, homework, exams, and labs dominate the student's life. Most of the Students express a very enjoying life in engineering. They can manage the heavy study loads that result into positive emotions about learning and full of social engagement with friends and family.
For example: *"going to bed at 1 A.M. Still have capacity of 5 hrs to do homework and studying . . .."*

2) Full of social engagement: Student can manage the time for enjoying the social life. A student enjoys the special occasion with friends and family, plays various games with friends with by doing an everyday study. Social engagement can provide support for releasing stress for students.
For example: *"I have got distinction in B.E with enjoying the T-20 and IPL match ...love engineering."*

3) Positive emotions: As there lots of positive emotions flowing in tweets about the study, exam, etc. The student is enjoying the creative environment of study.This environment able the student to think about innovative ideas of the project, research, etc.
For example: *"I have achieved first prize in State level project competition....proud to become an Engineer."*

Probable Words in each Category:[1]

Engineering Problem:

Category	Probable Words
Heavy Study load	Hour, homework, exam, day, classes, negtoken, problem, all, study, week, toomuch, labs, worked, still, out, time, page, library, spend, today, long, school, due, engineer, already
Lack of social engagement	Negtoken, Friday, homework, out, study, work, weekend, life, class, engineer, exam, drink, break, Saturday, people, social, lab, spend, tonight, watch,game,miss,party,sunny,beautiful
Negative Emotions	Hate, f***, hit, exam, negtoken, week, class, hell, engineer, suck, tried, damn, death, hard, study, hour, homework, time, equate, FML, lab, sad, bad, feel
Sleep Problem	Sleep, hour, night, negtoken, bed, allnight, exam, homework, nap, coffee, day, long, morning, wake, awake,

	no, sleep, time, study, more, work, class, dream, ladyEngineer, late, week
Diversity Issues	Girl, class, only, negtoken, guy, engineer, Asia, professor, speak, English, female, hot, kid, foreign, out, white, black, more, too much, walk, people, teach, understand, chick, China

IV. EXPERIMENTAL SETUP

Dataset:

We have gathered engineering student Twitter post using Twitter Search API. For an accessing Twitter database, we have used the twitter4j-core-4.0.4.jar .We have collected 1000 tweets related to engineering problems and engineering perks. Following are some example of engineering student Tweets.

Negative Tweets:

1. *RT @HugotInhinyero: Wish we could turn back time to the good old days. When our mama sings us to sleep but now we're stressed out because of exams.We want a long break from this work …#engineeringProblem*
2. *Engineering is only spending too much time on study, exam, unable to enjoy the weekend, no sleep..hate engineering #engineeringProblem.*
3. *RT @engineerproblem: Engineering is Over-simplifying the problem, over complicating the solution #engineeringMajors*
4. *The format my geology professor wants his lab reports gives me anxiety. #EngineeringProblems*
5. *If you haven't spent countless hours in front of excel sheets... Are you an engineer??? #EngineeringProblems*

Positive Tweets:

1. *This project proposal is the first paper I've had to write in a year and a half. #engineeringperks*
2. *Having nicer jobs than all your non-engineering friends #engineeringperks*
3. *The library is nice to study and do homework, but I think having a whole building to yourself wins. #EBSLiving #seniorlife #engineeringperks*
4. *Nothing better than feeling like you're doing something to help someone #engineeringperks*
5. *I feel so powerful because of the fact I can get into thorn when other students can't #engineeringperks #whoruntheworld.*

Data Cleaning Phase:

As the social media content, i.e., Twitter Tweets are the shortest message; which may contain spelling mistakes, meaning is unclear, some special symbols, etc. For perceiving in-depth knowledge about student learning experiences from such data; Text preprocessing is the necessary phase. We have removed the common stop words using Porter stemmer, and Krovetz stemmer (kstem 3.7.jar) is used for stemming. Other special symbols, links, and Hashtag are removed through code.

Classification Phase:

In this step, we have used Naive Bayes Multinomial Classifier for classification of engineering problems and perks. Naive Bayes is a supervised learning technique. We implemented Multilabel classifier using one-vs.-all Heuristic; Multilabel classifier means one tweet may fall into multiple labels for example: *"Why am I not in business school?? Hate being in Engineering school. Too much stuff. Way too complicated. No fun"* falls into heavy study load, and negative emotion at the same time. We have calculated the probability of a tweet for each category and it the calculated probability is greater than the Threshold value (i.e., 0.3 for Negative category and 0.2 for positive category)then the label of that category is get assigned. Following equation 1 and 2 are used for probability calculation:

$$p(C|Ti) = \frac{p(T_i|C).p(C)}{p(T_i)} \alpha \prod_{K=1}^{K} p(W_{iK}|C).p(C) \quad \text{……..(1)}$$

$$p(C'|Ti) = \frac{p(T_i|C').p(C')}{p(T_i)} \alpha \prod_{K=1}^{K} p(W_{iK}|C').p(C') \text{…… (2)}$$

Where,
C-Category
T_i-Document i.e., Tweet
W_{ik}-k th word in ith Tweet

Evaluation Measures:

The standard evaluation measures to evaluate the performance of classifier are Accuracy, Precision, Recall and F1 tests. But these assessment measures cannot be directly applicable for Multilabel classifier. There are two evaluation measures for multilabel classifier as below:

1) Example Based Evaluation Measures: Example based measures are calculated for each example (tweet) and then average over the entire examples.

$$\text{Accuracy a}= \frac{1}{X}\sum_{i=1}^{X}\frac{M_i \cap N_i}{M_i \cup N_i} \quad \text{.....(3)}$$

$$\text{Recall r} = \frac{1}{X}\sum_{i=1}^{X}\frac{M_i \cap N_i}{M_i} \quad \text{.....(4)}$$

$$\text{Precision p}= \frac{1}{X}\sum_{i=1}^{X}\frac{M_i \cap N_i}{N_i} \quad \text{.....(5)}$$

$$\text{F1} = \frac{1}{X}\sum_{i=1}^{X}\frac{2.p_i.r_i}{p_i+r_i} \quad \text{.....(6)}$$

Where,
X=Total number of Tweet
M_i=True set of label assigned by us
N_i=Predicate set of label assigned by classifier

2)Label Based Evaluation Measures: As the label based accuracy is not correct measure to report label imbalanced, so we have not discussed that for performance evaluation of Multi-label classifier.

V. RESULT

We have achieved the result for the positive and negative category as in Fig.2.

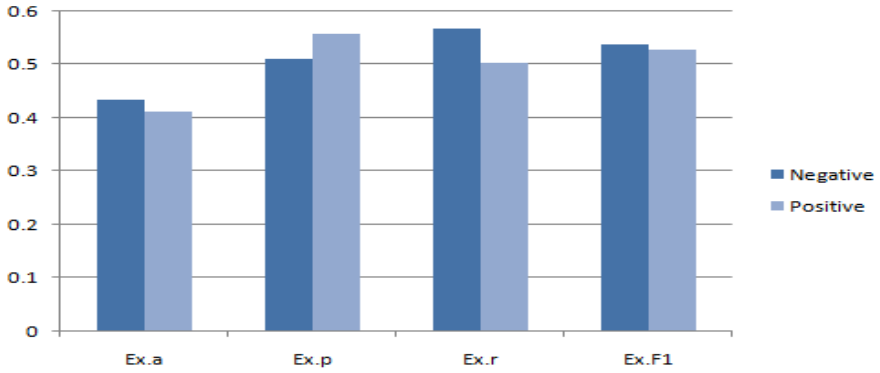

Fig 2.Result Analysis for Positive and Negative Categories

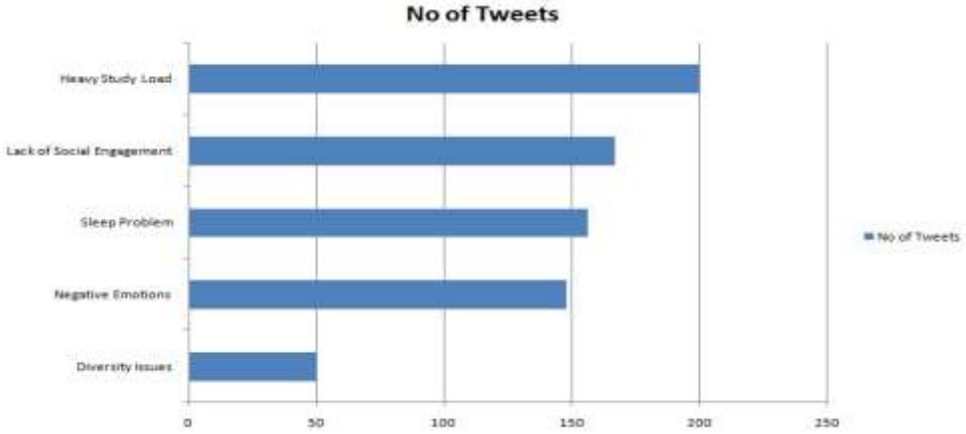

Fig.3. Number of Tweets for each Negative Category

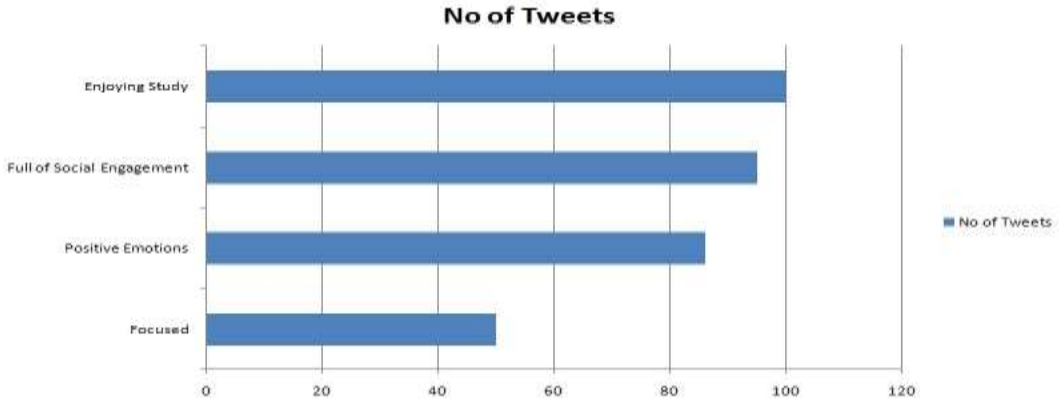

Fig.4..Number of Tweets for each Positive category

VI. CONCLUSION

This system provides a workflow for analyzing social media content for educational purpose that defeats the major drawbacks of manual analysis and large-scale data mining technique. We have analyzed both the positive and negative thought about learning though designed workflow. This analysis is useful for educational researchers and learning analytic for understanding learning related problems.

VII. REFERENCE:

[1] Xin Chen, Student Member, IEEE, Mihaela Vorvoreanu, and Krishna Madhavan " Mining Social Media Data for Understanding Students' Learning Experiences," IEEE transaction on Learning Technologies, vol. 7, no. 3, July-Sept. 2014.
[2] K.Bahar,C.Cemre,H.Elif,B.Guil and D.sercan "Using Twitter as a Diagnostic Teaching and Learning Assessment Tool".International Computer Institute, Ege University Izmir, Turkey, 35100,2014.
[3] K.E. Arnold and M.D. Pistilli, "Course Signals at Purdue: Using Learning Analytics to Increase Student Success," Proc. Second Int'l Conf. Learning Analytics and Knowledge, pp. 267-270, 2012.
[4] E. Goffman's The Presentation of Self in Everyday Life. Lightning Source Inc., 1959.
[5] A. Go, R. Bhayani, and L. Huang, "Twitter Sentiment Classification Using Distant Supervision," CS224N Project Report, Stanford pp. 1-12, 2009.
[6] D. Davidov, O. Tsur, and A. Rappoport, " Enhanced SentimentLearning Using Twitter Hashtags and Smileys," Proc. 23rd Int'l Conf. Linguistics: Posters, pp. 241-249, 2010.
[7] R. Baker and K. Yacef, "The State of Educational Data Mining in 2009: A Review and Future Visions," J. Educational Data Mining, vol. 1, no. 1, pp. 3-17, 2009.

Author Profile

Prof. Dinesh A. Zende is currently working as Assistant Professor, Department of Information Technology Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology, Baramati (Formerly College of Engineering, Baramati) Pune, Maharashtra, India. His research interests are Machine Learning, Social Media Data Mining and Analysis. He is professional member of IEEE, Life Member of CSI and ISTE.

Priyanka P. Lokhande received the Bachelor degree (B.E.) in Computer Engineering in 2013 from Maharashtra Institute of Technology, Kothrud, Pune University. Currently, she is pursuing Master's degree in Computer Engineering at Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology, Baramati, Pune University. Her current research interests include Data Mining.