ISSN 2395-1621

Mining Social Networking site for Digging Students Emotional Behaviour

#1Vidya Shendge, #2Sagar Manisha, #3Nayantara Daure, #4Prof. Suvarna Satkar

1vidyashendge27@gmail.com
2sagarmanu@gmail.com
3nayandaure18@gmail.com

#123Deparment of Computer Engineering
#4Prof. Department of Computer Engineering
G.H. Raisoni College of Engineering and Management
Wagholi,Pune.

ABSTRACT

Social media sites such as Twitter, Facebook, and YouTube provide great venues for students to share joy and struggle, vent emotion and stress, and seek social support. On various social media sites, students discuss and share their everyday encounters in an informal and casual manner. Students' digital footprints provide vast amount of implicit knowledge and a whole new perspective for educational researchers and practitioners to understand students' experiences outside the controlled classroom environment. In this paper, a work-flow is developed which combines both qualitative investigation and large-scale data mining scheme. It is found that certain issues like heavy study load, hectic schedule and lack of sleep are encountered by the students. Hence these issues are classified using Naive Bayes Multi-label Classifier algorithm. This classification can help in understanding the student's problem in a very efficient way.

ARTICLE INFO

Article History

Received :30th December 2015

Received in revised form :

31st December 2015

Accepted : 1st December , 2015

Published online :

2nd December 2015

Keyword: **Data mining, social media, text mining.**

I. INTRODUCTION

Students' informal conversations on social media(e.g Twitter, Facebook) shed light into their educational experiences-opinions, feelings, and concerns about the learning process. Data from such un-instrumented environments can provide valuable knowledge to inform student learning. Analyzing such data, however can be challenging. The complexity of students' experiences reflected from social media content requires human interpretation. However, the growing scale of data demands automatic data analysis techniques. Twitter posts of engineering students' is focused to understand issues and problems in their educational experiences. Engineering students encounter problems such as heavy study load, lack of social engagement and sleep deprivation are considered. A multi-label classification algorithms to classify tweets reflecting students' problems is implemented.

Nowadays, there are numerous social media sites like twitter, facebook, photo bucket etc. These sites provide people a way to express their thoughts and feelings in front of huge amount mass. Social media enables us to be connected and interact with each other anywhere and anytime – allowing us to observe human behaviour in an unprecedented scale. This provides golden opportunities to understand individuals at scale and to mine human behavioural patterns otherwise impossible.

Social media sites also provide a way to advertise and share with people hence they are now used in various fields like politics and educational systems. Social media having endless advantages also comes with few disadvantages like overuse of these sites by people, posting objectionable things or using it for harming others.

Social media comes in various categories like some sites are for sharing of data in purely text format while some are for sharing pictures and videos.

More people are becoming interested in and relying on the social media for information, breaking news and other diverse subject matters. They find out what other people's views are about certain product/service, film, school etc.

Organizations are now conscious of the significance of the opinion of consumers which they post on social sites to the development of their products or services. Moreover, personalities make efforts to protect their image and are being conscious of how they are perceived on these sites.

With the rise of social media, the web has become very vibrant and lively. Hence more and more people are actively participating in these sites. Social media has become an ever increasing field in today's world.

II. SOCIAL MEDIA MINING

Social media mining is the process of representing, analyzing, and extracting meaningful patterns from data in social media, resulting from social interactions of people.It is a field which encompasses techniques from computer science, data mining, , social network analysis, network science, sociology and mathematics. Mining is also called as "Knowledge Discovery" i.e to discover or gain knowledge from raw data. In social media mining we mine the knowledge from various posts generated by people. These are then analysed and used for producing the results.

In this process information is collected, analysed, classified according to need and then final results are produced. Hence the mining helps in understanding individuals better, which can be used to design better computing systems tailored to individuals' needs that will serve them and society better.

III. CHALLENGES IN SOCIAL MEDIA MINING

Social media mining comes with various challenges which poses certain difficulty in mining process.

The first challenge is its enormous size.

Second is the user generated data which comprises of noise and un-structured data.

Third is to classify the data which is of use and dump the remaining data.

IV. RELATED WORK

1. Text mining

Text mining is the process of formatting the given text (parsing, with addition of some linguistic features, addition of some subsequent data), forming the structured database and evaluation and interpretation of output. The good quality of text mining refers to the combination of relevance, interestingness.The real time application is to scan the set of documents which is natural language and document set for the predictive classification.

How to do text mining?

1. Information retrieval: Collecting or identification of set of text documents, taken from the social media which is posted by the user.
2. Natural language processing: In this we are recognize the part of speech which is tagged, syntactic parsing.

Text mining, also referred to as *text data mining*, roughly equivalent to text analytics, refers to the process of deriving high-quality information from text. High-quality information is typically derived through the devising of patterns and trends through means such as statistical pattern learning. Text mining usually involves the process of structuring the input text (usually parsing, along with the addition of some derived linguistic features and the removal of others, and subsequent insertion into a database), deriving patterns within the structured data, and finally evaluation and interpretation of the output. 'High quality' in text mining usually refers to some combination of relevance, novelty, and interestingness. Typical text mining tasks include text categorization, text clustering, concept/entity extraction, production of granular taxonomies, sentiment analysis, document summarization, and entity relation modeling (*i.e.*, learning relations between named entities).

V. PROPOSED WORK

First a sample is taken from student and then it conduct qualitative analysis on that sample which is associated to engineering students educational life. It found engineering students encounter problems such as heavy learning load, lack of social meeting, and sleep deficiency. stand on these outcomes, authors apply a multi-label classification algorithm to categorize tweets presenting student's problems. Then decision tee algorithm is applied to make more accurate result it will perform filtering after that used the algorithm to prepare a detector of student problems. This study presents a tactic and outcome that demonstrate how casual social media data can present insight into student's incident. In this study it implemented a multi-label classification model where we permitted one tweet to go down into many categories at the same time. Our categorization is compared with other generic classifications. Our work expands the range of data-driven approaches in teaching such as learning analytics and educational data mining. The important point in proposed study are, First, it propose a workflow to bridge and integrate a qualitative research methodology and large scale data mining techniques.

It base our data-mining algorithm on qualitative insight resulting from human understanding, so that it can gain deeper understanding of the data. Then apply the algorithm to another large-scale and unexplored dataset, so that the physical method is improved. Second, the paper provides deep insights into engineering student's educational experiences as reacted in informal, uncontrolled environments. Many issues and problems such as study-life balance, lack of sleep, lack of social engagement, and lack of diversity clearly emerge. These could bring awareness to educational researcher, policy-maker.

Mining Twitter Data

Twitter is one of the most popular site. This social media site is public i.e the data or the content of Twitter is very concise. It allows only 140 characters to be read or tweet by the users. Twitter provides an application interface which can be used to stream the data. Various analysis methods can be used to mine the data. Analysis methods include qualitative content analysis, quantitative analysis, etc. But mining social media data is a very difficult task. Many methodological difficulties arise during analysing the huge amount of dataset i.e. textual dataset.

Cluster

Clustering and classification are both the fundamentals of data mining. Classification are also known as the supervised machine learning and clustering also known as the unsupervised machine learning technique.

Clustering is the method of finding the similar data set and collect int the same group which is more similar to each other. The main base of clustering is to grouping the data based on the given data description.

VI. CONCLUSION

The conclusion is based on mining social media data which helps in recognizing the student's problems. Mining social media data is helpful to researchers in learning analytics, educational data removal, and learning skill.

REFERENCES

[1] Xin Chen, MihaelaVorvoreanu, and Krishna Madhavan."Mning Social Media Data for Understanding Students'Learning Experiences" IEEE tarnsactions on learning Technologies, ID, DOI 10.1109/TLT.2013.2296520.

[2] "Using the Twitter Search API | Twitter Developers https://dev.twitter.com/docs/using-search , 2013.

[3] Ferguson, "The State of Learning Analytics in 2012: A Review and Future Challenges," Technical Report KMI-2012-01, Knowl-edge Media Inst. 2012.

[4] IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS, VOL. 45, NO. 1, FEBRUARY 2015Integrating Human Behavior Modeling and DataMining Techniques to Predict HumanErrors in Numerical TypingCheng-Jhe Lin, Changxu Wu, Member, IEEE and Wanpracha A. Chaovalitwongse, Senior Member, IEEE.

