

Mining Social Networking site for Digging Students Emotional Behaviour

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



¹vidyashendge27@gmail.com

²sagarmanu11@gmail.com

³eyestarnayan@gmail.com

^{#123}Department of Computer Engineering

^{#4}Prof. 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.

Keyword: Data mining, social media, text mining.

ARTICLE INFO

Article History

Received: 13th May 2016

Received in revised form :

13th May 2016

Accepted: 16th May 2016

Published online :

17th May 2016

I. INTRODUCTION

Social media is a group of internet-based applications that build on the ideological and technological foundations of Web. It has become part and parcel of life. It is not only popular and affordable but mostly used platform where people share their experiences with the world. Social media shatters all the geographical boundaries and allows communicating with the whole world.

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. LITERATURE SURVEY

Earlier offline procedures were carried out to study such problems [2][3]. These problems included surveys, focus groups, interviews and other such classroom programs. Such

programs are generally carried out in front stage environment. A front stage environment is a controlled environment, where a person is likely to express superficially and not transparently [4][5], whereas a backstage environment is a relaxed environment, where one has no pressure to answer a question in a particular way. Such a platform can be online social network like Facebook, Twitter which is very frequently used by the students and it is their spontaneous hub too. Twitter is one of the many popular social networking websites. There is a provision of API which is free of cost, which can be used to stream data. Therefore, the analysis of tweets can be done on twitter. Twitter allows 140 characters per tweet so its conciseness also helps in easy streaming of data. A hash tag is a word that begins with a „#“ which means all the content related to the hash tag name will be tagged or added in that particular hash tag. Analysis was carried out on engineering students because engineers are said to be the future of any nation. Their learning process has to be strong and has to be upgraded for better adaptability to technology [6]. The hashtag #enggproblems was taken into consideration and was examined. Here the students posted more about their problems faced in their learning system. The tweets were worked upon as a large process where the tweets were said to fall under various category such as heavy study load, diversity issues, lack of social engagement, negative emotion and sleep problems. These categories were built by human examination of tweets falling under #enggproblems. The human inspection of such a data is framed as inductive content analysis. [7].

III. 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.

3. Name entity recognition: In this we are identifying the named text like:- the name of the peoples, place, organization, gestures and the postures symbols, etc.

4. Pattern identified entities: Some characteristics like the telephone no., e-mail id, and address can be identified by using the regular expression or by matching of data by using concern data sets.

5. Conference: Identifications of grammatical noun phrases using the posted dataset.

6. Sentiment analysis: In this process, we are identifying the user's mood, emotion, opinion, sentiments; by using this interpreter helps to identifying the concepts and the opinion holder and opinion objects.

7. Qualitative text analysis: The semantic or the grammatical relationship between the words is extracted from the posts made by the users on the social media to find the meaningful text.

III. PROPOSED WORK

This paper targets only the student's generated data on the sites. Students regularly post their experiences, thoughts, ideas on these sites. This raw and unformatted data is taken, classified into various categories and according to those categories results are produced. Sentiment analysis can be referred to as discovery and recognition of positive or negative expression of opinion by students. It deals with the feelings of users. Opinions expressed on social media by users are often convincing and these indicators can be used to form the basis of choices and decisions made by people. Consequentially it has become necessary to analyse sentiment expressed by users using data mining techniques in order to generate a meaningful framework that can be used as decision support tool. This can be made possible by employing algorithms and techniques to ascertain sentiment.

Cluster

Clustering and classification are both the fundamentals of data mining. Classification is 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.

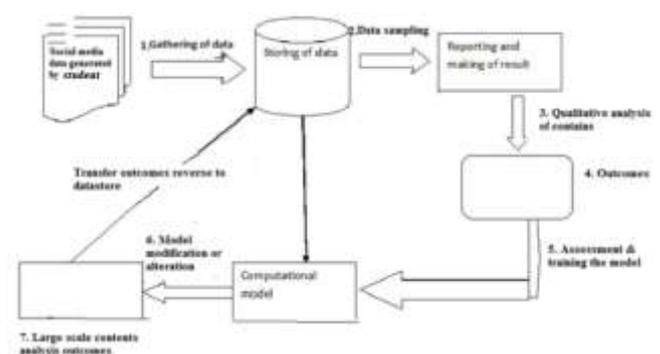


Fig 1. System architecture

IV. ALGORITHM

Algorithm:

Step 1: Start

Step 2: User login

Step 3: User can submit feedback using facing problem.

Step 4: User comments can analysis using navy bays algorithm

Step 5: Remove all unnecessary symbols.

Step 6: Check all keywords and classify one class

Step 7: Clustering all comments and identify facing problem

Step 8: Final output can analysis by admin panel actually what type of problem student can faces.

Step 9: Stop

V. RESULT



Fig 2. Overall analysis result depending on student feedback

Above fig 2. Shows that overall analysis of all classification result. When student send feedback then automatically classify the emotion. Here admin check student problems using this analysis graph.



Fig 3. Girls and boys analysis count

Above fig 3. Shows that how many users (girls or boys) are registered for sending the emotions. Here we count total no of users for checking which user facing problem more.

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. It gives a way

to examining social medium statistics that conquer the main restrictions of both physical qualitative analysis and huge scale computational study of user produced textual Content.

REFERENCES

- [1] Al-Khalifa, H. S., and Davis, H. C. Harnessing the wisdom of crowds: How to semantically annotate web resource using folksonomies. In Proceedings of IADIS Web Applications and Research (WAR2006) (2006).
- [2] Alani, H., Dasmahapatra, S., O'hara, K., and Shadbolt, N. Ontocopi - using ontology based network analysis to identify communities of practice. IEEE Intelligent Systems 18 (2003), 18–25.
- [3] Andriessend, E., Soekijad, M., Veld, M. H. i. t., and Poot, J. J. Dynamics of knowledge sharing communities. Tech. rep., Delft University of Technology, 2001.
- [4] Angeli, C., Bonk, C., and Hara, N. Content analysis of online discussion in an applied educational psychology course. [On-line] (1998).
- [5] Arcee, R. Using computer-mediated communication in an educational context: Educational outcomes and pedagogical lessons of computer conferencing. The Electronic Journal of Communication 3, 2 (1993). http://www.cios.org/getfile/ARCHEE_V3N293 [restricted access].
- [6] Bales, R. F. Interaction Process Analysis. A Method for the Study of Small Groups. Cambridge, MA: Addison-Wesley, 1950/1951.
- [7] Balsamo, A. Technologies of the Gendered Body: Reading Cyborg Women. Duke University Press, New York, December 1995.
- [8] Bandura, A. Self-efficacy: Towards a unifying theory of behavioral change. Psychological Review 84 (1977), 191–215.
- [9] Berglund, A., and Eckerdal, A. What do cs students strive to learn? Computer Science sEducation 16, 3 (2006), 185–195.
- [10] Berkson, L. Problem-based learning: Have the expectations been met? Academic Medicine 68, 10 (1993), S79–S88.