Interpreting Students Learning Behavior And Issues Using Social Media Data Mining.

[1]Supriya Sumbe, [2]Pragati lendghar, [3]Priyanka Pachpute, [4]Komal Mundra, [5]Prof. Chhaya R. Jadhav

[1,2,3,4,5] *Department of Information Technology, D. Y. Pimpri, Pune - 412 105*

Abstract — *Students' casual discussions on on-line networking (e.g., Twitter, Face book) shed lightweight into their academic experiences. Their opinions, emotions, and considerations regarding the training method. Info from such instrumented environmental will provide valuable data to tell students learning. Anal zing down such information, on the opposite hand, is difficult. The complexness of students' experiences mirrored from social media substance needs human translation. Then again, the growing size of information demands automatic data analysis strategies. In this paper, we tend to developed work method to integrate each chemical analysis and large-scale data processing. We tend to targeting engineering students' Twitter post to grasp drawback and issue in their academic experiences. We tend to at first conducted a chemical analysis on samples taken from around 25,000 tweets associated with engineering students' life. We tend to discovered that engineering students expertise problems, like overwhelming study burden, absence of social engagement, and lack of sleep. In lightweight of those results, we tend to enforce a multi-label classification rule to classify tweets reflective students' problems. We then used the rule to coach a detector of scholars problems from around thirty five,000 tweets streamed at the geo-area of Purdue University. This work, astonishingly initial time, introduces associate degree approach and results that show however informal social media information will give insights into of students' experiences.*

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

I.   INTRODUCTION

Data mining analysis has effectively made many technique, tools, and algorithms for managing immense amounts of information to answer real-world troubles. As social media is wide used for numerous functions, immense amounts of user created knowledge be gift and may be created out there for data processing. Data processing of social media will enlarge researcher's ability of understanding innovative expertise, to the employment of social medium and develop business intelligence to gift smart services and extend innovative opportunities. Main objectives of the knowledge mining procedure square measure to jointly handle large-scale data, extract unjust patterns, and gain perceptive information.. Social media sites like Twitter, Face book, and YouTube gift grand place to students to share happiness and struggle, sentiment and tension, and gain social support. On numerous social media sites, students refer their everyday encounters in an exceedingly snug and informal manner. This Student's digital info provides immense quantity of implicit info and an entire new viewpoint for instructional researchers to grasp student's experiences outside the prohibited room atmosphere. This understanding will enhance education quality, and therefore improve student employment, preservation, and accomplishment. The immense quantity of knowledge on social sites provides prospective to acknowledge student's downside, however it raises some method complexities in use of social media knowledge for instructional reasons. The complexities like absolute knowledge volumes, the miscellany of web slangs, the modification of locations, and moment of scholars posting on the net. Pure physical analysis cannot contract with the ever growing scale of information, whereas pure automatic algorithms cannot capture in-depth significance within the info.

II.   LITERATURE SURVEY

1] Title : Representation and Communication: Challenges in Interpreting Large Social Media Datasets
Authors:M. Rost, L. Barkhuus, H. Cramer, and B. Brown
Online services offer a spread of opportunities for understanding human behavior through the massive mixture
data sets that their operation collects. The information sets they collect don't unproblematic ally model or mirror the globe events. During this paper we have a tendency to use information from Foursquare, a preferred location arrival service, to argue for the importance of analyzing social media as a communicative instead of naturalistic system. Drawing on logs of all Foursquare check-ins over eight we have a tendency toes we highlight four options of Four square's use: the link between attending and check-ins, event check-ins, business incentives to arrival, and in conclusion

ridiculous check-ins These points show however giant information analysis is stricken by the top user uses to that social networks are place.

2] Title: The State of Learning Analytics in 2012: A Review and Future Challenges
Authors: R. Ferguson
We propose that the planning and implementation of effective Social Learning Analytics (SLA) gift vital challenges and opportunities for each analysis and enterprise, in 3 necessary respects. the primary is that the learning landscape is awfully turbulent at the present, in no tiny half thanks to technological drivers. Online social learning is rising as a major development for a range of reasons, that we tend to review, so as to motivate the construct of social learning. The second challenge is to spot differing kinds of SLA and their associated technologies and uses. we tend to discuss 5 classes of analytic in respect to on-line social learning; these analytics area unit either inherently social or will be liberal. This sets the scene for a 3rd challenge, that of implementing analytics that have pedagogic and moral integrity in a very context wherever power and management over data area unit currently of primary importance. we tend to think about a number of the considerations that learning analytics provoke, and suggest that Social Learning Analytics could offer ways that forward. We tend to conclude by revisiting the drivers and trends, and think about future eventualities that we tend to may even see unfold as SLA tools and services mature.

3] Title: Microblogging in Classroom: Classifying Students' Relevant and Irrelevant Questions in a Micro blogging-Supported Classroom
Authors: S. Cetintas, L. Si, H. Aagard, K. Bowen, and M. Cordova-Sanchez
Microblogging could be a in style technology in social networking applications that lets users publish on-line short text messages (e.g., but two hundred characters) in real time via the net, SMS, instant electronic messaging shoppers, etc. Microblogging may be a good tool within the room and has of late gained notable interest from the education community. This paper proposes a unique application of text categorization for 2 styles of microblogging queries asked during a room, particularly relevant (i.e., queries that the teacher desires to deal with within the class) and digressive queries. Empirical results and analysis show that exploitation personalization beside question text ends up in higher categorization accuracy than exploitation question text alone. it's additionally helpful to utilize the correlation between queries and on the market lecture materials likewise because the correlation between queries asked during a lecture. moreover, empirical results additionally show that the elimination of stop-words ends up in higher correlation estimation between queries and ends up in higher categorization accuracy. On the opposite hand, incorporating students' votes on the queries doesn't improve categorization accuracy, though the same feature has been shown to be effective in community question respondent environments for assessing question quality.

4] Title: Online Identity Management Literacy for Engineering and Technology Students
Authors: M. Vorvoreanu, Q.M. Clark, and G.A. Boisvenue,
This paper establishes the necessity for together with on-line identity management acquirement in college man education, as a part of making ready students for coming into the work market. It discusses the impact of on line data on employment, and presents original interview knowledge regarding engineering and technology college man students' on-line identity management practices. The paper argues for the necessity to show students social media acquirement and proposes a particular arrange for on-line identity management which will be integrated into college man curricula.

5] Title: Predicting Popular Messages in Twitter," Proc. 20th Int'l Conf. Companion on World Wide Web
Authors: L. Hong, O. Dan, and B.D. Davison
Social network services became a viable supply of data for users. In Twitter, data deemed necessary by the community propagates through rewets. Learning the characteristics of such standard messages is very important for variety of tasks, like breaking news detection, customized message recommendation, infective agent selling et al.. This paper investigates the matter of predicting the recognition of messages as measured by the number of future rewets and sheds some lightweight on what sorts of factors influence data propagation in Twitter. We tend to formulate the task into a classification downside and study 2 of its variants by investigation a large spectrum of options supported the content of the messages, temporal data, information of messages and users, likewise as structural properties of the users' social graph on an outsized scale dataset. We tend to show that our methodology will with success predict messages which is able to attract thousands of rewets with sensible performance.

6] Title:Large Scale Max-Margin Multi-Label Classification with Priors
Authors: BharathHariharan, LihiZelnik-Manor, S. V. N. Vishwanathan, ManikVarma
We propose a max-margin formulation for the multi-label classification drawback wherever the goal is to tag an information purpose with a collection of pre-specified labels. Given a collection of L labels, a data point will be labelled with any of the 2L possible subsets. The most challenge so lies in optimising over this exponentially large label house subject to label correlations. Existing solutions take either of 2 approaches. The first assumes, a priori, that there aren't any label correlations and severally trains a classifier for every label (as is done in the 1-vs-All heuristic). This reduces

the problem complexness from exponential to linear and such strategies will scale to giant problems. The second approach expressly models correlations by pair wise label interactions. However, the complexness remains exponential unless one assumes that label correlations are sparse. Moreover, the learnt correlations replicate the coaching set biases

7] Learning Analytics and Educational Data Mining: Towards Communication and Collaboration
Authors: George Siemens, Ryan S J.D. Baker
Growing interest in information and analytics in education, teaching, and learning raises the priority for inflated, high-quality analysis into the models, methods, technologies, and impact of analytics. Two analysis communities – instructional data processing (EDM) and Learning Analytics and data (LAK) have developed separately to handle this want. This paper argues for inflated and formal communication and collaboration between these communities so as to share analysis, methods, and tools for data mining and analysis within the service of developing each LAKand EDM fields

III. PROPOSED SYSTEM

We planned a work flow to place along each qualitative investigation and large-scale data processing theme. Initial a sample is taken from student so analysis conducted on it sample that is associated to engineering student's instructional life. Therefore solely tweets associated with engineering student are collected. It's found that engineering students encounter issues like significant learning load, lack of social meeting, and sleep deficiency. Supported this outcome, a multi-label classification formula that's Naive mathematician Multi-label Classifier formula is applied to categorise tweets presenting student's issues. Then call tee formula is applied to form additional correct result it'll perform filtering. The formula prepares a detector of student issues.

Advantages of Proposed System:
1 .It proposes a work flow to bridge and integrate a qualitative analysis methodology and huge scale data processing techniques.
2. It provides deep insights into engineering student's academic experiences as reacted in informal, uncontrolled environments.

IV. SYSTEM ARCHITETURE
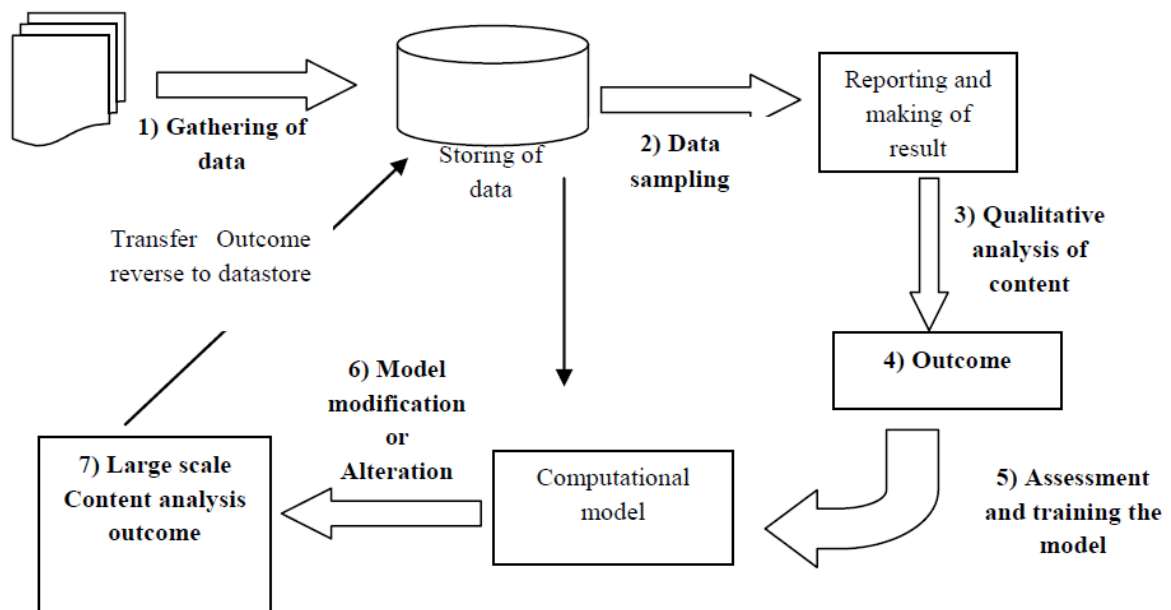
V. OBJECTIVE

1) To demonstrate a workflow of social media data sense-making for educational purposes, integrating both qualitative analysis and large-scale data mining techniques.
2) To explore engineering students' informal conversations on Twitter, in order to understand issues and problems students encounter in their learning experiences.

VI. MOTIVATION

Data mining analysis has effectively made many technique, tools, and algorithms for managing immense amounts of information to answer real-world troubles. As social media is wide used for numerous functions, huge amounts of user created information be gift and may be created on the market for data processing. Data processing of social media will enlarge researcher's ability of understanding innovative expertise, to the utilization of social medium and develop business intelligence to gift smart services and extend innovative opportunities. Main objectives of the information mining procedure are to together handle large-scale data, extract unjust patterns, and gain perceptive information. Social media sites like Twitter, Face book, and YouTube gift grand place to students to share happiness and struggle, sentiment and tension, and gain social support.

VII. CONCLUSION

Mining social media information is useful for researchers in learning analytics, instructional information removal, and learning talent. It offers the way to examining social medium statistics that conquer the most restrictions of each physical analysis and large scale procedure study of user made matter content. 2 algorithmic rules are helpful for this classification 1st the Naive Bayes Multi-label Classifier and Second the Tree construction algorithm. It notifies instructional manager, and alternative applicable assessment manufacturers to expand additional acceptance of engineering students' establishment understanding.

ACKNOWLEDGMENT

REFRENCES

[1] G. Siemens and P. Long, "Penetrating the Fog: Analytics in Learning and Education," Educes Rev., vol. 46, no. 5, pp. 30-32, 2011.

[2] M. Rost, L. Barkhuus, H. Cramer, and B. Brown, "Representation and Communication: Challenges in Interpreting Large Social Media Datasets," Proc. Conf. Computer Supported Cooperative Work, pp. 357-362, 2013.

[3] M. Clark, S. Sheppard, C. Atman, L. Fleming, R. Miller, R. Stevens, R. Streveler, and K. Smith, "Academic Pathways Study: Processes and Realities," Proc. Am. Soc. Eng. Education Ann. Conf. Exposition, 2008.

[4] C.J. Atman, S.D. Sheppard, J. Turns, R.S. Adams, L. Fleming, R. Stevens, R.A. Streveler, K. Smith, R. Miller, L. Leifer, K. Yasuhara, and D. Lund, Enabling Engineering Student Success: The Final Report for the Center for the Advancement of Engineering Education. Morgan & Claypool Publishers, Center for the Advancement of Engineering Education, 2010.

[5] R. Ferguson, "The State of Learning Analytics in 2012: A Review and Future Challenges," TechnicaReport KMI-2012-01, Knowledge Media Inst. 2012.

[6] R. Baker and K. Yacef, "The State of Educational Data Mining in2009: A Educational Data Mining, vol.1, no. 1, pp. 3-17, 2009.

[7] S. Cetintas, L. Si, H. Aagard, K. Bowen, and M. Cordova-Sanchez, "Microblogging in Classroom: Classifying Students' Relevant and Irrelevant Questions in a Microblogging-Supported Classroom,"IEEE Trans. Learning Technologies, vol. 4, no. 4, pp. 292-300, Oct.- Dec. 2011.

[8] C. Moller-Wong and A. Eide, "An Engineering Student Retention Study," J. Eng. Education, vol. 86, no. 1, pp. 7-15, 1997.

[9] Nat'l Academy of Eng., The Engineer of 2020: Visions of Engineering in the New Century. National Academies Press, 2004.

[10] E. Goffman, The Presentation of Self in Everyday Life. LightningSource Inc., 1959.