EFFECTIVE MINING SOCIAL MEDIA DATA FOR UNDERSTANDING STUDENTS LEARNING EXPERIENCES

Payal S.Jain, [1] Pallavi S.Panhale, [2]Praneshwari A.Deokar[3]

[123]Sanghavi College of Engineering, Nashik

Abstract - Students confabulations on gregarious media (e.g. Twitter, Facebook) shed light into their scholastic experiences opinions, feelings, and concerns about the cognition process. Data from such uninstrumented environments can provide valuable cognizance to apprise student learning. Analyzing such data, however, can be arduous. The involution of students experiences reflected from gregarious media content requires human interpretation. However, the growing scale of data demands automatic data analysis techniques. In this paper, we developed a workflow to integrate both qualitative analysis and astronomically immense-scale data mining techniques. We fixated on engineering students Twitter posts to understand issues and quandaries in their edifying experiences. We first conducted a qualitative analysis on samples taken from about 25,000 tweets cognate to engineering students college life. We found engineering students encounter quandaries such as heftily ponderous study load, lack of gregarious engagement, and slumber deprivation. Predicated on these results, we implemented a multi-label relegation algorithm to classify tweets reflecting students quandaries. We then utilized the algorithm to train a detector of student quandaries from about 35,000 tweets streamed at the geo-location of Purdue University. This work, for the first time, presents a methodology and results that show how informal gregarious media data can provide insights into students experiences.

Keywords—Education, computers and education, social networking, web text analysis

1. Introduction

Data mining research has effectively produced several technique, tools, and algorithms for managing huge amounts of data to answer real-world troubles. As social media is widely used for various purposes, vast amounts of user created data be present and can be made available for data mining. Data mining of social media can enlarge researchers ability of understanding innovative experience, to the use of social medium and develop business intelligence to present good services and extend innovative opportunities. Main objectives of the data mining procedure are to collectively handle large-scale data, extract actionable patterns, and gain insightful knowledge.. Social media sites such as Twitter, Face book, and YouTube present grand place to students to share happiness and struggle, sentiment and tension, and gain social support. On various social media sites, students talk about their everyday encounters in a comfortable and informal manner. This Students digital information gives huge amount of implicit information and a whole new viewpoint for educational researchers to know students experiences outside the prohibited classroom environment. This understanding can enhance education quality, and thus improve student employment, preservation, and achievement. The vast amount of information on social sites provides prospective to recognize students problem, but it raises some methodological complexities in use of social media data for educational reasons. The complexities such as absolute data volumes, the miscellany of Internet slangs, the change of locations, and

moment of students posting on the web. Pure physical analysis cannot contract with the ever growing scale of data, while pure automatic algorithms cannot capture in-depth significance inside the data.

2. Motivation of Project

Many times, students gets shy or trepidacious of clearing their quandary during the classroom and this convivial media avail them to just post whatever they feel at that time about their emotions. The schools and departments have been struggling with student recruitment and withholding issues. Graduates play a paramount role in nation future workforce and which directly effects the nation economic magnification and ecumenical competency. The concept of amalgamating Students learning

experiences for enhancing E-learning experiences is innovative for amending the training style or edifying style by which student to redress them at some required time without perpetual concealing or surveying. Predicated on understanding of issues and quandaries in students life, policymakers and educators can make more apprised decisions on opportune interventions and accommodations that can avail students overcome barriers in learning. Student can be trained or amend inculcation quality as they have been relegated. Students learning experiences from convivial media will preserve the time to a mass the data manually.

3. Statement of Scope

- Engineering schools and departments have long been struggling with student recruitment and retention issues. Engineering graduates constitute a paramount part of the nations future workforce and have a direct impact on the nations economic magnification and ecumenical competency.

- Predicated on understanding of issues and quandaries in students life, policymakers and educators can make more apprised decisions on felicitous interventions and accommodations that can avail students overcome barriers in learning.

- Twitter is a popular gregarious media site. Its content is mostly public and very concise (no more than 140characters per tweet). Twitter provides free APIs that can be acclimated to stream data. Therefore, we opted to commence from analyzing students posts on Twitter.

3.1 Goals and Objective

- To classify student based on content shared in the social media.

- To make institution to perform better decision making on student interventionsfor at-risk students, improvement of education quality, and thus enhance student recruitment, retention, and success.

- To integrate both qualitative analysis and large-scale data mining techniques.

4.Architecture

Engineering schools and departments have long been struggling with student recruitment and retention Issues. Engineering graduates constitute a paramount part of the nation future workforce and have a direct impact on the nation economic magnification and ecumenical competency Predicated on understanding of issues and

quandaries in students life, policymakers and educators can make more apprised decisions on felicitous interventions and accommodations that can avail students overcome barriers in learning.3. Twitter is a popular gregarious media site. Its content is mostly public and very concise (no more than 140 characters per tweet). Twitter provides free APIs that can be acclimated to stream data. Therefore, we chose to commence from analyzing students posts on Twitter. In this paper, we went through an exploratory process to locate the pertinent data and pertinent Twitter hash tags (a Twitter hash tag is a word beginning with a denotement, used to accentuate or tag a topic). We accumulated 25,284 tweets utilizing the hash tag engineering Quandaries over a period of 14 months, and a second dataset of 39,095 tweets utilizing the geocode (longitude and latitude) of Purdue University, West Lafayette. This corresponds Three researchers conducted an inductive content analysis on samples of the engineering. Quandaries dataset, which corresponds to steps 2 and 3 in Fig. In step 4,we found that major quandaries engineering students encounter in their cognition experiences fall into several prominent categories. Predicated on these categories, we implemented a multi-label Naive Bayes relegation algorithm. We evaluated the performance of the classifier by comparing it with other state-ofthe- art multi-label classifiers. We utilized the relegation algorithm to train a detector that could avail detection of engineering students quandaries at Purdue University The results could avail educators identify at-risk students and make decisions on felicitous interventions to retain them. This paper makes two major contributions. the performance of the classifiers is estimated by comparing it with other multilabel classifiers. In step 6 The classification algorithm is applied by System to prepare a detector that help recognition of engineering students problems. The results are provided by step 7 help educators to identify at risk students and make decisions on proper interference to preserve them.
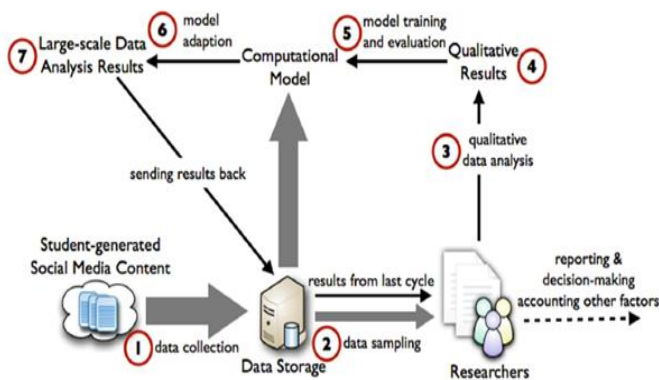
4.1 Navie Bayes Clasifier Algorithm

This algorithm considers each sub words in the review and accordingly classifies the reviews in different categories

Let S is the Sentence

Step 1: Define categories   c={c1,c2,c3,...,cn}

Step 2: Read data from a database.

Step 3: Divide S into sub works{w1,w2,w3...wn} split.

Step 4: Check sub words {w1,w2,w3…..wn} for every categories

Step 5: if words match with categories {c1,c2.c3….cn} increment the counter for that categories Else put that in "other" categories.

Step 6:Find probability of each category

5.Conclusion

Our study is benign to researchers in learning analytics, inculcative data mining, and learning technologies. It provides a workflow for analyzing gregarious media data for inculcative purposes that surmounts the major constraints of both manual qualitative analysis and sizably voluminous scale computational analysis of utilizer-engendered textual content.

Our study can apprise edifying administrators, practitioners and other pertinent decision makers to gain further understanding of engineering students college experiences. As an initial endeavor to instrument the uncontrolled gregarious media space, we propose many possible directions for future work for researchers who are intrigued with this area. We hope to optically discern a proliferation of work in this area in the near future. We advocate that great attention needs to be paid to forfend students privacy when endeavoring to provide good edification and accommodations to them.

6.References

[1] ost, L. Barkhuus, H. Cramer, and B. Brown, Representation and communication: challenges in interpreting astronomically immense convivial media datasets, in Proceedings of the 2013 conference on Computer fortified cooperative work, 2013, pp. 357362.

[2] lark, S. Sheppard, C. Atman, L. Fleming, R. Miller, R. Stevens, R. Streveler, and K. Smith, Academic pathways study: Processes and realities, in Proceedings of the American Society for Engineering Edification Annual Conference and Exposition, 2008.

[3] . 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 prosperity: The final report for the Center for the Advancement of Engineering Inculcation, Morgan Claypool Publishers, Center for the Advancement of Engineering Inculcation, 2010.

[4] erguson, The state of learning analytics in 2012: A review and future challenges, Cognizance Media Institute, Technical Report KMI-2012-01, 2012.

[5] aker and K. Yacef, The state of scholastic data mining in 2009: A review and future visions, Journal of Edifying Data Mining, vol. 1, no. 1, pp. 317, 2009.