

Mining Data for Analyzing Student's Learning - A Survey

Remya R.S,Smitha E S
M.Tech Student, Associate professor
Computer Science Engineering
LBSITW, Thiruvananthapuram,India

Abstract- Social media provide a range of opportunities for understanding human behaviour through the large aggregate data sets that their operation collects. Data Mining is very useful in the field of education especially when examining students learning behaviour in online learning environment. Student's casual conversations on social media show the educational experiences, opinions, feelings and concern about the learning process. The data extracted from the social media provide valuable knowledge to inform student learning experience about how they positively taking the education and what are the benefits they gain from the education and what are their problems and issues.

Keywords- web mining, learning analytics and knowledge, educational data mining, sentiment analysis.

I INTRODUCTION

Today's engineering students will solve tomorrow's problems in a world that is advancing faster and they are facing critical challenges than ever before. Some students don't have a clear vision regarding their reason for being in college. Perhaps they may be here at their parent's insistence while not feeling that they are doing what they really want to do with their life. Many students feel engineering is boring and hard, which leads to lack of motivation to study and negative emotions. Diversity issues also reveal culture conflicts and culture stereotypes existing among engineering students. Sometimes courses and majors are chosen to please others but have little or no relationship to the student's true interests. Many students just aren't sure about that they really want to do in their future career.

Academic problems of students come in various forms such as difficulty in math subject, lack of motivation, study habits, strict teachers and failed in major examinations. Identifying these problems along with their negative attitude towards the engineering program would provide better understanding of student's situation and behavior inside the classroom. We are mainly concentrating on engineering students because they constitute a major part of the nation's future workforce. Engineering students are struggling with recruitment and retention issues.

II LITERATURE SURVEY

Traditionally, educational researchers have been using methods such as surveys, interviews, various classroom activities to collect data related to students learning experiences. Usually these methods are very time consuming, thus cannot be duplicated or repeated with high frequency. And the scale of such studies is usually limited.

For instance, the Academic Pathways Study (APS) is a five year, multi-institution study. The APS is the largest element of the Center for the Advancement of Engineering Education (CAEE)[5]. The primary goal of the Academic Pathways Study was to create a rich and wide-ranging portrait of the undergraduate engineering learning experience. It uses multiple research methods and rely on the students own words for much of the data. The APS addresses the following research questions:

- How do engineering student's skills and knowledge change over time?
- How do engineering student's skills and knowledge develop over time?
- How does one's identity as an engineer evolve?
- What elements of engineering education contribute to the students' skills/knowledge and identity?
- What skills do engineers need when they enter the workplace?

1. LEARNING ANALYTICS AND KNOWLEDGE

Learning analytics and knowledge is the measurement, collection, analysis and reporting of data about learners and their contexts for the purpose of understanding and optimizing learning and environments in which it occurs[7]. It is used for prediction purposes, personalization and adaptation and intervention purposes.

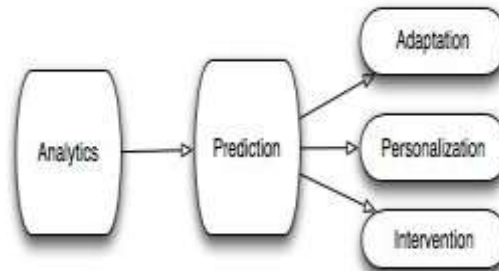


Figure 1: learning analytics

It is used to access academic progress, predict future performance and spot potential issues. Data are collected from explicit student actions and tacit actions. Explicit student actions include completing assignments and taking exams. And tacit actions include online social interactions, extracurricular activities and other activities that are not directly assessed as part of the student's educational progress. The goal of learning analytics is to enable teachers to tailor educational opportunities to each student's level of need and ability.

Learning analytics and knowledge includes the following:

- Social network analysis (e.g., analysis of student-to-student and student-to-teacher relationships and interactions to identify disconnected students)
- Social or attention metadata to determine what a user is engaged with.

This method is used to monitor and predict student's performance in learning [1]. And also used to identify their issues early. These interventions can be used to identify at risk students. It helps us to answer these questions:

- When are students ready to move on to the next topic?
- Whether a student is in at risk or not for completing a course?
- What is the next best course for a given student?
- Should a student be referred to a counselor for help?

2. eMUSE

eMUSE stands for empowering MashUps for Social E-learning. It is a social learning environment. It integrates several Web 2.0 tools (wiki, blog, microblogging tools, social bookmarking tool, media sharing tools) [5]. eMUSE was built in order to support both the students and the instructors to manage their activity on several different social media tools. The platform retrieves student's actions with each Web 2.0 tool. These tools include post blog entry, post blog comment, upload youtube video, post delicious bookmark, add delicious friend to network, add slideshare document, create picasa album, post - tweet, revise wiki page, upload wiki file etc and stores them in a local database for further processing.

The Web 2.0 tools are selected by the instructor for the course at hand. eMUSE is different from mash-up personal learning environments such as MUPPLE. The instructor can configure the learning environment by selecting the Web 2.0 tools to be used by the students before the start of the course.

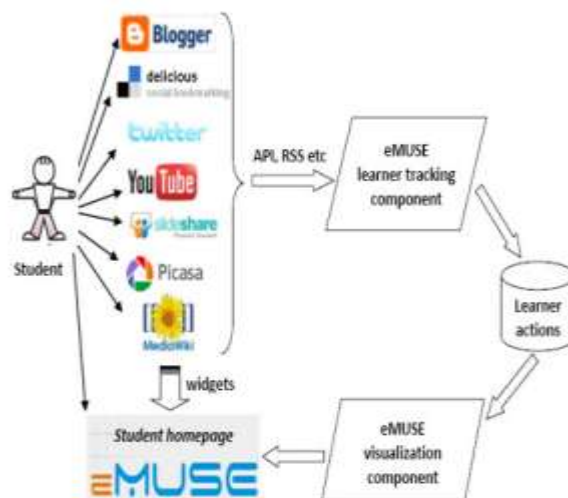


Figure 1: eMUSE schematic architecture - learner tracking module

3. CROWD BASED DESIGN ACTIVITIES

Crowdsourcing is an online, distributed problem-solving and production model. Crowdsourcing plays a powerful role in educational organization. Use of crowdsourcing can keep organizations forefront of education innovation, prepare students for online world challenges, and most importantly to stimulate closer communication between students, faculty and administration and beyond the campus with other institutions and groups. It is very helpful in solving on-campus problems that institutions face. Crowdsourcing can help find solutions for on-campus issues and provide fresh feedback or even ideas on how to solve problems differently. Crowdsourcing give students real world experience in coming up with creative solutions to important problems. Students can apply classroom knowledge to real world problems and learn the ins and outs of their chosen fields from a practical perspective.

4. EDUCATIONAL DATA MINING

Data mining can be defined as the process involved in extracting interesting, interpretable, useful and novel information from data. The objective of data mining in each application area is different. Educational Data Mining is the application of data mining techniques to educational data [2][3]. Its objective is to resolve educational research issues. The first phase of the EDM process is discovering relationships in data. Discovered relationships must then be validated in order to avoid overfitting [7]. Validated relationships are applied to make predictions about future events in the learning environment. Predictions are used to support decision-making processes and policy decisions.

5. SENTIMENT ANALYSIS

Social Networking Internet services are changing the way to communicate with others and entertain. The rapid growth in popularity of social networks has enabled large numbers of users to communicate, create and share content, give and receive recommendations. Social media is the social interaction among people in which they create, share or exchange information and ideas in virtual communities and networks. Social media sites such as Twitter, Facebook, MySpace and YouTube provide great venues for students to share their happiness, stress and and seek social support. On various social media sites, students discuss and share their everyday encounters in an informal and casual manner. Most existing studies found on social media data classification are either binary classification or multi-class classification.

Sentiment analysis is also known as opinion mining [6]. It is a three-class classification on positive, negative, or neutral opinions. Sentiment analysis is very useful for mining customer opinions on products. It can help marketers evaluate the success of an ad campaign or new product launch, determine which versions of a product or service are popular and identify which demographics like or dislike particular product features. Only knowing the sentiment of student-posted comments in social media does not provide much knowledge on relevant interventions and services for students. The purpose is to achieve deeper and finer understanding of student's experiences especially their learning-related issues and problems.

III CONCLUSION

This study is helpful to researchers in learning analytics, educational data mining, and learning technologies. Many studies have been conducted to determine the problems of students. But no research found to directly mine and analyze student-posted data from uncontrolled spaces such as social media. Sentiment analysis involves building a system to collect and examine opinions made in blog posts, comments, reviews or tweets. But it is not appropriate for deeply understanding the learning experience of students. Furthermore it is assumed that an appropriate multi label classification algorithm can provide better results.

IV ACKNOWLEDGMENT

The authors would like to thank the publishers and researchers for making their resources available. We also thank the college authority for providing the required infrastructure and support. Finally we would like to extend our heartfelt gratitude to friends and family members.

REFERENCES

- [1] 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.
- [2] 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.
- [3] C. Romero and S. Ventura, "Educational Data Mining: A Review of the State of the Art," *IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Rev.*, vol. 40, no. 6, pp. 601-618, Nov. 2010.
- [4] Elvira Popescu, Dan Cioiu, "eMUSE-Integrating Web 2.0 Tools in a Social Learning Environment", *Advances in Web Based Learning*, pp.41-50, Springer, 2011.
- [5] 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.
- [6] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs Up?: Sentiment Classification Using Machine Learning Techniques," *Proc. ACL-20Conf. Empirical Methods in Natural Language Processing*, vol. 10, pp. 79-86, 2002.
- [7] G. Siemens and R.S. d Baker, "Learning Analytics and Educational Data Mining: Towards Communication and Collaboration," *Proc. Second Int'l Conf. Learning Analytics and Knowledge*, pp. 252-254, 2012.