Research Article

A Study on the Impact of Social Media Site Using Mining Algorithm

D. Chinnapparaj, C. Sasikala and C. Jothi Venkateswaran

PG's & Research Department of Computer Science, Presidency College (Autonomous), Chennai - 600 005, Tamil Nadu, India.

Abstract

This is a study of social media sites use regarding students learning enhancement through social media. The objective of this research to analyze the impact of social media sites on education system from school to college. Using association rules and chi-square test were used to find students learning process associated with social media sites. The main purpose of this study found the relationship between social media sites and students learning activities. Result highlighted that frequent uses of social media sites for learning by the students have positive impact and strong relationship. Social media sites thus enhance students learning activity faster and easier. Association rule mining used for searching the relationship of social media sites on students learning activities performed better than chi-square test analysis.

Keywords: Social Media Sites (SMS), Association Rule Mining Algorithm (ARMA), Learning activities, Apriori, Chi-Square, Data Mining.

Introduction

Association rule mining algorithm is used to find relationship between social media and learning activities of students who are all using social media sites to enhance their education. Knowing the relationship between social media sites helps and plays an important role in most of the students daily life. Based on the apriori algorithm result obtained and statistical interpretations issues hypothesis to find the impact of social media sites use on students learning activities.

Social media sites have enormous study related materials in the forms of structured and unstructured collection of social education hyperlinked documents which makes faster searching relevant documents with useful study materials from lots of social media experts who have already associated with social educational websites. At any time, more than one social media sites may visited by many students for chatting, entertainment and learning purpose. One social media site can be accessed by a students many times for different purposes.

Data mining refers to extracting knowledge from the huge database which used to find the relationship of social media sites on students learning activities. The main objectives of these studies have been carried out to investigate the impact and use of social media sites in school and college students academically. Even though the literature covers a wide variety of such studies, the purpose of the study will focus on the following themes: 1) to understand the impact of social media sites use on students learning activities 2) to study the relationship between social media sites and students learning activities.

1.1 SOCIAL MEDIA SITES:

Social media sites are computer mediated tools that allows any user to create, exchange of user generated information, upload pictures and videos from expert communities over the networks.

There are lots of social media sites which are used in education related activities given below:

1.1.1 Facebook is one alternative means for shyer students to be able to voice their thoughts in and outside of the classroom. It allows students to collect their thoughts and articulate them in writing before committing to their expression (Moody, 2010).

1.1.2 Twitter- Domizi utilized twitter in a graduate seminar requiring students to post weekly tweets to extend classroom discussions. Students reportedly used twitter to connect with content and other students. Additionally, students found it "to be useful professionally and personally (Domizi, 2013).

1.1.3 YouTube is the most frequently used social media tool in the classroom. Students can watch videos, answer questions, and discuss content. Additionally, students can create videos to share with others (Moran *et al*., 2012). Sherer and Shea claimed that YouTube increased participation, personalization (customization), and productivity (Sherer and Shea, 2011). YouTube also improved students' digital skills and provided opportunity for peer learning and problem solving found that videos kept students' attention, generated interest in the subject, and clarified course content. Additionally, the students reported that the videos helped them recall information and visualize real world applications of course concepts (Eick and King, 2012).

1.1.4 Wikipedia - In early 2013, Steve Joordens, a professor at the University of Toronto, encouraged the 1,900 students enrolled in his introductory psychology course to add content to Wikipedia pages featuring content that related to the course.

2. REVIEW RELATED WORK

Chika Euphemia Asogwa *et al,* (2015) in his study paper named use of social networking sites and academic performance among students of selected tertiary institutions in kogi state clearly demonstrate the effect of facebook usage on the academic performance of students of tertiary institutions in kogi state. The results of this study to highlight the social networking sites have positive impact and useful instrument in improving academic performance of students of tertiary institutions in Kogi state. Vidhya Sadhu Kshirsagar and Sharmila Ajitkumar Kulkarni, (2013) in their study analyzed social media sites can be an used an effective tool for learning activities. Results indicated that college students using atleast one social media site for learning activity.

Rahul Mishra et al, paper titled comparative analysis of apriori algorithm and frequent pattern algorithm in web log data stated that weblog database give correct information about what kind of users will access which kind of web pages. Analyzing and exploring regularities in weblog records can identify potential customers for electronic commerce and also enhance and improve the quality of internet information services to end users (Mishra and Choubey, 2012).

Pappu Rajan et al, in their study paper named application of retail analytics using association rule mining in data mining techniques with respect to retail supermarket states that to identify each products category in which has associated with that products which are likely to buy with other products as a bundle of item sets. Using the association rule mining and market basket analysis, customer profitable associated product item sets are found. From this we can identify customer's pattern of buying behavior and it is then used to devise cross selling strategies to improve the sales of the products, increasing the profit of the retail outlet and retaining its customers (Pappu Rajan and Suresh, 2015).

Navdeep Kaur et al, analyzed that how pattern recognition in large database can be done and how they get benefit from a large number of companies by increasing their sales. If we look at the customer point of view, it is concluded that market basket analysis helps a buyer make better choice and easily search the items of his choice in a supermarket. Thus improving customer satisfaction and profit to the companies from this study (Navdeep Kaur and Prabhsimran Singh 2015).

3. SYSTEM ARCHITECTURE

This system architecture shows the relationship (association rule mining) between social media sites use and students learning activities. Use of social media sites depends upon the gender, age, marital status, education level, university name, type of institution, learning activities through social media sites. Hence, the students learning activities depend upon the use of social media sites have illustrated in this figure-1 and all the students are aware of social media sites usage.
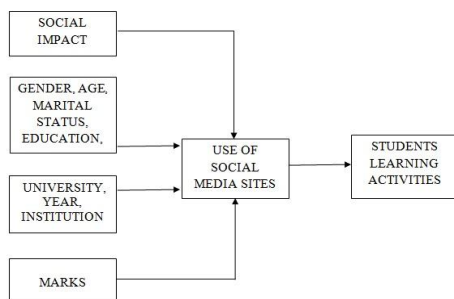

Figure 1: System architecture

4. ASSOCIATION RULE MINING ALGORITHM

To find all possible social rules from the given social transaction database and find all frequent item sets and then generate strong social association rules from the frequent item sets.
Support: The rule $1 \rightarrow 2$ holds with support s if s% of social transactions in D contains $1 \cap 2$. Support = (consist of item set combination) / (total number of record).
Confidence: The rule $1 \rightarrow 2$ holds with confidence c if c% of the social transactions in D that contain 1 also contains 2. Confidence $(1 \rightarrow 2)$ = support count $(1 \rightarrow 2)$ / support count (1).
Frequent Item set: If the given item sets which satisfies minimum support threshold, then it is called frequent item set.
Strong association rules: The rules which satisfy both minimum support threshold and minimum confidence threshold are called strong association rules mining.

4.1 APRIORI ALGORITHM:

Apriori algorithm is to mine association rules and uses breadth first search (level wise) to count the support of item sets and uses a candidate generation function which exploits the downward closure property of support (Patel, 2010).

Finding frequent itemsets using candidate generation

The Apriori algorithm for discovering frequent itemsets for mining Boolean association rules. Apriori employs an iterative approach known as a level wise search, where k-itemsets are used to explore (k+1) itemsets. First, the set of frequent 1-itemsets is found by scanning the database to accumulate the count for each item, and collecting those items that satisfy minimum support. The resulting set is denoted L1. Next L1 is used to find L2, the set of frequent 2-itemsets, which is used to find L3, and so on, until no more frequent k-itemsets can be found. The finding of each Lk requires one full scan of the database (Poonkuzhali and Saravanakumar, 2009).

The Apriori principle:
Any subset of a frequent itemset must be frequent.
The Apriori property consists of two step processes:
a. Join step: Candidate item set of size k (Ck) is obtained by joining frequent item sets of size k-1 (Lk-1) with itself.
b. Prune step: Any k-1 item set that is not frequent cannot be a subset of a frequent k-item set. Hence, it is discarded (Poonkuzhali and Saravanakumar, 2009).

4.2 Data Collection:

According to this research, primary data was used. The primary data collected from 405 students through a questionnaire which played as the main source of data for the analysis of the research and these collected data lead this research study to the exploration of the impact of social media sites on students learning activities. Random sampling method has been used for data collection from different Government and private of colleges.
4.3 Software Used: In this study Matlab software was used and data set add-ons in MS Office 2007. Using Excel it creates a connection to Matlab to find the relationship of social media sites use on students learning activities.

5. DATA ANALYSIS AND INTERPRETATION
5.1 Chi - square (χ^2) analysis:

Data collected were statistically analyzed using chi-square test (χ^2) = 13.42. The table value of chi square at alpha (α) = 0.01 and degrees of freedom = 6 is **16.812**. Which is greater than the calculated chi square test therefore accept the null hypothesis (H0). This means that the frequency uses of social media sites by the students have strong impact on their learning activities.
5.2 Apriori rule mining algorithm result:

The output of Apriori algorithm is given below in the following steps. To analyze these social transactions to calculate support count for each item table-1 shows generation and pruning of candidate 1-itemset from the social media sites dataset.
Step 1: Generation and pruning of candidate 1-itemset
From the above table-1 shows frequent item sets for Z= {1, 2, 3, 4,5,6,7,8,9,10,11,12,13}. A data set that contains k items can generate up to 2k-1 frequent item sets. To find frequent item sets to determine support count for every candidate item set, apriori approach is used (Tan *et al*., 2006).

Read the database to count the sub count of 1-itemset. Then the set of frequent 1-itemset L1 is formed after removing the itemset which is less than minimum support. From the above sub count items 9, 10, 11, 12, 13 are removed therefore L1 consists of {1,2,3,4,5,6,7,8} item set (Poonkuzhali and Saravanakumar, 2009).

Table-1: Generation and pruning of candidate 1-itemset

To discover the set of frequent item sets have relationship with social media sites, the algorithm uses the join L1 on L1 to generate a candidate set of 2-itemsets C2. So in the candidate generation step we get variables {2,3,4,5} which have strong relationship with social media sites, meanwhile the variables{6,7,8} have less relationship with social media sites.

Table-2: Relationship matrix to social media sites

FIND STRONG ASSOCIATION RULES FROM FREQUENT ITEMSETS:

Once the frequent item sets have found from social transaction in a database D have been found to generate strong association rule between social media sites and students learning activities from them which work based on the conditional probability as follows:

Confidence $(1 \rightarrow 2)$ =support count (1 U 2) / support count (1)

C3

3 - Itemsets	Confidence
1->2,3	87.28%
2->1,3	71.42%
3->1,2	99.75%
1,2->3	86.53%
1,3->2	99.26%
2,3->1	99.75%

Table - 3: Strong association rule mining

From the above table -3 the rules which satisfied minimum support threshold and minimum confidence threshold. Hence it is proved that social media sites have strong relationship with students learning activities.

CONCLUSION

Association rule mining algorithm has used for many purposes and studied how to find relationship between social media sites use and students learning activities. To measure the social rules there are two steps: support and confidence. According to this paper, we studied both steps using Apriori algorithm. The purpose of this analysis is to find how the relationship between social media sites use and students learning activities and details of students which retrieved from social transaction database and how students are improving their learning activities from social media sites. Apriori algorithm when use for searching the relationship of social media sites on students learning activities performed better than chi-square test analysis. If we see at the student point of view, it is concluded that social media sites helps a student positively and easily search the study related materials uploaded and shared from the experts from many social media sites which can be accessed faster than the books. Thus enhances students learning activities by using social media sites and guaranteed through this study.

References

Asogwa, C. E., Ojih, E. U. and Onoja I. B., 2015. Use of Social Networking Sites and Academic Performance among Students of Selected Tertiary Institutions in Kogi State", *International. J. Afric. Asi. Stud.,* www.iiste.org, (6): 46-57.

Domizi, D. P., 2013. "Microblogging to foster connections and community in a weekly graduate seminar course. *Tech. Trend.,* 57(1): 43-51.

Eick, C. J. and King, D. T., 2012. Non-science majors' perceptions on the use of YouTube video to support learning in an integrated science lecture. *J. Coll. Scien. Teach.,* 42(1), 26-30.

Patel, K. K., 2015. A Survey of Cyclic Association Rule. *Int. J. Engi. Develop. Res.,*3(1): 453-458.

Moody, M., 2010. Teaching Twitter and Beyond: Tip for Incorporating Social Media in Traditional Courses. *J. Mag. New Med. Res.,* 11(2): 1-9.

Moran, M., Seaman, J. and Tinti-Kane, H., 2012. How today's higher education faculty use social media (PDF). *Pearsonlearningsolutions.com*

Navdeep Kaur and Prabhsimran Singh 2015. A brief study of association rule mining. *Inter. J. Inform. Tech. Comp. Scien. Perspec. Pezzo. J.,* 4(1): 1388-1392.

Pappu Rajan, A. and Suresh S., 2015. Application of Retail Analytics Using Association Rule Mining In Data Mining Techniques With Respect To Retail Supermarket. *Intern. J. Engi. Manag. Res.* 5(1): 1-9.

Poonkuzhali, S. and Saravanakumar C., 2009. Data Warehousing & Data Mining. ISBN, 978-81-904915-0-1, pp 5.8.

Mishra, R. and Choubey A., 2012. Comparative Analysis of Apriori Algorithm and Frequent Pattern Algorithm for Frequent Pattern Mining in Web Log Data. *Intern. J. Comp. Scien. Inform. Technol.,* 3(4): 4662 –4665.

Sherer, P. and Shea, T., 2011. Using online video to support student learning and engagement. *Coll. Teach.,* 59(2): 56-59.

Tan, Ning, P. Michael, Steinbach, and Kumar, Vipin, 2006. *Introduction to data mining,* 6.

Vidhya, S., Kshirsagar, Sharmila, Ajitkumar and Kulkarni, A., 2013. A Study on Effects of Social Networking Sites as an Educational Tool on College Students. *Intern. J. Scien. Res.,* 4 (1): 1814-1816.