

WEB USAGE MINING: LEARNER CENTRIC APPROACH FOR E-BUSINESS APPLICATIONS

B. NAVEENA DEVI*

Abstract

Emerging of web has put forward a great deal of challenges to web researchers for web based information management and retrieval, especially online business competitiveness. Users find it very difficult to extract useful and relevant information from the huge amount of information particularly for business, education and e-learning. The problem can be solved by web usage mining with the experience using open web application programming interfaces(APIs) which are available by major companies for example RapidMiner, Digg.com, Amazon, eBay are opened access to their services and data through APIs, and we can make use of their services for the development of web usage mining research applications. This paper presents an experiment with pattern classification for user behavior prediction for an e-commerce business application. The paper concludes that web usage mining can be an approach to build innovative business applications and has potential to help improve learning performance with the help of APIs with in short period of time.

Keywords: Data Mining, Web Usage Mining, APIs, E-Learning.

* Department of CSE, Mahatma Gandhi Institute of Technology, Hyderabad, Andhra Pradesh 500075, India.

1. Introduction

The World Wide Web is a vast resource of multiple types of information in varied formats. Researchers are beginning to investigate human behavior in this distributed Web data warehouse and are trying to build models for understanding human behavior in virtual environments. Data mining, often called Web mining when applied to the Internet, is a process of extracting hidden predictive information and discovering meaningful patterns, profiles, and trends from large databases. Web research academia is requested to develop more efficient and effective techniques to satisfy the increasing demands of web users, such as retrieving the desirable and related information creating good quality web communities.

In order to effectively utilize the power of the web, information technology professionals need to have sufficient knowledge and experience in various web technologies and applications. In recent years courses in Internet and web related topics have been offered in many universities to equip students with such knowledge. However, to built web usage mining applications from raw data not an easy task that every student can complete in a semester. Now a days many major internet companies providing their services and data in the form of application programming interfaces. Such type of real time data provides ideal playground for students to gain practical skills in business application development and experimenting with the various pattern analysis techniques.

2. Potential of Web usage Mining

Web usage mining approaches can be applied to API-based learning and have promise for helping explain system usage. Web usage mining can be used to articulating and identifying patterns of use in learning systems may provide better understanding of how students undertake Web-based learning and guidance for better understanding of online business activities. There are many potential benefits of using Web mining for exploring learning behavior and patterns in API-learning. However, there is scant literature on the use of Web mining with APIs. This article is a case report of how Web usage mining method, classification could be applied on APIs data sets. This case report shows how patterns emerge from the application of Web mining approaches. The key purpose of this report is to illustrate the potential of Web usage mining and to identify issues

in its application with currently available APIs. As a way of setting the context for the case report the following examples show how Web usage mining could benefit learning process.

2.1. Understand learner or customer behavior

University administrators and instructors may be able to improve the implementation of business design systems by understanding the dynamic behavior of customers .

2.2.Determine e-learning or e-commerce system effectiveness

Patterns of behavior may be associated with system performance and enable more customized system configuration. Administrators and instructors may be able to discover high and low use areas of the business system and adjust resources to optimize the technical performance of the system.

2.3.Measure the success of instructional efforts

In API-learning systems, students use email, the Web forum, feedback forms, etc. to express their concerns and ask questions. These data are completely recorded in the learning system. Web usage mining can provide quantitative feedback to instructors about the outcomes of their activity.

The above features show how Web usage mining could provide new insights about student activity in API based learning and to address information needs as well as suggest customization of approaches for implementing e-learning, e-business, education by APIs with experience of real time data sets broad and rich benefits may accrue from better understanding patterns of behavior.

Table 1. Comparison of web usage mining process questions in business, education and e-learning.

Business	Education	E-learning
Who are my most profitable customers?	Who are the students taking most credit hours?	Who are the students with highest frequency of logging-in?

Who are my repeat Website visitors?	Who are the ones likely to return for more classes?	Who are the students most engaged on discussion boards?
Who are my loyal customers?	Who are the persisters at our university, college?	What pages do the students access most?
What clients are likely to defect to my rivals?	What type of courses can we offer to attract more students?	What kinds of students are likely to get a high score online?

3. Technical Preparation to Integrate APIs

Recently there are a large number of web services that we can use and many of them are open source based. Web services are APIs that facilitate the communication between applications for example RapidMiner, Digg.com, Amazon, eBay are opened access to their services and data through APIs, and we can make use of their services for the development of web usage mining research applications. The concept of Web APIs enables direct access to the website functionalities in order to leverage third party efforts on value adding services. However, the number of companies, services or web sites that gather information about users increasing continuously.

In recent years, due to the rapid development of web, more techniques related to internet and web applications have been introduced in many universities around the world to better equip students to effectively utilize the power of the web.

```
<!DOCTYPE html-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head><meta http-equiv="content-type"content="text/html;charset=utf8"/>
<script Src="httpmaps.google.com/maps?file=api&v=2&key=abcdefg" type="text/javascript"></script> <script type="text/javascript">
Function initialize()
{
    If(GBrowserIsCompatible())
    {
        Var map = new GMap2(document.getElementById ("map_canvas"));
    }
}
</script> </head> <body onload="initialize()"onunload="GUNload()">
<div id="map_canvas"style="width:500px;height:300px"></div>
</body> </html>
```

This example can be edited and play around with it, but we have to replace the key in the file with our own maps API key. Once register a key for a particular directory it works for all sub directories as well. Technical Preparation of developing e-business application using Open Web APIs: The Amazon e-commerce service example was implemented based on REST protocol. In order to illustrate the e-business application to integrate the web APIs into a specific website to perform selected business operations, depending on the system functionality. Many systems may need to store data such as Oracle, Microsoft SQL server or My SQL. This application is a Business Model constructed by using Amazon e-commerce service API for item listing data in the cellular phone category of auctions. The application then analyzes the captured data for market research and analysis purposes. Application provides tools such as statistical analysis, data analysis, and data visualization to aid sellers and buyers in achieving their business goals. We mined the Amazon e-commerce service API for all necessary product information, the use of the Amazon e-commerce service API is most crucial to the success of our Business Model. Through the Amazon e-commerce service API the data obtain information such as item ,price, quantity ,time and date of purchase.

4. Research Goals

Given the limited use of Web mining in education and the potential benefits to online education of making it more student centric that might emerge from effective utilization of Web mining, we set two research goals for this project: (1) to see if patterns of behavior could be used to predict achievement in effective learning in applying various pattern analysis techniques in a set of API data, and (2) to develop a better understanding about the process of applying Web mining to business applications, as well as to minimize preprocessing work with existing current versions of real time data sets. While Web usage mining is a recognized approach for building knowledge and value in business and commercial information systems. In this sense, this research is primarily exploratory and while the objective is to build new insights about learning activity, an equally important objective is to examine the fit of Web mining approaches to e-commerce. What are the challenges in extracting data from APIs and applying Web mining approaches? What strategies for data formatting and data analysis are needed for building meaningful insights? What changes are needed in Web usage mining solutions to improve the yield of Web mining in E-

commerce systems? A key outcome of this research will be suggestions for understanding how best to utilize open resources for Web mining in business applications and as well as for education.

5. Process of Mining

There are various data mining techniques such as statistics, classification, association rules, sequential patterns, and clustering which can apply to the Web domain. Classification is the form of data mining used in this study and is a technique that uses a set of pre-classified examples to develop a model that can classify the products of records. There are many algorithms for classification such as decision tree, neural network classification, etc. The classification algorithm starts with a training set of pre-defined example transactions. The classifier training algorithm uses these pre-defined examples to determine the set of parameters required for proper discrimination. The algorithm encodes these parameters into a model called a classifier. After an effective classifier is developed, it will be used in a predictive mode to classify new records into these same pre-defined classes. For instance, one classifier that is capable of identifying customer behavior performance could be used to help in the decision of whether to provide a specific recommendation to an individual customer. In this study, the decision tree software C4.5 (Quinlan, 1993) is used, which is shown in detail in the case report. C4.5 is an algorithm introduced by Quinlan for inducing Decision Trees from data.

6. Results and Discussion

In this research, mobile information retrieved from Amazon e-commerce API. A binary decision tree was built to classify the access log file. The log file contains name of the product, visited time, price, and quantity. The log data were examined to identify which product having no demand by analyzing behavior of customer. This could predict grades.

Select an item using the following drop-downs:		
Item	Price	Quantity
itemtitle: unlocked Quad-Band Dual SIM TV WiFi Mobile Phone F006 startprice: 135.0000 quantity: 1000 currency: 135.0000 country: HR	itemtitle: unlocked Quad-Band Dual SIM TV WiFi Mobile Phone F006 startprice: 135.0000 quantity: 2500 currency: 135.0000 country: HR	itemtitle: unlocked Quad-Band Dual SIM TV WiFi Mobile Phone F006 startprice: 135.0000 quantity: 4000 currency: 135.0000 country: HR
itemtitle: unlocked Quad-Band Dual SIM TV WiFi Mobile Phone F006 startprice: 135.0000 quantity: 1500 currency: 135.0000 country: HR	itemtitle: unlocked Quad-Band Dual SIM TV WiFi Mobile Phone F006 startprice: 135.0000 quantity: 3000 currency: 135.0000 country: HR	itemtitle: unlocked Quad-Band Dual SIM TV WiFi Mobile Phone F006 startprice: 135.0000 quantity: 4500 currency: 135.0000 country: HR
itemtitle: unlocked Quad-Band Dual SIM TV WiFi Mobile Phone F006 startprice: 135.0000 quantity: 2000 currency: 135.0000 country: HR	itemtitle: unlocked Quad-Band Dual SIM TV WiFi Mobile Phone F006 startprice: 135.0000 quantity: 3500 currency: 135.0000 country: HR	itemtitle: unlocked Quad-Band Dual SIM TV WiFi Mobile Phone F006 startprice: 135.0000 quantity: 500 currency: 135.0000 country: HR

Table.1. Sample log file from Amazon e-commerce API

Table.1. shows browsed information. This kind of information could predict the users next action and offer personalized website content and service based on corresponding forecast.

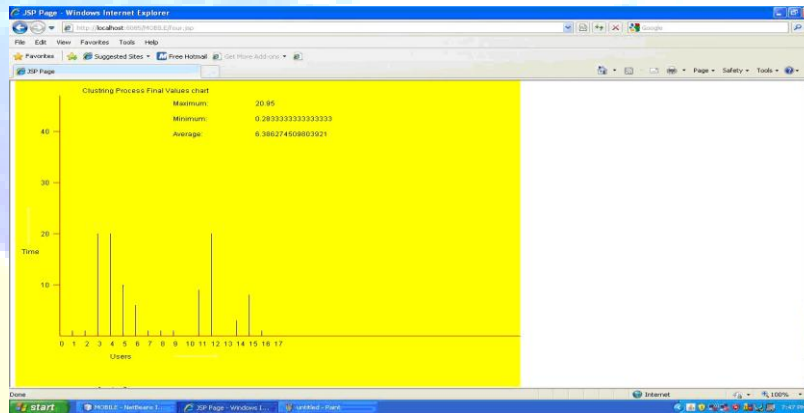


Fig.1. Graph showing the Frequency of browsing.

Fig.1 showing the Frequency of browsing by the customers. It will help vendor hypothetically a new or customized recommendation could be made for future demand.

7. Conclusion

In this research, we explained the use of Web usage mining approaches using APIs and identified some illustrative learning patterns that can be found by using Web-mining approaches. Although some interesting patterns were found, the exploratory state of Web Usage mining tools in education suggests replication and confirmation from other forms of research to build an innovative business application. The primary findings of this research are to suggest that Web usage mining can be an approach that educational researchers can use, and when combined with other forms of data collection has potential for adding to the way we build knowledge about pattern analysis. A second contribution of the current study is to draw implications for how to improve the process of Web Usage mining with API services.

The current research has shown a innovative business web usage mining application by utilizing Amazon API with minimum time by avoiding preprocessing phase and main concentration on analysis has promise for identifying patterns within the large processes. This achievement might be improved by availing various APIs and by introducing new mining methods. These results and our ability to use Web usage mining in e-learning and e-commerce are quite preliminary, and there is a need for further exploration and possible adaptation of the forms and usage of web mining to best suit education and business.

References

- [1] Buchner, A. and Mulvenna, M.D. (1998). Discovering Internet Marketing Intelligence through Online Analytical Web Usage Mining. SIGMOD Record, 27(4):54-61
- [2] Chakrabarti, S., Dom, B., and Indyk P. (1998). Enhanced hypertext categorization using hyperlinks. In Proc. of the ACM SIGMOD Conference on Management of Data, Seattle, Washington: ACM Press, pp. 307—318
- [3] Cooley, R. Mobasher, B. Srivastava, J. (1997). Web Mining: Information and Pattern Discovery on the World Wide Web. In Proc. of IEEE International Conference on Tools with Artificial Intelligence, Newport Beach, CA, pp558~567
- [4] Etzioni, O. (1996).The World-Wide Web: Quagmire or Goldmine? Communications of the ACM, vol. 39, no. 11, pp65-68.
- [5] Luan, J. (2002). Data Mining and Its Applications in Higher Education. New Directions for Institutional Research, 2002, 113 (Spring), pp17-36.
- [6] Padmanabhan, B. and Tuzhilin, A. (1998). A Belief-Driven Method for Discovering Unexpected Patterns. In Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98). AAAI, Newport Beach, California, USA, pp94--100.
- [7] Quinlan, J. (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA.