ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

User Behavior Analysis Based On Predictive Recommendation System for E-Learning Portal

Toshi Sharma
Department of CSE
Truba College of Engineering & Technology
Indore, India toshishm.25@gmail.com

Ritu Tondon
Department of CSE
Truba College of Engineering & Technology
Indore, India bansalritu10@gmail.com

Abstract

Web page access prediction has increased its importance in the field of e-commerce, e-learning systems and e-businesses. And involves personalizing the Web users' browsing experiences, helps in the improvement of the Web site structure and Web users in navigating the site and accessing the information they need. As the information grows on the Web, the need of recommendation systems to ease user navigations becomes important. In general term, if any system can predict what a user look for in web domain then this kind of system is known as Recommendation System. As we know that the Web is becoming a vast and preferred source of information retrieval.[1] It is required that the retrieved data is of users interest and should be related to what the user have queried. So there are many recommendation systems that suggest the items of their interest. Such kind of system design is known as collaborative filtering technique of recommendation system design. In order to implement the desired system a clustering technique and a predictive algorithm is required. Thus K-mean clustering algorithm and Hidden Markov Model is employed. Implementing this helps to predict the next page to be accessed by the Web user based on the user's previous browsing patterns i.e. to know the users behaviour pattern to increase accuracy and decrease space complexity.

Keywords – ***Recommendation System, Clustering, Hidden Markov Model(HMM)***

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

I. Introduction

With the rapid growth of World Wide Web, Web mining has been of great importance to many researches. Web mining is the use of data mining techniques to automatically discover and extract information from Web documents and services. Web mining is mainly classified as **Web structure mining, Web content mining** and **Web usage mining**. Web usage mining finds user access patterns from Web servers and tries to discover valuable information from users' transactional data.

As a result, prediction of users requested pages is an important issue for personalization applications. A Web personalization system is defined as any system that tailors the Web experience for a particular user/a group of users.[2]

As this web Recommendation System is very popular these days in the field of E-Commerce and E-Businesses, Consumers today in the world with the internet and its associated information explosion are faced with the problem of too much choice. Right from looking for a restaurant to looking for good investment options, there is too much information available. To help the consumers cope with this information explosion, companies have deployed recommendation systems to guide the consumer. The research in the area of recommendation systems has been going on for several decades now, but the interest still remains high because of the abundance of practical applications and the problem rich domain. A number of such online recommendation systems are implemented and are in use such as the recommendation system for books at Amazon.com and Libra, for movies at MovieLens.org, CDs at CDNow.com (from Amazon.com), etc.

Web prediction is a problem of classification in which we have to predict the further sets of Web pages that a user may visit based on the knowledge of the previously visited pages. User's behaviour prediction can be applied in various applications in the internet environment. Therefore, web usage mining techniques are used to analyze the web usage patterns for a web site. The web access log of the users is used to fetch the user access patterns, thus are used in the prediction process. [4]

According to the previously visited categorized Web pages, the interests and tastes of browsers

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

are captured. Thus, the search engine can utilize a prediction model for a certain Web site to cache the next set of web pages that might be visited by the users. In Web prediction, challenges are to be faced in both pre-processing and prediction. Pre-processing challenges include handling large amount of data that cannot be accommodated in the computer memory, to choose optimum sliding window size, to identify sessions of the websites, and seeking/extracting domain knowledge. Prediction challenges includes long training/ prediction time, low prediction accuracy, and memory limitation.[4]

II. Background Research

Recommender systems or **recommendation systems** are a subtype of information filtering system that seek to predict the 'rating' or 'preference' that user would give to an item. We can also define Recommendation systems as special types of expert systems in the sense that they combine the knowledge of the expert in a given domain (for the product type being recommended) with the user's preferences to filter the available information and provide the user with the most relevant information. Personalization of the recommendations works by filtering a candidate set of items (such as products or web pages) through some representation of a personal profile. Two main paradigms for the filtering are **content-based approach** and **collaborative approach**. Most recommendation systems use a **hybrid approach**, which is a combination of these two approaches. A content based recommendation system recommends items that are considered sufficiently similar
to the content descriptions in the user profile. Collaborative filtering systems also referred as to match the rating of a current user for items with those of similar taste of users in order to produce recommendations for items not yet rated or seen.

Recommender systems can use data mining techniques for making recommendations using knowledge learnt from the action and attributes of users [8]. The objective of data mining is to discover interesting and useful knowledge using a variety of techniques such as prediction, classification, clustering, association rule mining and sequential pattern discovery. Currently, an increasing interest is seen in data mining and educational systems, making educational data mining new and growing research community [6][7]. The data mining approach to personalization uses all the available information about users/students on the web site (in the web course) in order to learn user models and to use these models for personalization. These systems can use different recommendation techniques in order to suggest online learning activities or

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

optimal browsing ways to students, based on their preferences, knowledge and the browsing history of other students with similar characteristics.[5]

III. Related Work

Nachiketa Sahoo et al. present a method to make personalized recommendations when user preferences change over time. When the user is observed over a long period of time a dataset on employees' blog reading behaviour is collected, which show that users' product selection behaviours change over time. So proposed a hidden Markov model to correctly interpret the users' product selection behaviours and make personalized recommendations. He evaluates the model using three real world datasets that includes data on employee blog reading behaviour inside a firm, users' music listening behaviour collected through last.fm and users' movie rating behaviour at Netflix with different characteristics. Then comparison of the recommendation performance of the proposed model with that of a number of collaborative filtering algorithms and proposed prediction algorithm is done. He found that the proposed HMM based collaborative filter performs as well as the best among the alternative algorithms when the data is sparse or static. [3]

Paula Cristina Vaz and David Martins De Matos presented a Hybrid Recommendation System for improving literary book recommendation system through author ranking. Here books are recommended for the book readers to decide which book to read next. It uses two item based collaborative filtering algorithms to predict books and authors that users will like. Author prediction is used to recommend the books.[11]

Wahidah Hussain have given a Personalized Location Based Traveller Recommendation System that gives information to the travellers about the local area attraction such as local food, shopping spots etc.[12]

A real Time Route Recommendation system is presented by Henan Wang and Guoliang Li. This system provides real time traffic aware routes that overcome the problem of recommending same routes for all users and do not take full advantage of real time traffic.[13]

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

IV. Recommendations with Different Algorithms

Clustering is a process of grouping objects into classes of similar objects. It is an unsupervised classification or partitioning of patterns (observations, feature vectors or data items) into groups or subsets (clusters). The principle of clustering is maximizing the similarity inside an object group and minimizing the similarity between the object groups. There are many clustering methods [9], including hierarchical and function based algorithms.

K-means algorithm is one of the most well-known and commonly used algorithms that tries to minimize the distance of the objects to the centroid or mean point of each cluster. To recommend items (pages) to the users in simple k-means algorithm, firstly we evaluate the best cluster for each evaluation data point (session) by calculating the distance between these data points with cluster centres. After that, we sort the pages of the best cluster based on the sum of times of user views on those pages. The most important pages of each cluster are pages in which this sum is maximized. We recommend most important pages of the assigned cluster which user has not seen yet.

Hidden Markov Model (HMM) is a statistical model which the system modelling in it, is considered as a Markov process and the challenge is to estimate the model parameters based on observable states. Model can be used in secondary analyses after estimating these parameters. This model is a simple model of Bayesian networks. In the regular Markov state, all the states are observable and the parameters are only probabilities of state changes. In hidden Markov model, states are not observable but we can observe the parameters affected by states. These types of models can be applied to temporal pattern recognition, bioinformatics and sound recognition. We can see a global architecture of Markov model when running, which is called Trellis, in Figure 1. Each oval shows a variable which can take some values; $x(t)$ is the hidden state in time t, $y(t)$ is the observation in time t and the arrows show the conditional dependence of variables. As we can see, state at time t just depends on the data at time $t-1$ and the observed value depends on the state at the same time. An HMM cannot be observed directly but can only be viewed through another set of stochastic processes that produce a set of observations. One of the simplifying assumptions of the HMM is that the observed variable in a given time period is assumed to only depend on the value of the hidden variable in that time period. HMMs have been widely applied in speech recognition, cryptanalysis, part-of-speech tagging, machine

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

translation, gene finding, and alignment of bio-sequences, software developer learning, and customer relationship management.

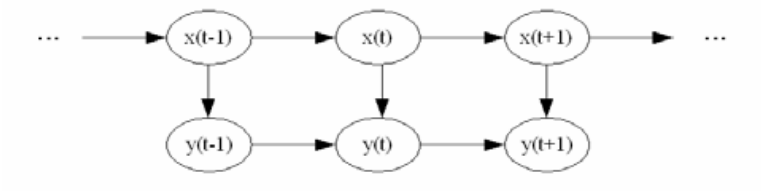

Fig 1. A global architecture of hidden Markov model while running.

V. Experimental View

In this study, we describe a personalized recommender system that uses web mining techniques for recommending a student which (next) links to visit within an adaptable educational hypermedia system. To prepare this system we use web access log data and clustering the documents topic wise by adopting K-Mean clustering technique. These resultant clusters are then fed to Hidden Markov Model (HMM) to find probability distribution, using similar properties and characteristics of user group and also the probability distribution of documents according to the topics as per users interest.

The processes of the given system can be summarized using the below given steps:

Input : D[I, j];
Output : Url U;

1. read data set D

2. D[I, k]=Preprocess(D); // number of columns are reduced in processing the data therefore the reduced columns are given using K.

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

3. store(D[I,k]);

4. CLData[User,D[n,m]] = kmeanclustering(D[I,k]);

5. Pdata[p,q]= find(User, CLData[User,D[n,m]]);

6. T = createTransition(Pdata);

7. O = createObservation(Pdata);
8. return U = (π, T,O)

VI. Result Analysis

This section provides the results of the developed recommendation system by analysing accuracy, error rate, learning time and decision time. Thus, according to the proposed system we are using the data logs as the instances for the above named analysis.

For the proposed system the accuracy is a measurement of prediction during the recommendation of the e-learning material. It (in fig. 2) is directly proportional to instances i.e. as instances increases accuracy also increases.

Error rate for this system provides the information how far the accurate prediction is, thus that can be calculated using the incorrectly identified instances during classification (in fig. 3) . That means whenever the accuracy increases error rate decreases also it can be defined as the instances maximize the error rate gradually minimizes.

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

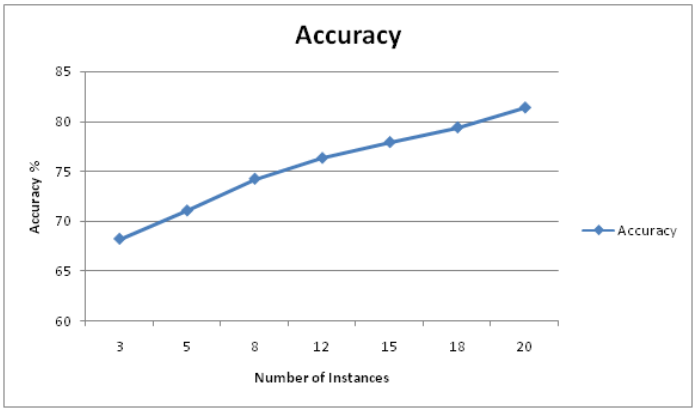

Fig. 2 Accuracy

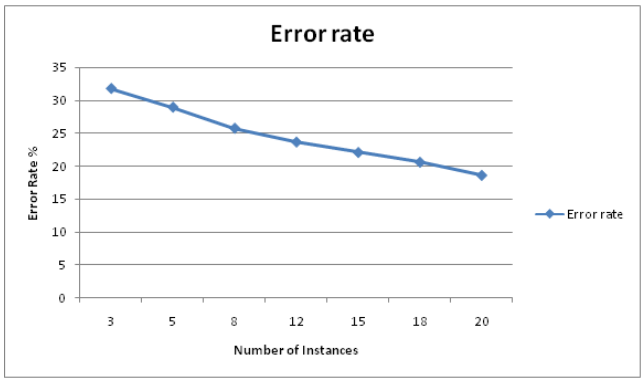

Fig. 3 Error Rate

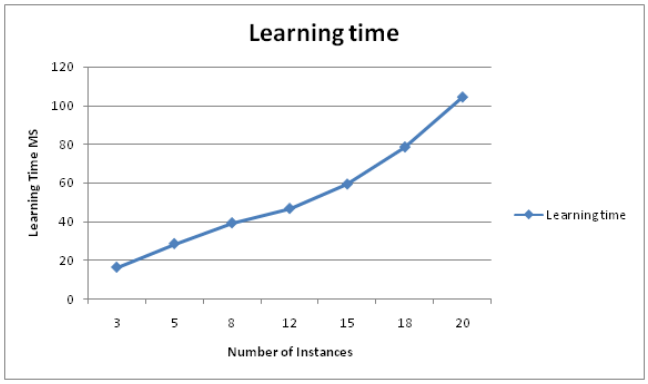

Fig. 4 Learning Time

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

Here learning time is the time required to learn the patterns from the raw data given above in (fig. 4). The evaluation of the learning time is performed using by finding the time difference between initialization of learning and completion of the learning.

Finally the decision time (in fig. 5) is calculated by analysing the amount of time for predicting a URL for the system. The decision time is not much affected by the instances for learning and time required learning.

Fig. 5 Decision Time

VII. Conclusion & Future Work

As we know that the Web is becoming a vast and preferred source of information retrieval. It is required that the retrieved data is of users interest and should be related to what the user have queried. So there are many recommendation systems that suggest the items of their interest. But a new Recommendation system is designed for the students which will help them in searching their required documents topic wise. This can be done by approaching a new design of a system which will use the clustering technique and HMM that will cluster the documents and student user in separate groups with respect to their similarities.

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

VIII. Acknowledgement

With due respect we would like to inform, that the article which we had proposed is prepared for academic scenario, that is not feasible for industrial research purpose the references which we had taken from listed here, one or not only references there are many other references which we had taken from real world and from daily life aspects. So it is vary through to remember all those references.

References

[1] Toshi Sharma & Prof. Ritu Tondon, "A Survey on Recommendation System Design Techniques", National Conference on Advances in Computer Science & Technology, ACST-2014.

[2] Shaghayegh Sahebi, Farhad Oroumchian & Ramtin Khosravi, "Applying and comparing Hidden Markov Model and Fuzzy Clustering Algorithms to Web usage data for Recommender

System".

[3] Nachiketa Sahoo, Param Vir Singh & Tridas Mukhopadhyay, "A Hidden Markov Model for Collaborative Filtering"

[4] Sampath P., and Ramya D "Performance Analysis Of Web Page Prediction With Markov Model, Association Rule Mining (Arm) And Association Rule Mining With Statistical Features (Arm-Sf)"2013

[5] Cristóbal Romero1, Sebastián Ventura1, Jose Antonio Delgado1, Paul De Bra, "Personalized Links Recommendation Based on Data Mining in Adaptive Educational Hypermedia Systems"

[6] Romero, C., Ventura, S.: Data mining in e-learning. Wit Press (2006).

[7] Romero, C., Ventura, S., Bra, P.D.: Knowledge discovery with genetic programming for providing feedback to courseware author. User Modeling and User-Adapted Interaction: The Journal of Personalization Research, 14:5 (2004) 425–464.

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

[8] Schafer, J.B.: The application of data-mining to recommender systems. In J. Wang (Ed.), Encyclopaedia of data warehousing and mining. Hershey, PA. Idea Group (2005) 44-48.

[9] Jain, A.K., Murty M.N., and Flynn P.J.: Data Clustering: A Review, ACM Computing Surveys, 31:3 (1999) 264-323.

[10] Alan Said, Domonkos Tikk, Andreas Hotho, The Challenge of Recommender Systems Challenges, Copyright is held by the author/owner(s). RecSys'12, September 9–13, 2012, Dublin, Ireland. ACM 978-1-4503-1270-7/12/09

[11] Paula Cristina Vaz, David Martins de Matos, Bruno Martins, PavelCalado, "Improving an Hybrid Literary Book Recommendation System through Author Ranking", Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries, Pages 387-388, ACM New York, NY, USA ©2012

[12] Wahidah Husain and Lam YihDih, "A Framework of a Personalized Location-based Traveller Recommendation System in Mobile Application", International Journal of Multimedia and Ubiquitous Engineering Vol. 7, No. 3, July, 2012