

**INTELLIGENT RECOMMENDER SYSTEM FOR STUDENT USING BAYES  
CLASSIFICATION ALGORITHM**

Miss Nayan N. Khandekar

Computer Science and Engineering, HVPM COET Amravati

**Abstract** —With the advanced development of resource technology, more and more digital resources are available in Internet. To maintain college study material repository online as i propose a “Intelligent Recommender system for student using Bayes Classification Algorithms”. this system provides document classification as well as recommender system for college students. It helps staff to upload study material and students to get the study material by convenient way and allow to provide rating for every study material to re-rank the links of study material uploaded by the staff. Re-ranking concept increases the interest of the student as he will get what he wants easily and overcome the resource restriction problem by providing additional live recommender system to student by using educational search engine.

**Keywords-** Document classification; Recommender system; Re-ranking concept; Educational search engine.

**I. INTRODUCTION**

Today's generation is a digital generation, every thing is going online. To maintain college study material repository online we propose a document classification as well as recommender system for college students. System proposed different modules like Admin panel, Staff, classification Analysis, Student , Recommender system.

In our system, staff members will upload any document on server. The document will get classified first by Rake Algorithm and then this output will get provided by weka tool using bayes document classification techniques and stored on server, The uploaded documents will be given to students as per their profile along with this we propose recommender system which recommend rating and liking wise study materials to student.

In case if the study material is not available for any subject, our system will recommend live data to the student by using educational search engine enabled.

**II. CLASSIFICATION TECHNIQUE**

Classification predicts categorical labels (classes), prediction model continuous-valued functions. In' Intelligent Recommender system for student using Bayes Classification Algorithm', the classification techniques are defined as, first RAKE Algorithm is used. when any type of subject related document is uploaded in this study portal so this document can be classified by RAKE Algorithm as, it will abstracted the keyword Phrases in that document. And this will be the first step in this classification techniques.

After that first step of classification, this document can further classified by the WEKA Classification technique. in that weka classification I used Bayes classification Algorithm. also in that Bayes classification for more efficient result I used Naive Bayes. the RAKE Algorithm output can be passed by the WEKA classifier-er techniques can classifying document page by page format. so that time reduction of producing result is puts an advantages over that system

In Weka tool the .Bayes classification Algorithm is worked as,

- Structure learning of Bayesian networks using various hill climbing (K2, B, etc) and general purpose (simulated annealing, tabu search) algorithms.
- Local score metrics implemented; Bayes, BDe, MDL, entropy, AIC.
- Global score metrics implemented; leave one out cv, k-fold cv and cumulative cv.
- Conditional independence based causal recovery algorithm available.
- Parameter estimation using direct estimates and Bayesian model averaging.
- GUI for easy inspection of Bayesian networks.
- Part of Weka allowing systematic experiments to compare Bayes net performance with general purpose classifiers like C4.5, nearest neighbor, support vector, etc.
- Source code available under GPL<sup>1</sup> allows for integration in other open-source systems and makes it easy to extend.

$$U = \{x_1, \dots, x_n\} \quad n \geq 1$$

- Let  $B_S$  be a set of variables. A Bayesian network  $B$  over a set of variables  $U$  is a network structure, which is a directed acyclic graph (DAG) over  $U$  and a set of probability

tables  $B_P = \{p(u|pa(u)) | u \in U\}$  where  $pa(u)$  is the set of parents of  $u$  in  $B_S$ . A Bayesian network represents a probability distributions  $P(U) = \prod_{u \in U} p(u|pa(u))$ .

network represents a probability distributions

- Where as Naive Bayes is work as following way,

Naive Bayes classifiers are a collection of classification algorithms based on **Bayes' Theorem**. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e. every pair of features being classified is independent of each other.

- In spite of their apparently over-simplified assumptions, naive Bayes classifiers have worked quite well in many real-world situations, famously document classification and spam filtering. They require a small amount of training data to estimate the necessary parameters.
- Naive Bayes learners and classifiers can be extremely fast compared to more sophisticated methods. The decoupling of the class conditional feature distributions means that each distribution can be independently estimated as a one dimensional distribution. This in turn helps to alleviate problems stemming from the curse of dimensionality.

In Intelligent Recommender system for student using Bayes Classification Algorithm, in student section I provided 'Educational search engine'. behind that search engine NLP Algorithm is used. This NLP Algorithm is used to identifying the fake keyword. when any student trying to search rather than the educational related keyword/data so in that case this NLP Algorithm is identifying that keyword as fake. And not producing any search related information. So by using that classifications techniques this puts good advantages over that system.

### III. LITERATURE SURVEY

With the advanced development of resource technology, more and more digital resources are available in Internet. The DL is introduced to provide readers with information and knowledge service at anytime from anywhere. Also, with the extension of network applications, more and more information is delivered and shared in the digital world. Digital Library (DL) has gradually been viewed as one of the most important digital information and knowledge resources because it could easily represent a variety of digitalized contents with text, image, video and audio without considering time and location limitation. Therefore, DL has become a convenient and successful solution for information services. It replaces large part of services with the capability of information delivery via Internet.

Many intelligent mechanisms applied in DLs have been proposed, which could help users retrieve information that are required with less time [2,3,4]. For example, Song et al. (2007) [5] suggested a document automatic classification system with an intelligent agent. To support on-line users to conveniently browse and search news, a hierarchical news map is offered in an automatic generation system [6]. In particular, the map gradually becomes one of the main trends that provide users with an integrating mechanism to efficiently gather different kinds of knowledge contents from different resources in DLs.

On the other hand, as the volume of digital materials and information sources are getting larger and larger, the DL has to move from being passive with little personalization for users to being proactive with tailored information for individual users. Personalization can help satisfy individual's need by understanding their preference. It has gradually become one of the important ways to improve the service quality of DLs [1,7,8]. Research literature indicated that personalization could be achieved by the user-guided approach (called adaptable) or automatic approach (called adaptive) [1,7,9,10]. The former indicates that the personalized pattern is directly created by the information provided by each user. For example, the My Yahoo! and MyLibrary are introduced with adaptable personalization.

### IV. ANALYSIS OF PROBLEMS

In existing literature, personalized study material recommendation is proposed. To recommend study material, user's personal information is used, Existing system only recommends the study material in any order. The concept of re-ranking is not discussed. In existing system, the resources are restricted and dependent on administrator.

To overcome these drawbacks, we propose a new re-ranking technique in which system will use ratings given by x students/other students to re-rank the links of study materials. Re-ranking concept increases the interest of the student as he will get what he wants easily. We also overcome the resource restriction problem as we are providing additional live recommender system to student.

### V. PROPOSED WORK

Following are the Modules provided by the system:

#### 5.1. Admin Panel

To Manage the system process like Register staff, Create staff log in, View staff details, Approve pending student registrations, View students, Register branches, Register branch wise subjects, Allocate subjects to staff.

## 5.2. Staff

Staff will do Log in, View allotted subjects, Upload documents, Delete document, Approve pending student registrations, Communicate with student, Communicate with other staff member, manage keyword in document classification.

## 5.3. Document classification

The document uploaded by the staff member, will be processed by first RAKE Algorithm and then passes to the weka tool using bayes document classification techniques.

## 5.4. Student

Student will do Registration, Log in, View recommended notes/study material, View live recommended links for every subject, Search any educational study materials if necessary. View own search history, Communicate with staff.

## 5.5 Recommender system

System propose two types of recommendation techniques

### 5.5.1. Study material recommendation

The study material uploaded by the guide will recommend to student as per his profile.

### 5.5.2. Live study material recommendation

The relevant study materials will be extracted from Google using Google API.

## VI. RESULT

This section presents the detailed information about the classification techniques in Intelligent Recommender system for student using Bayes Classification Algorithm in order to demonstrate the complete process. firstly this is the home page which takes the input data set from user.

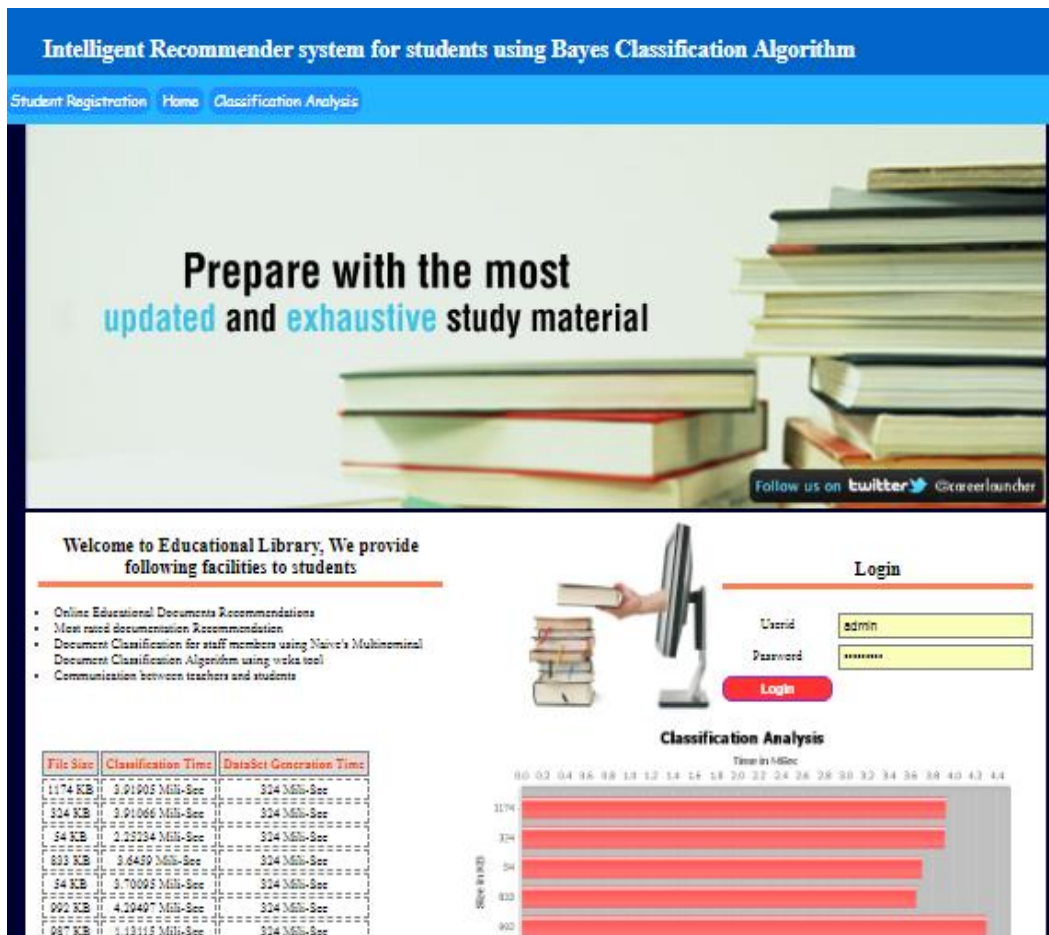


Figure 1. Home Page

### 6.1. Admin panel

In this admin panel is having following type of result is generating according to working of the admin.



*Figure 2. Admin Panel*

### 6.2. Staff

According to the staff working, the result is produces as following ways.



*Figure 3. Staff Panel*



### 6.3. Student panel

According to the working of the student this student panel is shown below.

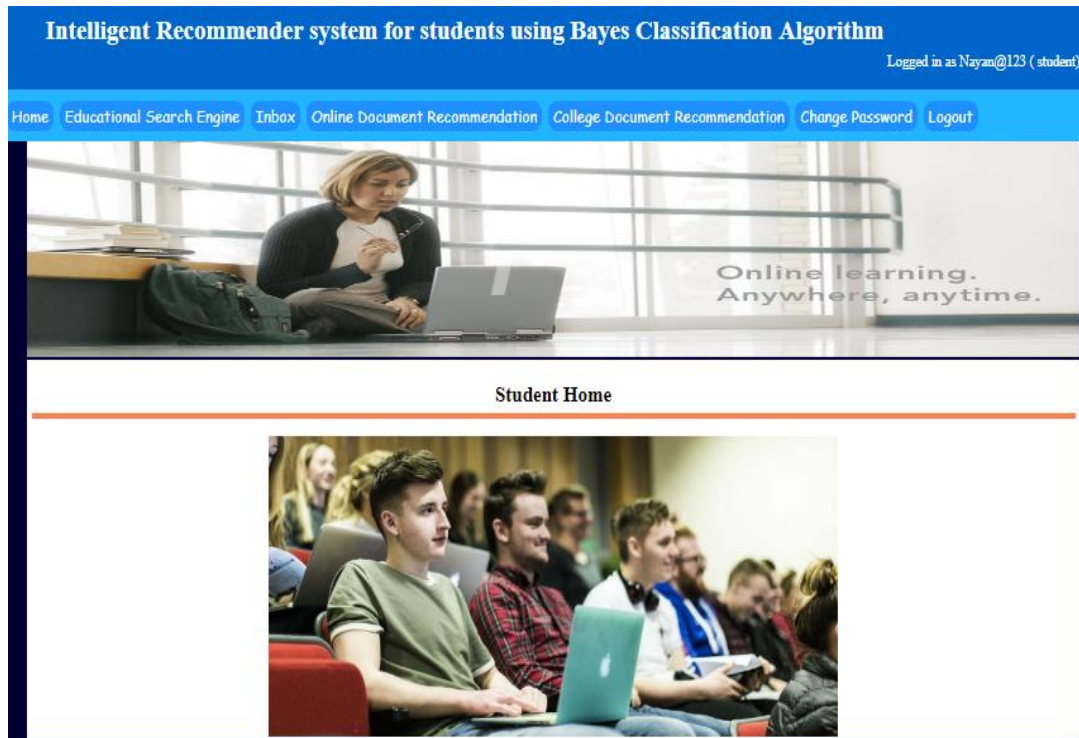


Figure 4. Student Panel

### 6.4. Classification Analysis

On the basis of all project working this final document analysis result is producing by following way.

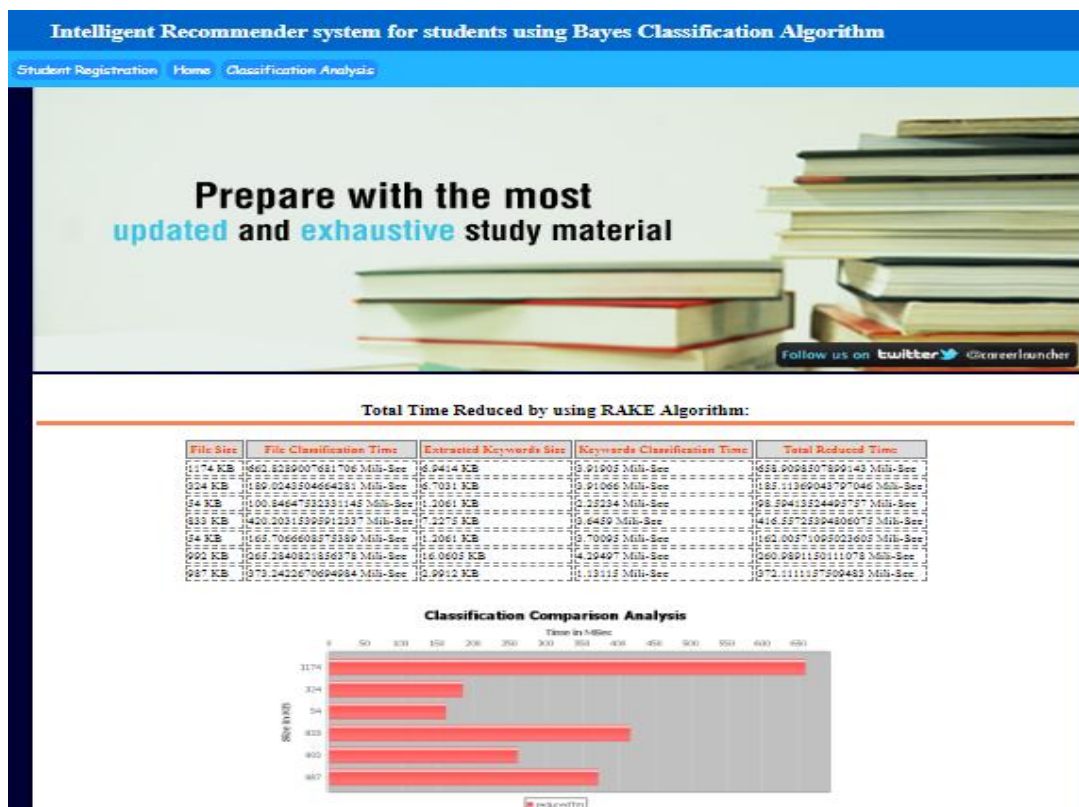


Figure 5. Classification Analysis

## REFERENCES

- [1] Callan, J., Smeaton, A., Beaulieu, M., Borlund, P., Brusilovsky, P., Chalmers, M., et al. (2003), "Personalization and recommender systems in digital libraries", NSF-EU DELOS Working Group Report.
- [2] Hong, J., Suh, E.H., Kim, J. & Kim S. (2009), "Context-aware system for proactive personalized service based on context history", *Expert Systems with Applications*, 36(4), pp. 7448-7457
- [3] Jiang, X. & Tan, A.H. (2009), "Learning and inferencing in user ontology for personalized Semantic Web search", *Information Sciences*, 179(16), pp.2794-2808
- [4] Chen, L.S. (2009), "Design and implementation of intelligent library system", *Library Collections, Acquisitions & Technical Services*, 32(3-4), pp.127-14
- [5] Song, D., Lau, R.Y.K. & Chen, P. (2007), "An intelligent information agent for document title classification and filtering in document-intensive domain", *Decision Support Systems*, 44(1), pp. 251-265
- [6] Ong, T.H., Chen, H. & Sung, W.K. & Zhu, B. (2005), "Newsmap: a knowledge map for online news", *Decision Support Systems*, 39(4), pp.583-597
- [7] Frias-Martinez, E., Magoulas, G., Chen, S. & Macredie, R. (2006)," Automated user modeling for personalized digital libraries", *International Journal of Information Management*, 26(3), pp.234-248
- [8] Feng, L. & Jeusfeld, M.A. & Hoppenbrouwers, J. (2005), "Beyond information searching and browsing: acquiring knowledge from digital libraries." *Information Processing and Management* 41(1), pp.97-120
- [9] Frias-Martinez, E., Chen, S.Y. & Liu, X. (2008), "Investigation of behavior and perception of digital library users: a cognitive style perspective", *International Journal of Information Management*, 28(5), pp.355-365
- [10] Frias-Martinez, E., Chen, S.Y. & Liu, X. (2009),"Evaluation of a personalized digital library based on cognitive style: adaptivity vs. adaptability", *International Journal of Information Management*, 29(1), pp.48-56