Improved e-Advising Model Using Data Mining: A Software Architecture Perspective

1PhD Research Scholar-Full Time, School of Computer Science, Engg.&Technology, Bharathidasan University, Tiruchirappalli, India
2Additional General Manager Bharat Heavy Electricals Limited Tiruchirappalli, Tamilnadu, India

DOI: 10.23956/ijarcsse/V7I5/0186

Abstract—A worldwide revolution in information and communications technologies (ICT) is in place. The Internet and the mobile phone are fundamentally changing our lives - affecting the way we work, learn and interact. Governments around the world are recognizing the value of e-Government. Properly designed and implemented, e-Government can improve efficiency in the delivery of government services, simplify compliance with government regulations, strengthen citizen participation and trust in government, and yield cost savings for citizens, businesses and the government itself. Not surprisingly, therefore, policymakers and managers are looking to adopt e-Government in countries around the world ranging from the most developed to the least developed [1]. This paper deals about using data mining effectively for e-Advising in higher educational institutions. A software architecture is presented to elaborate the process.

Keywords—Data Mining, e-Governance, e-Advising, ICT

I. INTRODUCTION

e-Government and e-governance can be defined as two very distinct terms. e-Governance is a broader topic that deals with the whole spectrum of the relationship and networks within government regarding the usage and application of ICTs. e-Government is actually a narrower discipline dealing with the development of online services to the citizen, more on any particular government service such as e-tax, e-transportation or e-health[2]. Data mining is generally thought of as a process of finding hidden, non-trivial and previously unknown information in a large collection of data. Association rule mining is an important component of data mining. Association rules are an important class of methods of finding patterns in data. Association rule mining has been used in many application domains. India has one of the largest higher education systems in the world. Despite having the largest higher education system, the quality of education, in general, cannot be claimed to be the best. Technical and vocational Education in India has seen enormous growth in recent years with a large increase in total number of institutes imparting higher education. On one hand, this growth promises to produce more skilled youth to fulfill needs of ever growing Indian industry and on the other hand it poses a huge challenge for the governing bodies like UGC, AICTE, etc. The introduction of e-governance in higher education is one such concept which can empower the governing bodies to administer the progress of the education plan in the whole country and serves various stakeholders in a much better way [3].

II. REVIEW OF LITERATURE

Devising a successful curriculum plan is very important in any higher education system. A good curriculum plan gives a successful career path to the students. As most of the universities used Information Systems for their successful e-Governance tasks, it is the appropriate time for the universities to use their historical data for better decision making in all the levels.

Introducing technology to the student decision making process aims at leveraging repetitive tasks on software and dedicating time to helping a student plan his/her education road map. An automated system for decision making or a technology-based advisory system helps a student plan the proper courses to take, by checking and listing courses for which he/she has satisfied the prerequisites, allowing students to do the work themselves, without referring to their advisors [4]. e-Advising answers the question of how to provide timely, high-quality advising services to geographically separated online students and/or onsite students with conflicting work or family schedules that cannot readily come to office hours [5].

It is important that, the University need to facilitate the students by giving the most current information regarding course selection, educational history and as well as current and future educational and career demands. At present most of the Universities maintain the database of the student performance for many years and have terabytes of data. Using these information systems, the student can select the courses based on the approved curriculum. Few Universities also implemented E-advising systems at certain level. These systems lacks proper decision making. If data mining is used in this decision making process, it helps the student in preparing effective educational plans. This is the appropriate time for the Universities to apply data mining in their data and retrieve useful knowledge and share it with all the stake holders [6].
This research proposes to conduct research on using a comprehensive data mining techniques, which will support for students in using the E-governance system in a more useful way. An architecture is presented to use the e-governance in higher educational institutions.

III. SCOPE OF THE RESEARCH WORK
Few higher educational institutions are not producing quality graduates. This leads to long range problems to students and institutions. This research addresses this issue by using the technology in enhancing the quality of students. The available data from student information system used by these institutions are considered initially to prove the validity of the developed model. The scope of the thesis is limited to e-Governance in higher educational institutions. This research work used the academic monitoring database available with the higher educational institutions. Data mining techniques are analyzed and identified to be used in the developed prototype. The software agents’ usage and its implications are well studied. Four new algorithms are designed, developed and tested in this research work. In order to further enhance the decision making, a new data structure is designed to serve as an expert assistance for the developed model. This new data structure contains the information from the industry experts which includes the current trends, expectations and job potentials in the chosen stream.

IV. E-GOVERNANCE ARCHITECTURE
The important aspects to be considered for the proposed framework architecture are the combination of the Apriori algorithm used for generating association rules, C4.5 algorithm used for generating classification tree and DBSCAN algorithm used for generating clusters. This kind of combination is not being attempted by any of the researchers. These algorithms working principles are sequenced so that the output of Apriori is used an input for C4.5 and the output of C4.5 is used an input to DBSCAN. Hence, each of these functional requirements is modelled as a separate module in the system. The usage of recommendation/recommender systems are vital in most of the e-Governance applications. The rapid advance of web and its applications has created a huge reputation for recommender systems. Being functional in various domains, recommender systems were planned to make recommendations such as items or services based on user interests. Basically, recommender systems experience many issues which replicates moderated effectiveness. Integrating authoritative data managing techniques to recommender systems can address such issues and the recommendations quality can be increased significantly.

A Student recommending system is an essential component of a successful academic involvement of any university or college. At present, most of the Universities/Colleges maintain the database of the student performance for many years and have terabytes of data. Using these information systems, the student can select the courses based on the approved curriculum. Few Universities also implemented E-Governance systems at certain level. It involves tasks where faculty members help students complete the requirements necessary to graduate. It also requires considerable planning on the part of both students and teachers. Various universities and institutions around the world use automated advising systems. They are helpful and beneficial for both advisors and advisees in that they contribute to assisting in making better-informed decisions and improved services. Introducing technology to the advising process aims at leveraging repetitive tasks on software and dedicating time to helping a student plan his/her education road map. If the Universities/Colleges uses the historical data decision making process, it helps the student in preparing effective educational plans. This is the appropriate time for the Universities to apply data mining in their data and retrieve useful knowledge and share it with all the stake holders. Figure 1 shows the conventional work flow method in a recommender system.

Fig. 1 Conventional Recommender System

In this research work, there are three distinct algorithms that can contribute to the success of the proposed framework:

- Apriori algorithm
- C4.5 algorithm
- DBSCAN algorithm
The Apriori Algorithm is a significant algorithm for mining frequent itemsets for boolean association rules. In this research work it helps to find the detailed analysis of association rule mining is carried out on the student-course registration data set by considering the following parameters:

- Volume of data
- Maximum number of learning outcomes for a course

The C4.5 algorithm help to generate a decision tree where each node splits the classes based on the gain of information from the processed data from the Apriori algorithm. The attribute with the highest normalized information gain is used as the splitting criteria. For this this existing information system, its functionality and data structure are studied. Clustering of processed data is done with the help of DBSCAN algorithm and it helps to cluster a set of points in some space, it groups together points that are closely packed together.

Thus, by using this framework, it could be a replacement for the existing conventional models where a student meets the faculty advisor for effective planning of course registration. The generic view of developed framework to accomplish the objectives of this research work is shown in below figure 2. It incorporates association rule mining algorithm along with Apriori algorithm and uses the student-course registration data set specialized database.

![Governance framework](image)

Fig. 2 e-Governance framework

V. ARCHITECTURE PROCESS

A model had been developed using the well-known Apriori technique in association rule mining and applied to the student-course registration data set. For the initial implementation, a dataset belongs to a college is considered. A larger data set across the university can also be used based on the need / objective. Java is chosen as the implementation tool and the open source data mining tool Rapid miner is used for the visualisation of the results produced by the Java program. A dataset contains the course registration detail is considered for implementation. This data set contains the student registration number and the list of course he/she has registered during study.

The results produced by the system benefits the students and faculty. This promotes the academic success whereby the overall advising quality is increased. The methodology used for the implementation is outlined as follows in table 3.1.

<table>
<thead>
<tr>
<th>Step</th>
<th>Detailed Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysing the association rule mining techniques, classification technique and clustering technique.</td>
<td>Detailed analysis of association rule mining is carried out. A best technique is chosen namely Apriori for association rule mining, C4.5 for generating classification trees and DBSCAN for generating clusters are chosen by considering the following parameters: Number of attributes, Volume of data.</td>
</tr>
</tbody>
</table>
Collection and pre-processing of data from University
The existing information system, its functionality and data structure are studied.
Made appropriate changes as needed without changing the functionality of the existing system
Archive database is built
Pre-processing such as data cleaning, transformation, reduction is carried out

Developing & implementing the proposed system by using the techniques collected in (a) & (b)
chooses JAVA as a front-end application for the development
Use Rapidminer as an visualisation tool

Testing the developed model and making it ready for the use
Test the model with various test cases

The user (student/faculty) logs in to the system using the credentials provided by the University. The proposed system retrieves the needed data from the information systems. The following information is available in the information systems:

- Student profile
- Curriculum, programme outcomes (Course Profile)
- Internal Assessment Marks
- Final Examination Marks
- Academic Calendar

The systems build a course plan with the help of an association rule mining technique by way of presenting interesting combination of courses chosen by the students. The association rule mining algorithm also links the data retrieved from the information systems with the data available in Archive. The user chooses a rule for further process. This selected rule is used as input to C4.5 algorithm. This algorithm generates classification models. The user further selects a branch of the generated tree. This is used an input to DBSCAN algorithm. This algorithm produces clusters which is used for making decision by the students. For the interaction between the user and between the data mining algorithms, different kinds of software agents are used. Five new agents are designed to meet this need. The agents and its working principles are provided in detail in Chapter 4. In this stage, the user checks the outputs with the expert database available with the archive. This expert database contains the data related to current trends, domain technologies etc., from the related field of specialisation of the user. The final results are then viewed with the help of a visualization tool. This tool presents the student with the set of possible outcomes by finding the correlation between two similar courses.

The proposed model can be restructured as in Figure 3.3. In the proposed e-governance model using data mining and intelligent agents, the user’s interaction is done through the user interface module. Then, the user requests are forwarded to the proposed module. Here, the user request is processed by proposed apriori algorithm. To predict the set of courses that a student can plan for the next semester, student information system data is used. Then, the calculated ratings are forwarded to the recommendation module, where the ratings are exploited to make list of subjects as recommendations along with the help of specialized database. The main ingredient of this work is trust between the users specialized database.

The user interface module phase describes the process of choosing a single service from the list of services returned by the user interface. In addition to selecting and returning a single value, the interface module builds additional values, a workflow that offers a classification of sub task facilities, that can be integrated to perform a single job.
Figure 4 depicts the class diagram that shows the relationship between the classes in the proposed architecture. The user interface class, expert database class and algorithm classes and its relationships are represented.

VI. ADVANTAGES

e-Governance should make full use of all existing modern technology, if necessary, to deliver the effective help to the students in higher educational institutions. However, technology as a means of offering advisory services can be viewed by some to be cold and impersonal. It is therefore encouraged that technology be harnessed not to fully replace, but to improve the efficiency of the overall advising process, still allowing students to physically meet a human advisor if necessary.

As a result, technology can enhance the e-Governance experience by assisting in the making of better informed decisions as well as providing improved services by migrating repetitive tasks on software. This would allow any student advisor time to be dedicated to helping a student select the most appropriate path or handling any non-academic issues that may have an impact on the student’s performance. Few of the major advantages of the proposed research work is discussed below.

Availability: The intelligent e-Governance system is available all the time, while the human academic advisor has fixed and limited time, and needs prior appointment.

Initiative: The Intelligent e-Governance system takes the initiative to contact the students, while the human academic advisor waits for students to come to contact him/her.

Comprehensiveness: The Intelligent e-Governance system gives much knowledge and information about academic requirements in different levels: university, college, and department with rational explanation for these requirements, while human academic advisor has a limited memory capacity to memorize all details and reasons.

Knowledge-based advice: The Intelligent e-Governance system can make short-term as well as long-term plans using the appropriate background knowledge while human academic advisor can only take the interests of the student in account, without applying previous cases like his/her.

Continuous monitoring: The Intelligent e-Governance system continuously monitors students’ progress toward educational goals by up-to-date information while the human academic advisor needs to access different records to keep track of the student record.

VII. FUTURE DIRECTIONS

The following points can be considered in future to improve the performance of the developed e-Governance model.

i. Incorporation of different data mining algorithms.

More data mining techniques and algorithms will be added in the developed model to further enhance the performance of the e-Governance model. At present the prototype uses only the three types of data mining algorithms specified in Chapter 1. The remaining types of data mining algorithms also will be considered to develop a full fledged system prototype.

ii. Data integration.

Data needed for data mining process is usually stored in different industrial firms and government agencies using different data formats. Data formats are both structure-specific and vendor-specific. There have been a lot of works on data integration and data exchange, but with little success. These issues have become crucial with the emergence of data warehouses.

VIII. CONCLUSIONS

E-Governance in higher educational institution is very important in the academic system and it is the duty of the Universities/Colleges to use the appropriate technologies and providing the best available knowledge to the student community. This research work clearly shows that, the e-Governance model will surely benefit the faculty/student. The developed model provides the administration with a set of possible courses that can be offered in a semester. With this input, the management can plan for expertise and other resources needed effectively.
REFERENCES


[6] Jiawei Han, Micheline Kamber, and Jian Pei, Data Mining: Concepts and Techniques, 3rd edition, Morgan Kaufmann, 2011.