

Predictive Analytics: A Step Ahead For Education System

Sunil Yadav

Assistant Professor, Information Technology, Shree L.R. Tiwari College of Engineering, India.

Supriya Yadav, Sumita Savaliya, Madhuri Thanvi and Vishalkumar Pandey

Software Industry Expert India.

Sunil.yadav@slrtce.in

Abstract: Year gap rates for students in correspondence and open courses are on increase. There is a need of analysis of factors causing increase in this rate. The discovery of hidden knowledge from the educational data system by the effective process of data mining technology to analyze factors affecting student performance can lead to a better academic planning and management to reduce students year gap from the course, as well as can generate valuable information for decision making of stake holder to improve the quality of higher educational system. Data mining technique can be used for such analysis and prediction. In this project, we will use real data from study centers like SHREE L. R. TIWARI COLLEGE OF ENGINEERING and data from various sources like university database, survey form etc. Various steps of mining will be applied on the available data to deduce useful result. Various scenarios will be compared and there accuracy will be calculated.

This project presents the work of data mining in predicting the drop out feature of students and analysis of data set using data mining algorithms. After analysis the outcome will be the major factors that affect student results of the open courses the most. Before applying classification algorithms some feature selection algorithms are also used so as to get refined prediction results. Such analysis and prediction information will help college management and teachers to make necessary changes for imparting better education. Mining of useful knowledge can be done by using many other mining techniques like association, clustering etc.

The educational databases contain hidden information for improvement of students' performance. Educational data mining is used to study the data available in the educational field and bring out the hidden knowledge from it. This prediction will help to identify the weak students and help them to score better marks and save them from getting year gap. The outcome predicts the number of students who are likely to get year gap. The results provide steps to improve the performance of the students who were predicted to fail. After the declaration of the results in the final examination the marks obtained by the students are fed into the system and the results were analyzed for the next session. The comparative analysis of the results states that this prediction will help the weaker students to improve and brought out betterment in the result.

Keywords: Association Classifier; predictive analysis.

I. INTRODUCTION

The development of Information Technology has generated large amount of databases and huge data in various areas. The research in databases and information technology has given rise to an approach to store and manipulate this precious data for further decision making. Data mining is a process of extraction of useful information and patterns from huge data. It is also called as knowledge discovery process, knowledge mining from data, knowledge extraction or data /pattern analysis.

Various algorithms and techniques like Classification, Clustering, Regression, Artificial Intelligence, Association Rules, Decision Trees etc., are used for knowledge discovery from databases.

Knowledge discovery in database (KDD) is the process of discovering useful knowledge from a collection of data. This is widely used data mining techniques. The major KDD application areas are marketing, fraud detection, telecommunication and manufacturing.

What is Predictive Analytics?

Predictive analytics is a technique of data mining that analyze current and historical trends in data to predict the future outcome or to get knowledge of occurrence of any unknown events. Predictive analytics can be applied to any type whether it is known or unknown or that event has been occurred in past, present or future. Predictive analytics is based on relation between predicted variables and the variables which are to be predicted to get the result.

Predictive Analytics in Education System

In recent years it has been seen that the number of student failures and year gap has been increased in higher education among different universities in India. Research has been done to identify the reason for such a low performance in their academics. The data stored about the students are very valuable as they help to determine the future outcome of student performance in his entire duration of higher education.

It must be identified that what information is related to the student performance in their academics and then perform analysis on that data. It can be done by using a technique known as knowledge discovery in database or data mining in educational system which is called as Educational data mining.

Collecting Required Data

In this step data is collected related to students which will be used in analysis and predicting the student performance. All the related data to the student is collected, the data collected play an important role it is important to identify the data which makes influence on the performance of the students. All the information collected is used for the process ahead and they are integrated into a single dataset which consist of all the related attributes to the students.

Data Mining Process

The data collected in the previous step is prepared for the actual processing which will produce the result. The data is prepared before applying it to algorithms, it includes making the data suitable according to the system specification. After this the dataset prepared in the final step is used and different algorithms are applied on it. Algorithms related to Classification, Clustering and Prediction are used to extract the data from the given information. A lot of process takes place in this step to get the result required to predict the student performance.

II. LIMITATIONS OF EXISTING SYSTEM

This is proposed to be an enhanced model of the present day existing prediction model. Many flaws in the present data mining tools haven't been able to exploit the full potential of data mining techniques.

This project gives a procedural on how a student's academic performance can be enhanced based on the details of performance available about the students in their previous semesters. Depending upon various criteria such as term test, lab work, assignment, family background etc. the conclusion is derived for the same.

III. SYSTEM ANALYSIS AND DESIGN

Systems analysis is a problem solving technique that decomposes a system into its component pieces for the purpose of the studying how well those component parts work and interact to accomplish their purpose.

Analysis modeling uses combination of text and diagrammatic forms to depict requirements for data, function, and behaviour in a way that is relatively easy to

understand and more important, straightforward to review for correctness, completeness and consistency.

ANALYSIS:

Analysis modeling often begin with data modeling. It is important to define all data objects that are processed within the system, the relationships between data objects and other information that is pertinent to the relationships

Data objects

A data object is a representation of almost any composite information that must be understood by software. The composite information means number of different properties and attributes.

The various data objects included in Predictive Analytics project include external entity that can be a user, a thing which includes report format of result, mining module which can be considered as a place etc.

Data Attributes

Data attributes define the properties of a data object. For Example: The various attributes of an external entity i.e. user can be user name, password etc. Analysis modeling leads to the derivation of various modeling elements which are as follows:

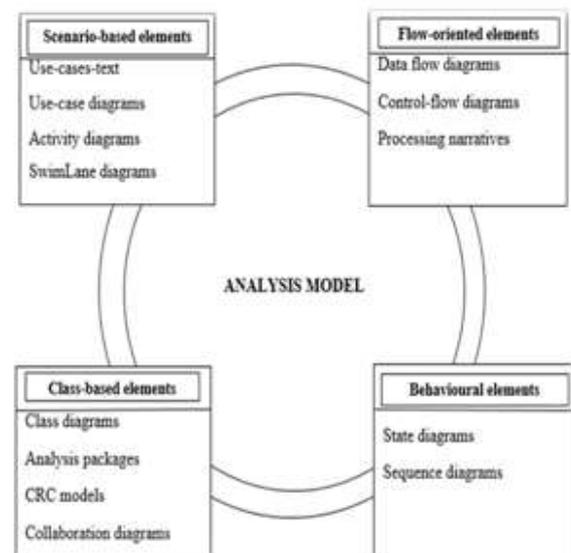


Fig.1 Analysis Model

A. Scenario based modeling: Scenario based modeling represents the system from user point of view.

B. Flow oriented modeling: Provides an indication of how data objects are transformed by processing functions.

C. Class based modeling: Class based modeling defines various objects, attributes, relationships.

D. Behavioural modeling: Depicts the states of the system and its classes and the impact of events on these states.

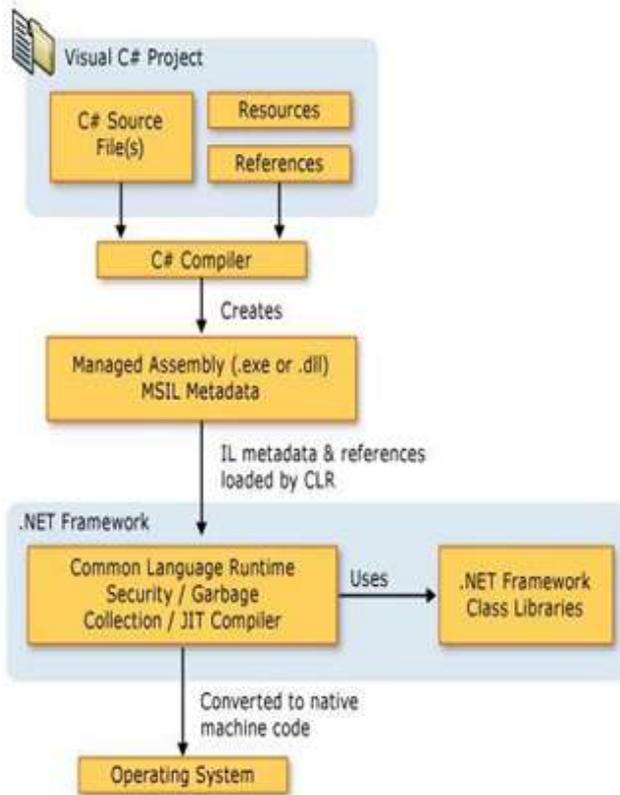


Fig.2 Architecture of System

REFERENCES

- [1] S.Yadav,A.Desai and V.Yadav , “Knowledge Management in CRM using Data mining Techniques”, International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July-2013 ISSN 2229-5518978.
- [2] E.W.T. Ngai, Li Xiu and D.C.K. Chau, “Application of data mining techniques in customer relationship management: A literature review and classification.” International journal of Expert Systems with Applications 36 (2009) 2592–2602 /ELSEVIER.
- [3] Kun Wu and Feng-ying Liu, “Application of Data Mining in Customer Relationship Management”, 978- 1-4244-5326-9/10/2010 IEEE.
- [4] P.Sonar, S.More, S.Yadav and S.Sall,” SALES AND INVENTORY PREDICTION USING DATA MINING”, International Journal For Technological Research In Engineering Volume 3, Issue 7, March-2016, ISSN: 2347 - 4718
- [5] S.Yadav,G.Patwa,P.singh , Deepak Maurya,” Association Classifier for Predictive Analysis in Health

Care Mining”, International Journal of Recent Trends in Engineering & Research (IJRTER) Volume 02, Issue 03; March - 2016 [ISSN: 2455-1457].

- [6] Jiawei Han and Micheline Kamber, Data Mining Concepts and Techniques, M.K. Publications, Second Edition.
- [7] Working with System design: http://www.tutorialspoint.com/software_engineering/software_design_basics.html
- [8] Working with Naive Bayesian Classifier: https://en.wikipedia.org/wiki/Naive_Bayes_classifier
- [9] Working with Naive Bayesian Analysis Model:http://www.saedsayad.com/naive_bayesian.htm
- [10] Working with SQL Server: [https://msdn.microsoft.com/enus/library/ms167593\(v=sql.105\).aspx](https://msdn.microsoft.com/enus/library/ms167593(v=sql.105).aspx)