

A Review of Educational Data Mining in Higher Education System

T.Thilagaraj

*Assistant Professor in Computer Applications
Kongu Arts and Science College,
Erode.
thilagaraj.t@gmail.com*

Dr.N Sengottaiyan

*Director Academics
Hindusthan College of Engineering and
Technology,
Coimbatore, Tamil Nadu 641032
nsriram3999@gmail.com*

Abstract

The discovery of hidden patterns in educational data is a promising research in Educational Data Mining. The students achievement rate were reduced continuously is the major problem in higher education. To increase the success rate of students the early forecast technique will help the management to counsel the poor students at right time. To discover the new patterns from various data the data mining approach is widely used. Likewise here the data mining is used in educational field to extract hidden patterns. Classification is used to classify the records based on the preparation set and also it uses the pattern to categorize the new records. This paper aims to show the various techniques of Educational data mining that guides the management to take better action on students at risk.

Keywords: Educational Data Mining (EDM), Classification, Data Mining, Prediction.

I. INTRODUCTION

The financial system of the country is directly depends on the students education which makes an impact in industries. The excellence of educational institutions is viewed by the success rate of students and the skill set of the institutions will be measured by retained rate of students at risk. The different aspects like individual, social and psychological will be useful to measure student academic performance. This may lead to discover the students who are in risk and it help the management to take timely action. The student academic performance will be measured by their socio-economic and previous academic performances. This process will be performed by using educational data mining techniques. The determination of classes will be made before examining the data so it is also referred to as supervised learning.

Based on the previous academic performances and Socio-economic situations the student performances were measured through Data mining techniques.

Classification maps the data into predefined sets or groups of classes. It is often referred to as supervised learning because the classes are persistent before examining the data. Patterns that are discovered by Data Mining methods from educational data can be used to enhance decision making in terms of identifying students at risk, decreasing student drop-out rate, increasing student's success and increasing student's learning outcome. The main objectives of this study are identification of different factors which affects a student's learning behavior and performance during academic career. Construction of a prediction model using classification data mining technique on the basis of identified predictive variables and extract valid information from existing students to manage relationships with upcoming students. And also to improve the performance of the student and validation of the developed model for higher education students studying in Universities or Institutions. The educational data mining methods are used to discover the patterns for improving the decision making to identify the students at risk. And also it is used to increase the students learning outcome and to decrease the drop-out rate of the students [1].

Educational data mining is emerging as a research area with a suite of computational and psychological methods and research approaches for understanding how students learn. Education Data Mining in data mining retrieve hidden knowledge by applying various techniques of data mining like clustering, rule mining, web based mining, test mining, neural network, Bayesian network, and many others which gives us a final result and if it need some requirement to changes then again the raw data if filtered according to need. Data collected from education institute can be aggregated over large numbers of students and can contain many variables that data mining algorithms can explore for model building. Data mining is also known as knowledge discovery in databases is a technology used in different discipline to search for significant relationship among variables in large database. It extract knowledge form hidden data base and find the necessary information according to the requirement. These tasks are done by applying various algorithm of data mining on training

dataset. These dataset passes through various steps of data mining and finally the filtered data is used by the user according to the requirement. It plays an important role in education institute. Measuring the student performance will help the placement officer to guide the student to decide their carrier in accurate path. According to the company's perceptive they expect the ability from the students like technical, communication, percentage, positive attitude and good performance. If the educational planners will conduct the awareness program in advance it may protect the students from risk in earlier and also it improve the overall productivity [2]. The learning needs are different among student groups and it can view by using educational data mining techniques [3].

II. EDUCATIONAL DATA MINING TECHNIQUES

The various techniques are used to acquire the hidden information in Educational data mining. They categorized into four groups:

A. Decision trees

To understand easily and to identify the most promising variables in fastest way the decision tree techniques are more liable. It is also used find the relativity between two or more variables. The new variables or features can be generated by the decision trees to predict the target variable. The algorithms used majorly in decision trees are ID3, CART, C4.5, Random tree and CHIAD.

B. Bayesian Classifier

This is very easy and simplest technique that requires less amount of data preparation to calculate the parameters. The unrelated features are insensitive and also the classifier is well-organized to deal both real and distinct data. The class conditional independencies are clear between subsets of variables in Bayesian classifier. A graphical model of causal relationship will provide to perform the learning process.

C. Neural Networks

By deriving the meaning from complex data to identify the trends and extract the risky patterns is the huge ability of neural networks. Discovering tasks based on the given data for training or preliminary experience is one the main advantage of neural networks. The Neural network has the capability to identify all possible interactions of the variables among various predictors. This is the main reason of using neural network in Educational data mining [4].

D. Clustering Techniques

It is a process of constructing a group of abstract objects into classes of parallel objects. The set of data into groups are partition first and then based on data comparison cluster analysis were made. Next mentioning labels to all groups and also here the data objects are treated as separate set. This techniques is flexible to make changes and it helps out individual positive features to distinguish different groups are main advantages of clustering in classification. The different areas like data analysis, pattern recognition and image processing are mainly focused by clustering techniques.

III. STUDENT ADVISORY SYSTEM FRAMEWORK

A. Framework Description

A student from particular branch is assured for recommendations of both classification and clustering techniques. This process needs an educational dataset to initiate the task. This framework is used to represent the student academic and non-academic attributes in higher education to hold complete information about students.

B. Classification Phase

To discover an efficient classifier in the educational dataset the classification algorithm is used. The recommended course must be viewed for the student is the main role of classifier. The current level failed student records are removed in this classification phase. The different classification algorithms were applied by using this training dataset with course attribute to improve the student performance.

C. Clustering Phase

Based on marks resemblance the student records are partition into number of clusters by applying clustering algorithm on educational dataset. The number of clusters is decided and previous grade attributes are removed in this phase. To identify the clusters and distribution of the each course percentage is identified using K-means algorithm.

D. Request an Output from the System

This phase will describe the procedure about how the data will be entered by the new student in the system. Here the system will read the data and check whether the data will be valid for the process or not. The classification phase will predict the department with certain norms and regulations which is already declared [6].

IV RELATED WORK

In [1], K.R.Kavyashree, LakshmiDurga has proposed that a country's growth is strongly measured by the quality of its education system. Education sector has witnessed sea change in its functioning. Today it is recognized as an industry and as an industry it is facing challenges. The challenges of higher education being decrease in students' success rate and their leaving the course without completion. An early prediction of students' failure may help the management provide timely counselling as well as coaching to increase the success rate and student retention. Data mining are widely used in educational field to find new hidden patterns from student's data which are used to understand the problem. Classification is one of the prediction type classifiers that classifies data based on the training set and uses the pattern to classify a new data. Aim of the project is to develop an internetworking application that uses data mining technique to predict the students' performance based on their behaviour. This paper explored the link between emotional skills of the students along with the socio economic and previous academic performance parameters using Naive Bayes Classifier technique.

In [2],TriptiDwivedi, Diwakar Singh has proposed that the survey of student background history which helps academic planners in institute to give right direction to student. If the class of student is predicted in midsession of institute in final year then it will be easy for the academic planner to plan some important workshop for the enhancement of performance of student which helps it in placement at the end of academic session. In educational institute data mining techniques plays an important role in each activities of institute whether it is academic, cultural, examination and training and placement etc. in which Educational Data Mining which is a field of data mining helps a lot to find the actual filtered data in various field of department in institute. Hidden knowledge through data mining techniques is extracted from large database which helps to predict the pattern in such activities. It plays a great role in predictions of student data for placement and performance.

In [3], AbdulmohsenAlgarinihas proposed that the data mining techniques are used to extract usefulknowledge from raw data. The extracted knowledge is valuableand significantly affects the decision maker. Educational datamining (EDM) is a method for extracting useful informationthat could potentially affect an organization. The increase oftechnology use in educational systems has led to the storageof large amounts of student data, which makes it importantto use EDM to improve teaching and learning processes. EDMis useful in many different areas including identifying at-riskstudents, identifying priority learning needs for different groupsof students,

increasing graduation rates, effectively assessinginstitutional performance, maximizing campus resources, andoptimizing subject curriculum renewal. This paper surveys therelevant studies in the EDM field and includes the data andmethodologies used in those studies.

In [5], Amirah Mohamed Shahiria, WahidahHusaina and Nuraini Abdul Rashidahhas proposed that the Predicting students performance becomes more challenging due to the large volume of data in educational databases.Currently in Malaysia, the lack of existing system to analyze and monitor the student progress and performance is notbeing addressed. There are two main reasons of why this is happening. First, the study on existing prediction methodsis still insufficient to identify the most suitable methods for predicting the performance of students in Malaysianinstitutions. Second is due to the lack of investigations on the factors affecting students achievements in particularcourses within Malaysian context. Therefore, a systematical literature review on predicting student performance byusing data mining techniques is proposed to improve students achievements. The main objective of this paper is toprovide an overview on the data mining techniques that have been used to predict students performance. This paperalso focuses on how the prediction algorithm can be used to identify the most important attributes in a students data.We could actually improve students achievement and success more effectively in an efficient way using educationaldata mining techniques. It could bring the benefits and impacts to students, educators and academic institutions.

In [6], Heba Mohammed Nagy, Walid Mohamed Aly, Osama FathyHegazyhas proposed that the educational data mining is a specific data miningfield applied to data originating from educational environments, itrelies on different approaches to discover hidden knowledge fromthe available data. Among these approaches are machine learningtechniques which are used to build a system that acquires learningfrom previous data. Machine learning can be applied to solvedifferent regression, classification, clustering and optimizationproblems.In their research "Student Advisory Framework"that utilizes classification and clustering to build an intelligentsystem. This system can be used to provide pieces of consultations toa first year university student to pursue a certain education trackwhere he/she will likely succeed in, aiming to decrease thehigh rate of academic failure among these students. A real casestudy in Cairo Higher Institute for Engineering, Computer Scienceand Management is presented using real dataset collected from2000–2012.The dataset has two main components: pre-highereducation dataset and first year courses results dataset. Results haveproved the efficiency of the suggested framework.

In [7], Monika Goyal and RajanVohrahas proposed that the data analysis plays an important role for decision supportirrespective of type of industry like any manufacturing unit andeducations system. There are many domains in which datamining techniques plays an important role. This paper proposesthe use of data mining techniques to improve the efficiency ofhigher education institution. If data mining techniques such asclustering, decision tree and association are applied to highereducation processes, it would help to improve studentsperformance, their life cycle management, selection of courses,to measure their retention rate and the grant fund management ofan institution. This is an approach to examine the effect of usingdata mining techniques in higher education.

In [8], U.K.Pandey and S.Pal has proposed that from ancient period in India, educational institution embarked to use class room teaching. Where a teacher explains the material and students understand and learn the lesson. There is no absolute scale for measuring knowledge but examination score is one scale which shows the performance indicator of students. So it is important that appropriate material is taught but it is vital that while teaching which language is chosen, class notes must be prepared and attendance. This study analyses shows the impact of language on the presence of students in class room. The main idea is to find out the support, confidence and interestingness level for appropriate language and attendance in the classroom. For this purpose association rule is used.

In [9], Dr. MohdMaqsood Ali has proposed that the Universities either public or private and its colleges enroll thousands of students into various courses or programs every year. They collect information from students at the time of admissions and store the same in computers. Understanding the benefits of data is essential from business point of view. Data can be used for classifying and predicting the students' behavior, performance, dropouts as well as teachers' performance. Therefore, this paper“Role of data mining in education sector” examines the role of data mining in an education sector. In addition, layemphasis on application of data mining that contribute to offer competitive courses and improve their business.

V EDUCATIONAL DATA MINING APPLICATIONS

A. Visualization of facts

Inserting a common term in a visual context is to guide the people to identify the significance of the data is known as data visualization.According to usertrends the reports were generated monthly or weekly schedule. Using of materials, studying topics

sequence, studying activity patterns and time schedules are noted as usage summary. It is very easy to know about patterns and correlations in visualization software, which are hidden in text-based data [7].

B. Predicting Student Performance

The Marks, knowledge and student performance are frequently predicted values in educational data mining. To improve the learning and teaching procedure, the performance of students will be predicted to guide the learners and educators in correct way.The percentage and grade are mostly used in educational data mining by the researchers. This technique is used to combine the labeled items based upon the quantitative traits and training sets which are gathered earlier in the process [5].

C. Enrolment Management

To structure the enrolment of the college to achieve the goals the enrolment management is used in higher education.The data analysis is used in enrollment management for achieving desired results of the management is the traditional way of educational data mining techniques. The educational institutions are planned to perform set of activities to influence frequently over the students enrolments will lead to reduce the drop-outs in higher education [8].

D. Grouping Students

The student groups are formed according to the personality and efficiency to improve the system.To construct a learning system to support the group learning methodology will make the learning techniques are more effective and easier for the students to develop themselves. Discover the student groups in related learning which is based on huge sequences are performed by clustering algorithm [7].

E. Predicting Students Profiling

At the time of admission the information collected by the management from students will hold the details like demographic, geographic and psychographic individual of the students. The best technique to identify the different types of students individuality is Neural networking [9]. The prediction of student performance are possible by using different techniques like Bayesian networks, decision trees and neural networks will make the management to take appropriate decisions to improve the students performance in higher education.

The prediction about final grade and course completion are major things found in this student profiling techniques to view the success rate of students [10].

F. Planning and scheduling

The Educational process like planning, course scheduling, resource allotment and going for new courses will make idea for admissions and counseling. These are the important aspects for the management to make an impact in the educational system. While planning the course activities the decision trees and link analysis are used to find course completion rates and preferences. To find the course classifications in educational training the clustering analysis, decision trees and back-propagation neural networks techniques are used to improve the student level in higher education [7].

VI. TYPES OF DATA CAN BE MINED

- **Flat files:** Flat files are actually the most common data source for data mining algorithms, especially at the research level. Flat files are simple data files in text or binary format with a structure known by the data mining algorithm to be applied. The data in these files can be transactions, time-series data, scientific measurements, etc.
- **Relational Databases:** Briefly, a relational database consists of a set of tables containing either values of entity attributes, or values of attributes from entity relationships. Tables have columns and rows, where columns represent attributes and rows represent tuples. A tuple in a relational table corresponds to either an object or a relationship between objects and is identified by a set of attribute values representing a unique key.
- **Data Warehouses:** A data warehouse as a store house, is a repository of data collected from multiple data sources (often heterogeneous) and is intended to be used as a whole under the same unified schema. A data warehouse gives the option to analyze data from different sources under the same roof.
- **Transaction Databases:** A transaction database is a set of records representing transactions, each with a time stamp, an identifier and a set of items. Associated with the transaction files could also be descriptive data for the items. For example, in the case of the video store, the rentals table.
- **Multimedia Databases:** Multimedia databases include video, images, audio and text media. They can be stored on extended object-relational or object-oriented databases, or simply on a file system. Multimedia is

characterized by its high dimensionality, which makes data mining even more challenging. Data mining from multimedia repositories may require computer vision, computer graphics, image interpretation, and natural language processing methodologies.

- **Spatial Databases:** Spatial databases are databases that, in addition to usual data, store geographical information like maps, and global or regional positioning. Such spatial databases present new challenges to data mining algorithms.
- **World Wide Web:** The World Wide Web is the most heterogeneous and dynamic repository available. A very large number of authors and publishers are continuously contributing to its growth and metamorphosis, and a massive number of users are accessing its resources daily. Data in the World Wide Web is organized in interconnected documents. These documents can be text, audio, video, raw data, and even applications. Conceptually, the World Wide Web is comprised of three major components: The content of the Web, which encompasses documents available; the structure of the Web, which covers the hyperlinks and the relationships between documents; and the usage of the web, describing how and when the resources are accessed.
- **Time-Series Databases:** Time-series databases contain time related data such stock market data or logged activities. These databases usually have a continuous flow of new data coming in, which sometimes causes the need for a challenging real time analysis. Data mining in such databases commonly includes the study of trends and correlations between evolutions of different variables, as well as the prediction of trends and movements of the variables in time.

VII. POWERFUL EDM TOOLS

Data mining has a wide number of applications ranging from marketing and advertising of goods, services or products, artificial intelligence research, biological sciences, crime investigations to highlevel government intelligence. Due to its widespread use and complexity involved in building data mining applications, a large number of Data mining tools have been developed over decades. Every tool has its own advantages and disadvantages. Within data mining, there is a group of tools that have been developed by a research community and data analysis.

They are offered free of charge using one of the existing open source licenses. An opensource development model usually means that the tool is a result of a community effort, not necessary supported by a single institution but instead the result of contributions from an international and informal development team. This development style offers a means of incorporating the diverse experiences data mining provides many mining techniques to extract data from databases. Data mining tools predict future trends, behaviors, allowing business to make proactive, knowledge driven decisions. The development and application of data mining algorithms requires use of very powerful software tools. As the number of available tools continues to grow the choice of most suitable tool becomes increasingly difficult [11].

A. RapidMiner (YALE)

Rapid Miner is a software platform developed by the company of the same name that provides an integrated environment for machine learning, data mining, text mining, predictive analytics and business analytics. It is used for business and industrial applications as well as for research, education, training, rapid prototyping, and application development and supports all steps of the data mining process. Rapid Miner uses a client/server model with the server offered as Software as a Service or on cloud infrastructures.

It is released on 2006 and the latest version available is Rapid miner 6. It is licensed by AGPL Proprietary and it is cross platform i.e. can be installed on any operating system. And also language independent it can be downloaded from www.rapidminer.com.

The general features of rapid miner are an environment for machine learning and data mining processes. It represents a new approach to design even very complicated problems by using a modular operator concept which allows design of complex nested operator chains for huge number of learning problems. Rapid miner uses XML to describe the operator trees modeling knowledge discovery process. It has flexible operators for data input and output file formats. It contains more than 100 learning schemes for regression classification and clustering analysis. Rapid miner supports about twenty two file formats. Rapid Miner has a lot of functionality, is polished and has good connectivity. Rapid Miner includes many learning algorithms from WEKA. It is solid and complete package. It easily reads and writes Excel files and different databases. We can program by piping components together in a graphic ETL work flows. If you set up an illegal work flows Rapid Miner suggest Quick Fixes to make it legal.

RapidMiner provides support for most types of databases, which means that users can import information from a variety of database sources to be

examined and analyzed within the application. Specialized for Business solutions that include predictive analysis and statistical computing.

Advantages:

- It has the full facility for model evaluation using cross validation and independent validation sets.
- Over 1,500 methods for data integration, data transformation, analysis and, modelling as well as visualization – no other solution on the market offers more procedures and therefore more possibilities of defining the optimal analysis processes.
- RapidMiner offers numerous procedures, especially in the area of attribute selection and for outlier detection, which no other solution offers.

Limitations:

- In RapidMiner are suited for people who are accustomed to working with database files, such as in academic settings or in business settings.
- The reason for this is that the software requires the ability to manipulate SQL statements and files.

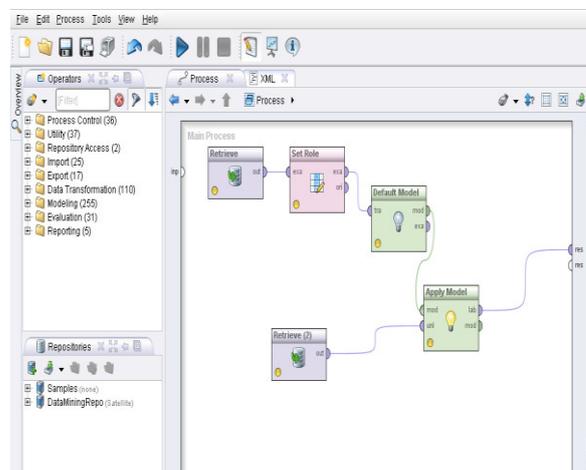


Fig 1. Flow diagram for RapidMiner

B. WEKA

Weka means Waikato Environment for Knowledge Analysis. It is a collection of machine learning algorithms for data mining tasks. This Tool is very useful and it is majorly used by the data mining researchers for predictive modeling and fact analysis. While compare this with rapidminer it have several advantages and supports standard educational data mining tasks like clustering, classification,

Advantages:

- It is very extensive statistical library.
- It is a powerful elegant array language in the tradition of APL, Mathematica and MATLAB, but also LISP/Scheme.
- Ability to make a working machine learning program in just 40 lines of code.
- The Numerical programming is better integrated in R and it has better graphics.
- R is more transparent since the Orange are wrapped C++ classes.
- Easier to combine with other statistical calculations. Import and export of data from spreadsheet is easier in R, spreadsheet are stored in a data frames that the different machine learning algorithms are operating on.
- Programming in R really is very different, you are working on a higher abstraction level, but you do lose control over the details.

Limitations:

- It is less specialized towards data mining.
- There is a steep learning curve, unless you are familiar with array languages.

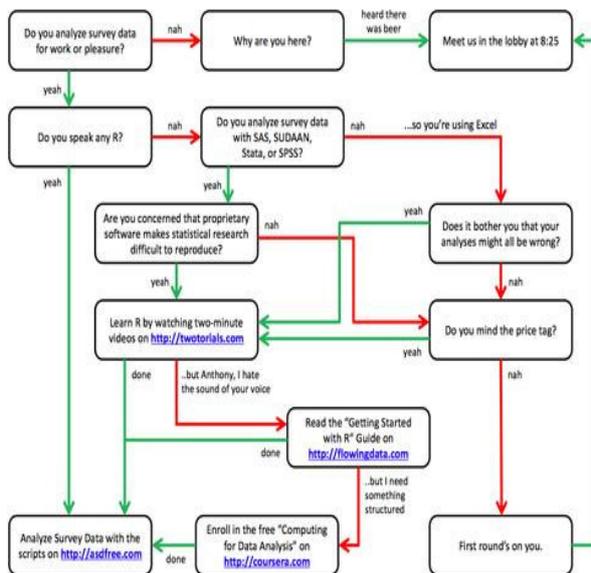


Fig 3. Flow diagram for R-Programming

D. Orange

The Orange is open source tool for educational data mining researcher and it is Python-based. The machine learning mechanism and the few features are added to bioinformatics and text mining is the main advantage of Orange tool. Orange is a component-based data

mining and machine learning software suite, featuring a visual programming front-end for explorative data analysis and visualization, and Python bindings and libraries for scripting. It includes a set of components for data preprocessing, feature scoring and filtering, modeling, model evaluation, and exploration techniques. It is implemented in C++ and Python. Its graphical user interface builds upon the cross-platform framework. It is developed in 2009. The Latest version available is Orange 2.7 and Licensed by GNU General Public License. It is Compatible with Python, C++,C.

The General features are it is component-based data mining and machine learning software suite. It includes a set of components for data preprocessing, feature scoring and filtering, modeling, model evaluation, and exploration techniques. Data mining in Orange is done through visual programming or Python scripting. It is open source data visualization and analysis for novice and experts. It contains components for machine learning and add-ons for bioinformatics and text mining.

Along with it's also packed with features for data analytics [11]. It is specialized for data visualization along with mining. The several advantages of Orange tool are open source data mining package build on Python, NumPy, wrapped C, C++ and Qt. Works both as a script and with an ETL work flow GUI. Shortest script for doing training, cross validation, algorithms comparison and prediction. Orange the easiest tool to learn and it is Cross platform GUI. It is written in python hence is easier for most programmers to learn. Scripting data mining categorization problems is simpler in Orange. Orange does not give optimum performance for association rules.

The limitations are it is not super polished. The install is big since you need to install QT. Limited list of machine learning algorithms. Machine learning is not handled uniformly between the different libraries. Orange is weak in classical statistics; although it can compute basic statistical properties of the data, it provides no widgets for statistical testing. Reporting capabilities are limited to exporting visual representations of data models.

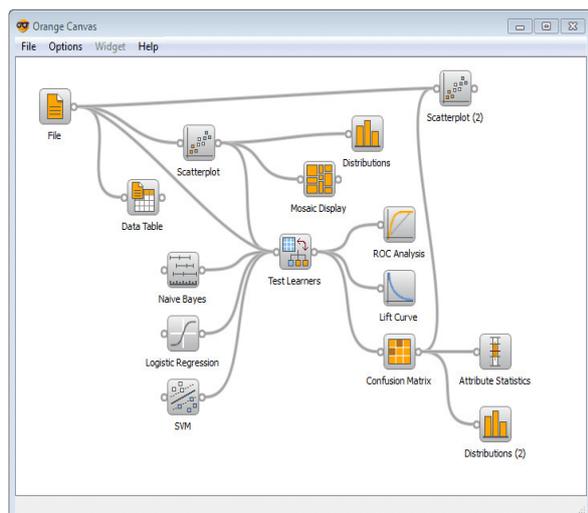


Fig 4: Flow diagram for Orange Tool.

E. KNIME

This KNIME tool is based on Eclipse and written in java. Transformation, extraction and loading are the powerful components of data preprocessing technique were performed well in this tool. The KNIME tool will permit the assembly of nodes for data processing through graphical user interface concept. The Business intelligence, financial data analysis, reporting, integration platform and data analytics were in open source scheme. Here it is easy to extend, add plug-in and to include the core version of data integration modules.

Konstanz Information Miner, is an open source data analytics, reporting and integration platform. It has been used in pharmaceutical research, but is also used in other areas like CRM customer data analysis, business intelligence and financial data analysis. It is based on the Eclipse platform and, through its modular API, and is easily extensible. Custom nodes and types can be implemented in KNIME within hours thus extending KNIME to comprehend and provide first-tier support for highly domain-specific data format. It is released on 2004. The latest version available is KNIME2.9 and licensed By GNU General Public License. Compatible with Linux ,OS X, Windows and written in java.

The general features are it is designed data mining tool that runs inside the IBM's Eclipse development environment. It is a modular data exploration platform that enables the user to visually create data flows (often referred to as pipelines), selectively execute some or all analysis steps, and later investigate the results through interactive views on data and models. The Knime base version already incorporates over 100 processing nodes for data I/O, preprocessing and cleansing, modeling, analysis and data mining as well

as various interactive views, such as scatter plots, parallel coordinates and others.

Integration of the Chemistry Development Kit with additional nodes for the processing of chemical structures, compounds, etc. Specialized for Enterprise reporting, Business Intelligence, data mining. The main advantages are it integrates all analysis modules of the well-known. Weka data mining environment and additional plugins allow R-scripts to be run, offering access to a vast library of statistical routines. It is easy to try out because it requires no installation besides downloading and un archiving. The one aspect of KNIME that truly sets it apart from other data mining packages is its ability to interface with programs that allow for the visualization and analysis of molecular data.

The Limitations are having only limited error measurement methods. Have no wrapper methods for descriptor selection. Does not have automatic facility for Parameter optimization of machine learning/statistical methods.

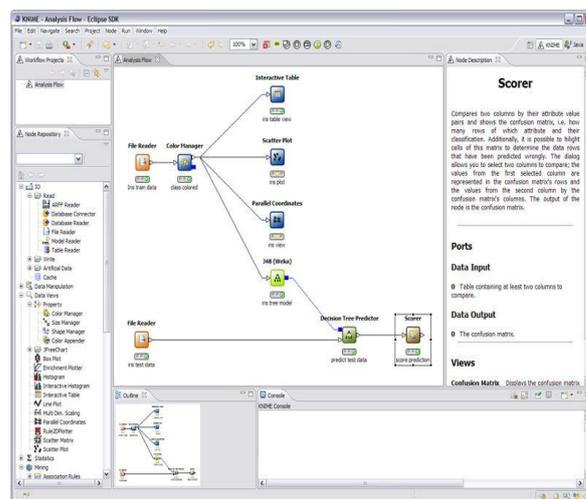


Fig 5. Flow diagram for KNIME

CONCLUSION

The development of Educational data mining techniques makes explosive growth in the field of education. The most promising tools and techniques in educational data mining for future are discussed in this paper. Now a days the academic success of students is the major issue for the management in all professional institutes. So the early prediction to improve the student performance through counseling and extra coaching will help the management to take timely action for decrease the percentage of poor performance by the students. The classification and clustering operations are used to predict more accurate results for improve the level of success rate of the students in higher education. To develop a decision support system and help the authorities to timely

actions on weak students these Educational data mining techniques were used.

REFERENCES

- [1] K. R. Kavyashree, LakshmiDurga – “A Review on Mining Students Data for Performance Prediction” - International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 4, April 2016
- [2] TriptiDwivedi, Diwakar Singh – “Analyzing Educational Data through EDM Process: A Survey” - International Journal of Computer Applications (0975 – 8887) Volume 136 – No.5, February 2016
- [3] Abdulmohsen Algarni - “Data Mining in Education” - (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 6, 2016
- [4] G. Gray, C.McGuinness, P.Owende – “An Application of Classification models to predict learner progression in tertiary education” - Advance Computing Conference (IACC), 2014 IEEE International, IEEE, 2014, pp. 549–554
- [5] Amirah Mohamed Shahiria, WahidahHusaina and Nuraini Abdul Rashida – “A Review on Predicting Students Performance using Data Mining Techniques” - The Third Information Systems International Conference, Procedia Computer Science 72(2015) 414 – 422
- [6] Heba Mohammed Nagy, Walid Mohamed Aly, Osama FathyHegazy – “An Educational Data Mining System for Advising Higher Education Students” - International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:7, No:10, 2013
- [7] Monika Goyal and Rajan Vohra - “Applications of Data Mining in Higher Education” - IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012
- [8] U. K. Pandey and S. Pal - “A Data mining view on class room teaching language” International Journal of Computer Science Issue, Vol. 8, Issue 2, pp. 277-282, ISSN:1694- 0814,2011
- [9] Dr. MohdMaqsood Ali - “Role of data mining in education sector” - International Journal of Computer Science and Mobile Computing Vol. 2, Issue. 4, April 2013
- [10] Dr. P. Nithya, B. Umamaheswari, A. Umadevi – “A Survey on Educational Data Mining in Field of Education” - International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 5 Issue 1, January 2016
- [11] KalpanaRangra, Dr. K. L. Bansal-“Comparative Study of Data Mining Tools” -International Journal of Advanced Research in Computer Science and Software EngineeringVolume 4, Issue 6, June 2014