**STUDENT PERFORMANCE PREDICTION USING DATA MINING**

**Final project**

**By:**

**Ashish Paul Stephen (805663515)**

**Mohit Kirange (805600418)**

[**CPSC 531-02 13541**](https://moodle-2016-2017.fullerton.edu/course/view.php?id=9524)

**Fall 2016**

**Professor: Dr. Chun-I Philip Chen.**

**Department of Computer Science**

**California State University, Fullerton**

**December 15, 2016**

**Table of contents**

Abstract…………………………………………………………………..………………………4

Introduction………………………………………………………………..……………………..5

Definition of Terms…………………………………………..…………………………………..6

Literature Review……………………………………………………………….………………..8

Methodology……………………………………………………………………………………..12

Results and Discussion…………………………………………………………………………..25

Conclusions, Implications and Recommendations……………………………………….……..30

References……………………………………………………………………………………….32

**List of Figures**

Figure 1: The steps of extracting knowledge from data ………..….……………………………..7

Figure 2: The cycle of applying data mining in educational systems ……………………………9

Figure 3: Research Framework…………………………………………………………………..16

Figure 4: System Architecture……………………………...…………………………. ………..21

Figure 5: Homepage screenshot …………………………………………………………………23

Figure 6: Collection of user data screenshot……………………………………………………..24

Figure 7: Predicted Grades ………………………………………………………………………25

Figure 8: Result by Naïve Bayes algorithm……………......…………………………………….26

Figure 9: Result by J48 algorithm…………….....………………………………………………27

Figure 10: Result by JRip algorithm…………………………………………….………………28

Figure 11: Decision Tree Logic Tree of the training data ……………………….……………29

**List of Tables**

Table 1: Description Of Attributes In The Dataset ……………………………………………..13

Table 2: List of different software available for data mining analysis…………………………..16

Table 3: Software requirement for Weka.………………..……………………………………21

Table 4 : Table of accuracy……………………………………………………………………25

**Abstract**

Role of education has always been a critical topic to deal with the development of any country. Now a day education system gathers and generates a lot of information and data related to the student. We can make use of this data to redeem hidden information such as pattern out of it, this can be made possible by making use of various data mining techniques and with the help of this information we can review our educational process and make certain improvement in our system.

Although data mining has been used effectively in business world, its use in education is still low, as this is used to identify and extracting new and potentially valuable knowledge from the data. Using the techniques of data mining, the aim was to develop a model/system that had the capability of predicting the student’s success. In this paper, we make use of data mining techniques to analyze grades of students based on evaluative assignments for the courses taken by them. For this purpose, we put into comparison different classifiers using the dataset consisting of various student records and use it to predict the performances of them. We apply different techniques of classification techniques on both numerical as well as categorized attributes. Different techniques and methods of data mining were put to comparison during the prediction of student’s success as the result obtained by one would not be able to provide the whole picture. That is the reason we made use of different classifiers such as J48, JRip, Naïve Bayes to predict the performance of the student in the final examination. For the analysis, we would be making use of a sample data set consisting of student data. On the other hand, we can improve the performance of the student based on the result obtained as this is taken and analyzed for the next semester. With this result, we can improve the performance of the student who are on the risk of failure and increase the chances of performing much better in the next examinations.

*Keywords: Educational Data Mining, Classification, Prediction, Naive Bayes, JRip*

**Introduction**

Educational data available for analysis nowadays continues to grow rapidly. The need to

study the massive amounts of data generated from the global educational ecosystem has spawned the field of (EDM). EDM is the process of applying data mining tools and techniques to analyze data in educational institutions. The application of data mining techniques to educational data will help the educational sector improve its learning process to better benefit students. Example of data stored in databases maintained by educational institutions includes: enrollment data, students’ performances, teachers’ evaluations, gender differences, and numerous other measures.

EDM can help universities better plan for the anticipated number of students enrolling in their programs, predict the dropout ratio, easily identify weak students, and make better use of available resources ranging from the number of faculties to the utilization of other resources. In addition, higher education assesses its students’ performances and progresses to improve its academic programs. More recently, literature has evolved in the field of educational data mining because of its potential benefits to education. Once these data have been correctly analyzed, they will help advance knowledge in the educational sector. This will help educational institutions assess, evaluate, plan, and decide their educational programs. This new knowledge is expected to reveal hidden patterns that will assist academic programs utilize resources more

effectively.

Academic programs typically evaluate students based on inner and outer assessment. Inner assessment consists of activities carried out during the semester, such as quizzes, midterms, projects, lab work, etc. Outer evaluation is based on student’s final scores. The assessment of students is based on examinations, activities, and assignments during their courses. A student passes or fails a course based on the total points scored in a semester. Hence, the ability to specify prior to the final exam the students who are likely to fail a course can prompt additional remedial efforts by both the teacher and the student to improve their performance. These efforts can help struggling students succeed in their courses. The evaluation can be applied by usual method of calculating the total points towards the end of the semester. Or it can be applied by data mining techniques. Various data mining techniques can be applied to educational report and information, and classification is one of them. Classification is a supervised learning technique that builds a model to classify a data item according to a predefined class label.

The aim of classification is to predict future output based on available data. Classification can be used to predict students’ performances in the most critical courses, such as programming courses in computer science. One of the most important courses in Computer department, MumbaiUniversity is the “Data Structures,” which has a high rate of failure. We gathered the data from two consecutive semesters in academic year of 2015 - 2016 of 200 students were enrolled in the courses ‘Data Structure’ and ‘Computer Architecture and organization’. Since data structure course, has a high failure among the CS students. we decided to use classification techniques to study students’ performance. We will focus on students enrolled in the "Data Structures" course to answer the following questions:

- How to predict student performance on the final exam?

- How to forecast the total points obtained by each student toward the end of the course?

- Focus on failing students and how to predict student’s failure? (mashael a. Al-barrak & mona s. Al-razgan, 2015)

**Definition of terms**

**Data mining-** Data mining is an interdisciplinary subfield of computer science. It is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Aside from the raw analysis step, it involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating. Data mining is the analysis step of the "knowledge discovery in databases" process, or KDD. (Wikipedia,2016)

**Educational data mining(EDM)-** Educational Data Mining (EDM) is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in. Educational Data Mining focuses on developing new tools and algorithms for discovering data patterns. EDM develops methods and applies techniques from statistics, machine learning, and data mining to analyze data collected during teaching and learning. EDM tests learning theories and informs educational practice. (educationaldatamining, 2016)

**Classification**- Classification is a data mining function that assigns items in a collection to target categories or classes. The goal of classification is to accurately predict the target class for each case in the data. For example, a classification model could be used to identify loan applicants as low, medium, or high credit risks. (Oracle,2016)

**Prediction-** Prediction is like classification but the output here is continuous and not discrete. It can construct a model and use that model to predict continuous output value for the given data and input. Usually prediction involves methods such as regression where many variants of the regression analysis contribute in predicting unknown or missing values.

**Literature Review**

The data mining has attracted a great deal of attention in the information technology industry, due to availability of large volume of data which is stored in various formats like files, texts, records, images, sounds, videos, scientific data, and many new data formats. There is imminent need for turning such huge data into meaningful information and knowledge. The data collected from various applications require a proper data mining technique to extract the knowledge from large repositories for decision making. Data mining, also called Knowledge Discovery in Databases (KDD), is the field of discovering novel and potentially useful information from large volume of data. Data mining and knowledge discovery in databases are treated as synonyms, but data mining is a step in the process of knowledge discovery. The sequences of steps identified in extracting knowledge from data are shown in Figure. 1.

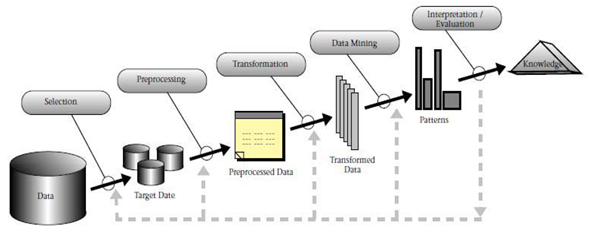


Figure 1: The steps of extracting knowledge from data (Barahate Sachin & Shelake Vijay, 2012)

The main functionality of data mining techniques is applying various methods and algorithms to discover and extract patterns of stored data. These interesting patterns are presented to the user and may be stored as new knowledge in knowledge base. Data mining and knowledge discovery applications have gotten a rich focus due to its significance in decision making.

Data mining has been used in areas such as database systems, data warehousing, statistics, machine learning, data visualization, and information retrieval. Data mining techniques have been introduced to new areas including neural networks, patterns recognition, spatial data analysis, image databases and many application fields such as business, economics, and bioinformatics.

The main objective of this paper is to survey the applications of data mining techniques and algorithms to traditional educational system. The first section is used to describe the history and current trends in the field of Educational Data Mining (EDM). We then review the different data mining techniques that have been applied in educational field grouping them by task. Finally, we discuss and summarize key applications of educational data mining. (Barahate Sachin & Shelake Vijay, 2012)

**Educational data mining**

The educational data mining community defines educational data mining as, “Educational Data Mining (EDM) is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the setting which they learn in”.

There are increasing research interests in using data mining techniques in educational field. This new emerging field, EDM, concerns with developing methods that discover knowledge from data originating from educational environments. Educational data mining techniques often differ from traditional data mining techniques, in explicitly exploiting the multiple levels of meaningful hierarchy in educational data. EDM focuses on collection, archiving, and analysis of data related to students learning and assessment. The analysis performed in EDM research is often related to

techniques drawn from variety of literatures, including psychometrics, machine learning, data mining, educational statistics, information visualization and computational

modeling. The application of data mining in educational system is an interactive cycle of hypothesis formation, testing, and refinement as shown in Fig. 2. Discovered knowledge should enter the loop of the system and guide, facilitate, and enhance learning. The system is used to turning data into knowledge as well as filtering mined knowledge for decision making. . (Barahate Sachin & Shelake Vijay, 2012)

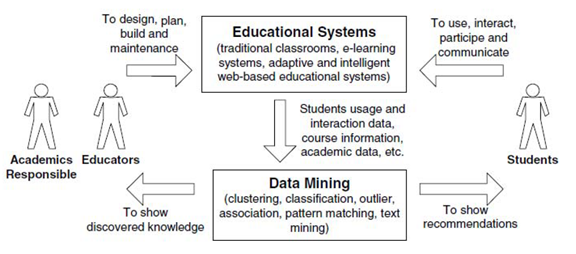


Figure 2: The cycle of applying data mining in educational systems (Barahate Sachin & Shelake Vijay, 2012)

**EDM Methods**

Romero and Ventura, and Baker categorize methods in educational data mining into the following general categories:

• Prediction

• Clustering

• Relationship mining

Romero and Ventura’s viewpoint is focused on applications of educational data mining to data. These methods are listed as web mining methods, and are quite prominent in mining web data and in mining other forms of educational data. These categories of educational data mining methods are largely acknowledged to be universal across types of data mining.

**Prediction**

The goal of prediction is to develop a model which can infer a single aspect of the data (predicted variable) from some combination of other aspects of data (predictor

variables). Prediction has two key uses within educational data mining. In first type, prediction methods can be used to study what features of a model are important for prediction, giving information about underlying construct. This approach is basically used to analyze the student's’ performance. In second type of usage, the prediction methods are used to predict what the output value would be in contexts where it is not desirable to directly obtain a label for

that construct. Prediction can be classified in three types viz. **Classification, Regression, and Density estimation.**

A classifier is mapping from (discrete or continuous) feature space X to discrete set of labels Y. Classification or discriminant analysis is used to predict class labels (describes future situation). Classification is supervised technique which is used to label newly encountered (still unlabeled) patterns from a collection of labeled (pre-classified) patterns. Some popular classification methods include logistic regression, support vector machines and decision trees. The decision tree algorithm produces a treelike structure of the model it produces. From the tree, it is then easy to generate rules in the form IF condition THEN

outcome. It is basically a predictive model in which an instance is classified by following the path of satisfied condition from root until reaching a leaf, which will correspond to class label. Some of the most well-known decision tree algorithms are C4.5 and ID3. (Barahate Sachin & Shelake Vijay, 2012)

**Problem related to the higher education system**

At present, most of the institutions or organization in India are facing the problem of student admission. Most of the engineering college or university are face problem of low admission in engineering stream. There are lot of reason for that like less placement record, less infrastructures; syllabus not updated, less qualified staff, poor teaching methodology. So, to increase education system in the college we need to provide these basic needs of the time. Without providing these features no college will sustain soon and face the problem of failure. So, to remain in the competition with other college they need to provide extra to the student which helps them a lot in their study. Educational data mining is the solution of the entire problem because with the help of educational data mining we can analysis the all the data which are produced by the educational setting.

With the help of analysis, we can predict the result of the student, dropout of any student, placement of the student, behavior of the student etc. If any student having a risk of failure and we can predict that risk in advance, then we can provide timely help to that student. Education data mining techniques can be applied on any types of educational data. There are lots of data mining techniques which are applied on educational data like classification and clustering algorithm. (Mukesh Kumar, Prof (Dr.) A. J. Singh, 2016)

**Methodology**

**Technical Approach** - **Building the models**

**Data Collection**

Student’s data for the course “Data Structures” and “Computer organization and architecture” was collected from the Computer department at Mumbai University, India, for two consecutive semesters. A total of 225 records were initially collected for each. Having deleted data for withdrawn students, the remaining number of records was 200. Each student record had the following attributes: student ID, student name, student grades in quiz1, quiz2, and quiz3, midterm1, midterm2, project, tutorial, final exam, and total points obtained. The distribution of points for the course was 60 for a year’s work (which included quizzes {1, 2, 3}, midterms {1, 2}, tutorials, the project and the final lab) and 40 points for the final exam. A student must have obtained at least 60 out of 100 to pass the course.

**Data Preparation and Pre-processing**

Data pre-processing is among the common steps prior to applying any data mining technique. We applied the following steps to prepare the data:

* Eliminating the records of students who withdrew from the course because some of their relevant values were consequently missing.
* Discretizing the total grade attribute into five categories: A, B, C, D, and F.
* Discretizing all attributes of the semester into four categories: excellent, good, average, and poor.

After pre-processing the data, we ran the Waikato Environment for Knowledge Analysis

(Weka) toolkit to apply the classification algorithms. Weka was developed at the University of Waikato in New Zealand, and is very popular data mining software that contains a wide range of algorithms implemented in Java.

For our research purposes, we used two forms of data: real numerical values of all attributes except the class attribute under examination (final exam or total) to be categorized, and the second one after categorizing all remaining attributes. The description of the attributes in the final data is listed in table 1.

|  |  |  |
| --- | --- | --- |
| Attribute | Description | Possible Numerical Values |
| Quiz 1, 2, 3 | Quiz 1, 2, 3 grade | Real numbers from 0-5 |
| Mid 1, 2 | First and second midterm exam grade | Real numbers from 0-10 |
| Project | Project grade | Real number from 0-10 |
| Tutorial | Tutorial grade | Real number from 0-5 |
| Final Lab | Final laboratory exam grade | Real number from 0-10 |
| Final Exam | Final examination grade | Real number from 0-40 |
| Total | Total points obtained by the student that determines his/her success or failure in the course | \*Used as categorized data in the experiment |

Table 1: Description Of Attributes In The Dataset (mashael a. Al-barrak & mona s. Al-razgan,

2015)

**Data Visualization**

Having applied the pre-processing techniques to the dataset, we loaded it to the Weka software. We first attempted to analyze the data through visualization prior to applying any classification algorithm. Specifically, we wanted to observe the distribution of the grades in the final exam as well as the total points. final exam points alone cannot be used as an indicator of the success or failure in the course. Prior to the final exam, most students must have use of what they learned during the semester through tutorials and projects. This was reflected in their performance in the final test. However, passing this course depended also on students’ performance in all the other methods of assessment during the semester.

**Classification**

Classification is the process of placing an object into a class or category. This process involves two primary steps:

1. Learning, where a set of training data is analyzed to build a model with a

predetermined set of classes.

2. Testing, where a different set of data is tested to determine the accuracy of the model.

Our goals in this research are to predict student’s final exam performances and the outcome of the class for the course. Classification is used as a predictive data mining technique that predicts the class of an attribute based on the value Journal of other attributes (in our case, student’s final exams and the outcome of the course). Many classification algorithms can be applied to an educational dataset, such as decision trees, neural networks, Naïve Bayes, and rule-based algorithms. We attempted to classify students’ performances using three classifiers: C4.5 decision tree (J48 in WEKA), Naïve Bayes, and the JRip rule-based algorithm.

**Research Framework**

We designed a methodological framework to organize the steps of our procedure. This methodology, which was used to find the best model to predict students’ final exam grades and the outcome of the course. We ran the three classification algorithms for the two forms of the dataset (numerical and categorical) several times to predict final exam performance and final course outcome. We examined the accuracy of the resulting models using 10-fold cross-validation.



Figure 3: Research Framework (mashael a. Al-barrak & mona s. Al-razgan, 2015)

**Proposed Methodology**

**Tools available for data mining analysis**

At present scenario, data is one of the most important in today’s world. Because by analyzing that data we can find some information which will be helpful in future. We have different types of data mining software for analysis. Every organization deals with different types of data in real life like data related to education, business, sales, marketing, hospital, hospitality etc. Software’s has their own features and properties and it depend on the data that which software is suitable for their analysis [6]. Here we present ten most important tools used for the data analysis in tabular form below:

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No** | **Software** | **Language used** | **Developed State** |
| 1 | RapidMiner | Java | Technical University of Dortmund |
| 2 | SAS Data Mining | C | North Carolina State University |
| 3 | WEKA | Java | University of Waikato, New Zealand |
| 4 | R-Software | C, Fortran, R | University of Auckland, New Zealand |
| 5 | Orange | Python | University of Ljubljana |
| 6 | KNIME | Java | University of Konstanz |
| 7 | NLTK | Python | University of Pennsylvania |
| 8 | DataMelt | Jython, Groovy | jWork.ORG community |
| 9 | Pentaho | Java | Hitachi Data Systems |
| 10 | Tanagra | DELPHI 6 | Lumière University Lyon, France |

Table 2 :List of different software available for data mining analysis

After reading different research paper about educational data mining we find that Rapid Miner and WEKA are the mostly used software for the analysis purpose. So, form the above discussion we are taken WEKA software tool for our analysis purpose. WEKA is an Open source software and easily available for the user under GNU public license. We can also implement our own algorithm on this software. Most of the data mining algorithms are available in WEKA software. WEKA is a complete package of different data mining or machine learning algorithm. It supports classification, clustering, regression, association rule and feature selection algorithm. It also able to shows you various relationships between data sets, cluster, visualization, predictive modelling, and association rule algorithms (Mukesh Kumar, Prof (Dr.) A. J. Singh, 2016)

**Classification techniques taken into consideration**

We have different types of data mining algorithms are available to make an analysis of our data like clustering, classification, association rule mining. But which data mining algorithm is suitable for your data is depend upon what types of information you want to take and what types of data set you have in your hand. Before selecting any algorithm make sure that what types of information you want to take from the dataset. Every data mining model is created with the help of a specific algorithm. We can solve any data mining problem with best possible way by using more than one algorithm. In this article, we want to make a prediction related to the result of the student in the coming semester. You will be the successful at data mining field even if you are not very much familiar with the inner working of each algorithm. But it is important to get the full understanding of the general features of each algorithm and their suitability with different dataset. Many classification algorithms can be applied to an educational dataset, such as decision trees, neural networks, Naïve Bayes, and rule-based algorithms. We attempted to classify students’ performances using three classifiers: **C4.5 decision tree (J48 in WEKA), Naïve Bayes, and the JRip rule-based algorithm.**

**J48-**Decision trees are powerful and popular tools for classification. A decision tree is a tree-like structure, which starts from root attributes, and ends with leaf nodes. Generally, a decision tree has several branches consisting of different attributes, the leaf node on each branch representing a class or a kind of class distribution. Decision tree algorithms describe the relationship among attributes, and the relative importance of attributes. The advantages of decision trees are that they represent rules which could easily be understood and interpreted by users, do not require complex data preparation, and perform well for numerical and categorical variables. The WEKA J48 classification filter is applied on the dataset during the experimental study. It is based on the C4.5 decision tree algorithm, building decision trees from a set of training data using the concept of information entropy.

**Naïve bayes-**Bayesian classifiers are statistical classifiers that predict class membership by probabilities, such as the probability that a given sample belongs to a class. Several Bayes’ algorithms have been developed, among which Bayesian networks and naive Bayes are the two fundamental methods. The Naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' Theorem with strong independence assumptions which assumes all the features are equally independent. Naive Bayes algorithms assume that the effect that an attribute plays on a given class is independent of the values of other attributes. However, in practice, dependencies often exist among attributes; hence Bayesian networks are graphical models, which can describe joint conditional probability distributions. Bayesian classifiers are popular classification algorithms due to their simplicity, computational efficiency, and very good performance for real-world problems. Another important advantage is also that the Bayesian models are fast to train and to evaluate, and have a high accuracy in many domains. In simple terms, a naive Bayes classifier assumes that the presence (or absence) of a feature of a class is unrelated to the presence (or absence) of any other feature.

**JRIP**-JRip (RIPPER) is one of the basic and most popular algorithms. Classes are examined in growing size and an initial set of rules for the class is generated using incremental reduced error JRip (RIPPER) proceeds by treating all the examples of a decision in the training data as a class, and finding a set of rules that cover all the members of that class. Thereafter it proceeds to the next class and does the same, repeating this until all classes have been covered. (Dr. Vaishali Parsania, Dr. N. N. Jani, Navneet H Bhalodiya, 2014)

**Design and functional requirements**

The data mining tool should able to take any size of dataset as input as this is the requirements. It should be able to process any dataset irrespective of its size and it should be able to apply any type of classification technique efficiently. Its execution time should be minimum, this thereby enhances accuracy as well as efficiency. The tool should be able to visualize the output processed data in any graphical manner.

**Data requirement**

The main aim of classification is to predict future output based on available data. Classification can be used to predict students’ performances in the most critical courses, such as programming courses in computer science.

**Source of Data**

Student’s data for the course “Data Structures” and “Computer Architecture and organization” was collected from the Computer department at Mumbai University, India, for two consecutive semesters in academic year of 2015-2016. A total of 225 records were initially collected. Having deleted data for withdrawn students, the remaining number of records was 200. Input data of each student record had the following attributes: student ID, student name, student grades in quiz1, quiz2, and quiz3, midterm1, midterm2, project, tutorial, final exam, and total points obtained. The distribution of points for the course was 60 for a year’s work (which included quizzes {1, 2, 3}, midterms {1, 2}, tutorials, the project, and the final lab) and 40 points for the final exam. A student must have obtained at least 60 out of 100 to pass the course. Numerical data consists of quiz (1, 2, 3), midterm (1, 2), project, tutorial, final lab, and final exam grade and the categorized dataset consists of all grades discretized into four categories: excellent, good, average, and poor.

**System requirements**

**Hardware requirement**

Minimum: Dual core, 2GHz processor, 4GB RAM, >1GB free disk space, 1280x1024

Operating System: Windows, Linux, MacOS X 10.8 or newer

**Software requirement**

The following matrix shows what Java version is necessary to run a specific Weka version. It also specifies when a changeover happened, listing date and [Subversion](http://weka.wikispaces.com/Subversion) revision (the number prefixed with *r*).

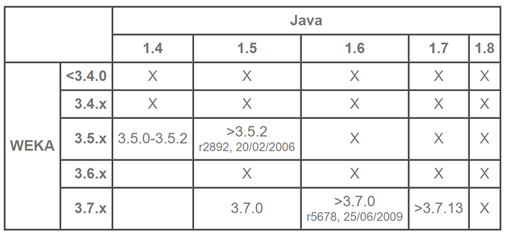
****

Table 3 : Software requirement for weka

**System architecture**

A total of 6 steps are taken into consideration for the system architecture. The flow of the architecture is shown in the following diagram where it starts from the phase of collecting data and it goes on to the next phase where we process and prepare the data for the next phase i.e data visualization. The architecture ends up with the phase of developing a model that accurately predicts the performance of the student.

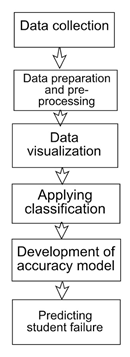


Figure 4: System Architecture

**System installation and implementation:**

During the implementation of our project, we developed the Website application to predict the students grade named as ‘FutureGrade’. The tool Visual Studio is used to develop the website and connecting with the database. In Visual Studio website is developed with help of Bootstrap which the most popular HTML, CSS, and JS framework for developing responsive, mobile first projects on the web and C# for the back-end coding.

**Home Page:** Currently website introduces the EDM by Mumbai university and allows grade prediction for two subjects i.e. ‘Data Structure’ and ‘Computer Organization and architecture’

And also, gives option to the user to see the graphical representation of previously collected data for subjects. Website is created with consideration of user friendliness.

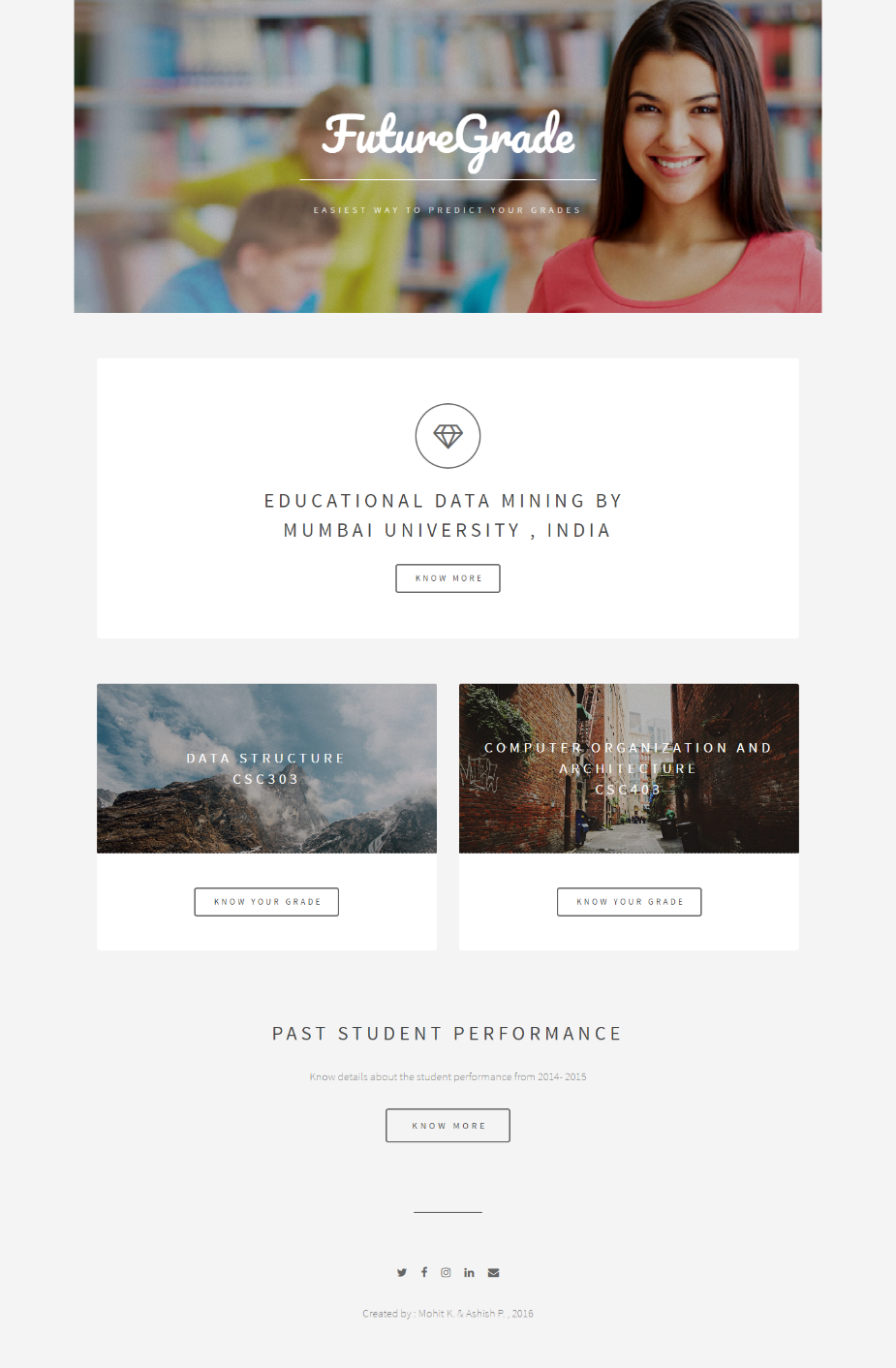


Figure 5 : Homepage Screenshot

**Collection of user data:** After clicking on the ‘know your grade’ new Subject page opens and user is asked to enter his available academic data into the form to predict his grades. Subject page also allows user to see details of subject as course objectives and outcomes.

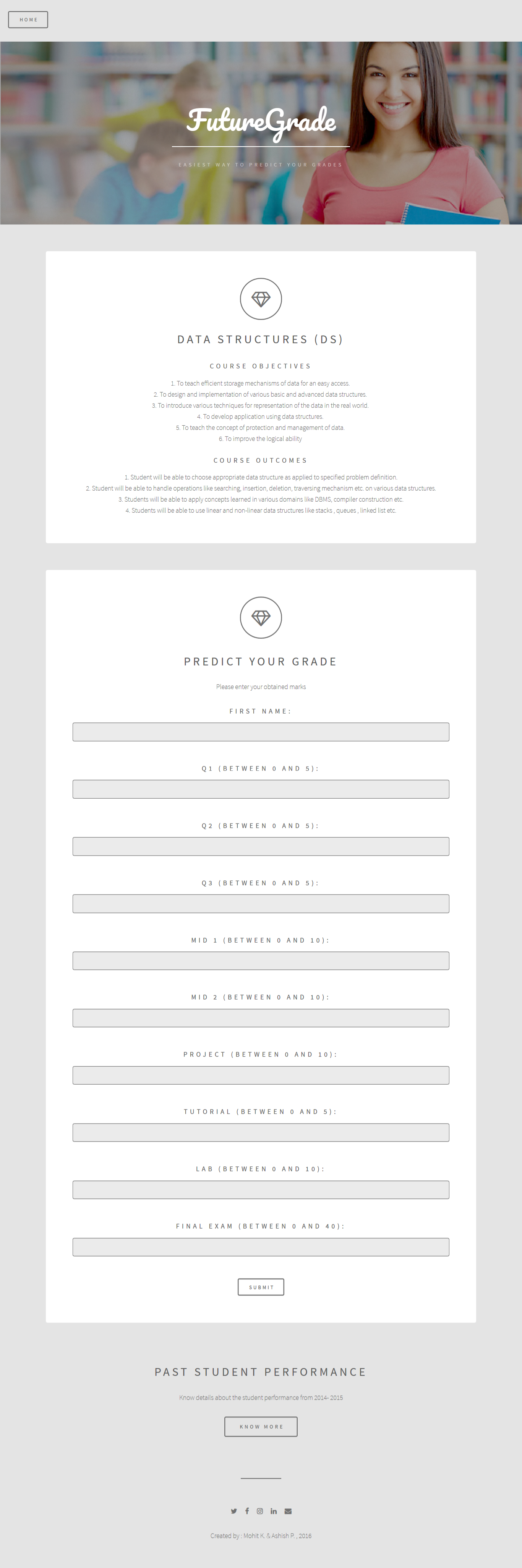


Figure 6 : Collection of user data Screenshot

**Predicted grade:** Based on user submitted data the website used the prediction logic which is implemented into the website to predict the possible grade. Hence by grade we can identify whether user is passed or failed. The logic used is extracted unknown pattern identified by applying various classification algorithms in data mining on the training dataset with help of weka.

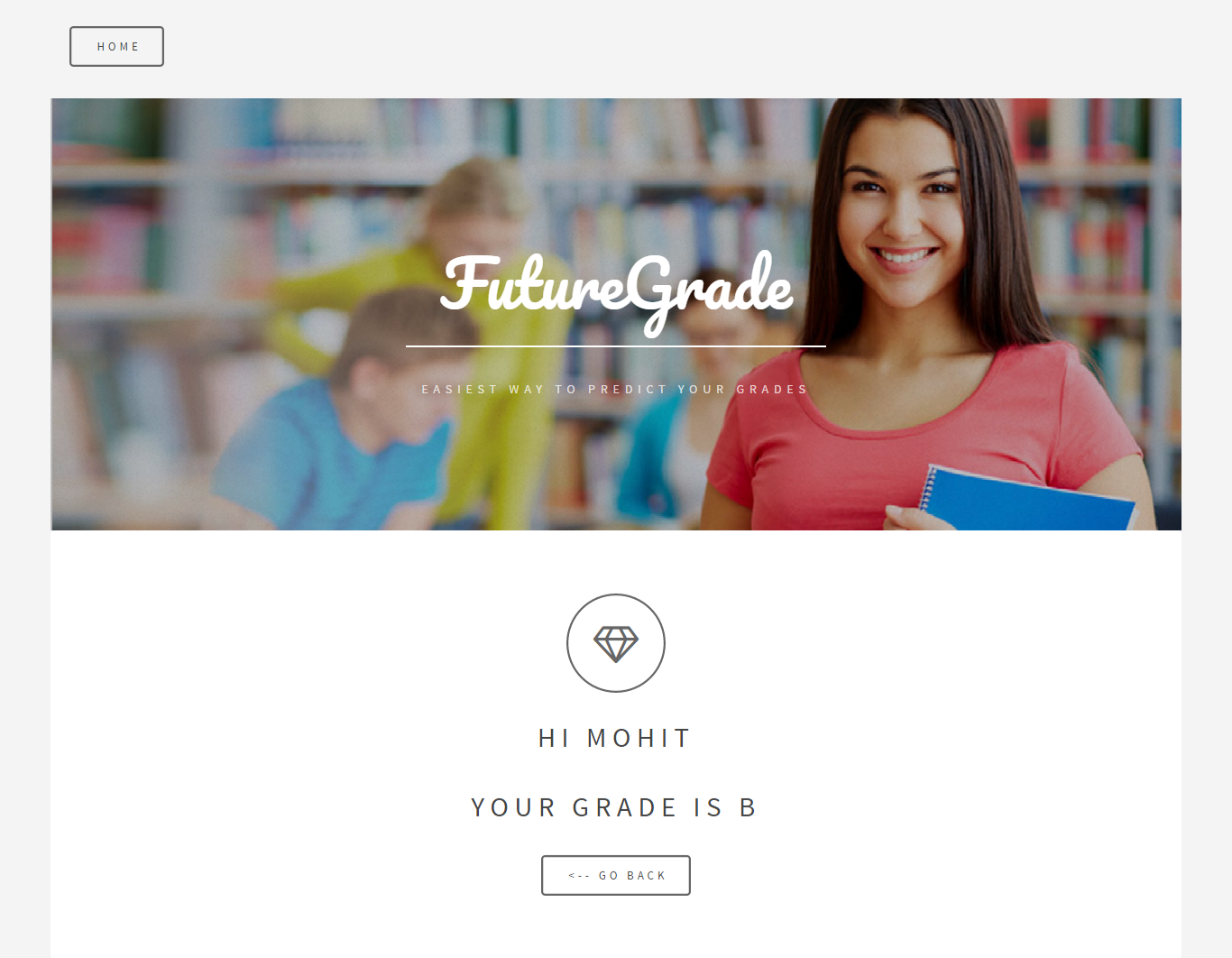


Figure 7 : Predicted Grades

The Weka tool is used to apply various classification algorithms on the training and test dataset like naïve bayes , J48 and JRip. Their results are discussed in next section.

**Results and Discussion**

To perform the student grade prediction, we firstly used the WEKA tool to extract the unknown pattern i.e. logic to be used in the website for prediction of the grade. In weka we applied 3 classification algorithms which are naïve bayes , J48 and JRip on the same training dataset which was collected. To analyse which algorithm gives the best results. The analyzation was done based on the accuracy of the output.

|  |  |  |
| --- | --- | --- |
| Algorithms | Correctly Classified Instances | Incorrectly Classified Instances |
| Naïve Bayes | 93.5% | 6.5% |
| J48 | 98.5% | 1.5% |
| JRip | 99% | 1% |

Table 4 : Table of accuracy

Hence based on comparison of the accuracies, JRip algorithm gives the 99% of accuracy which is better than other two algorithms. Hence Logic Model of Jrip Algorithm is used in

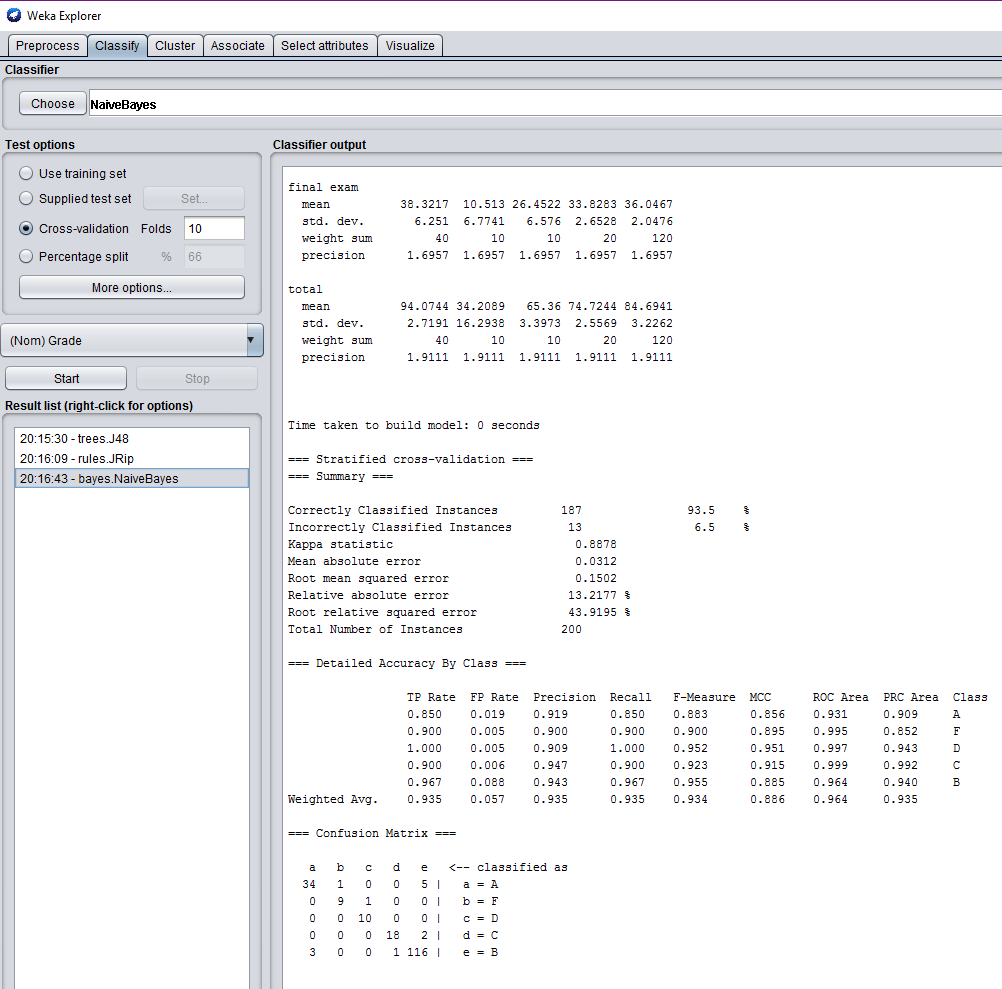


Figure 8 : Result by Naïve Bayes algorithm

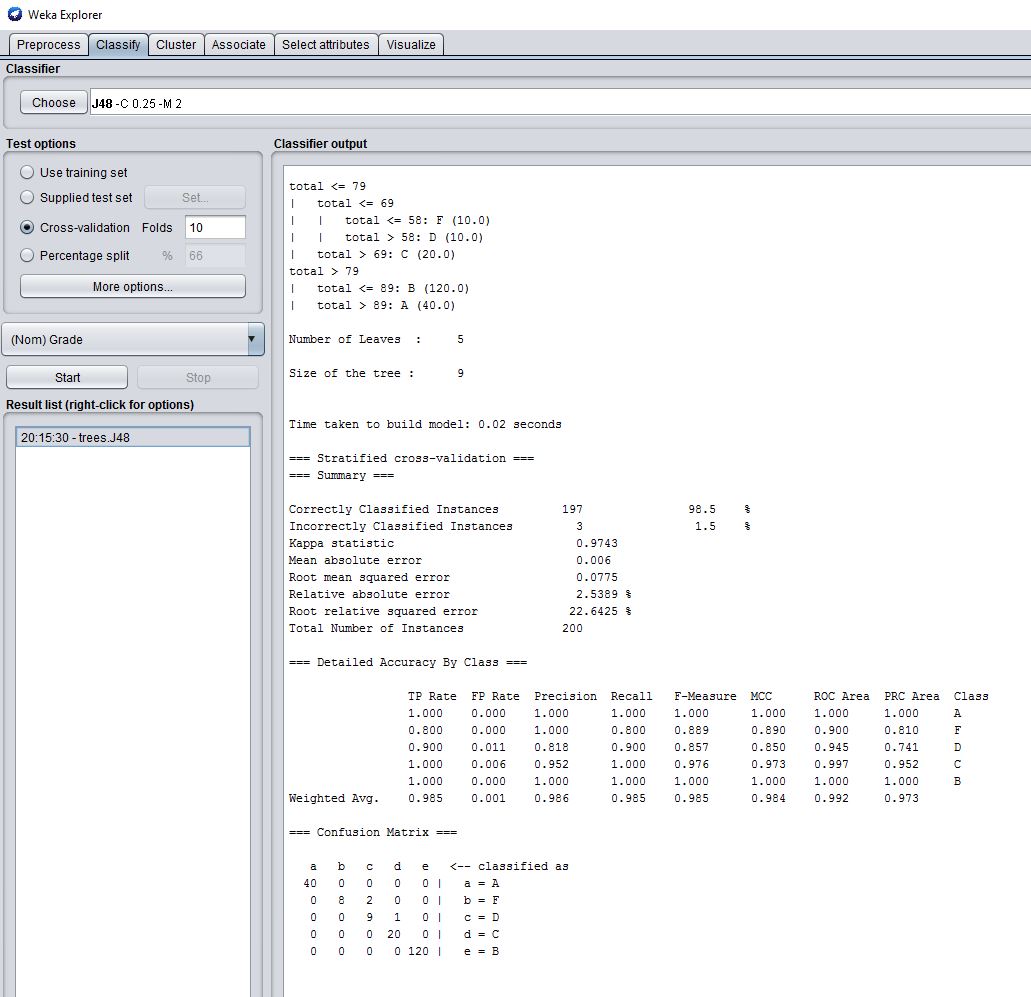
****

Figure 9 : Result by J48 algorithm

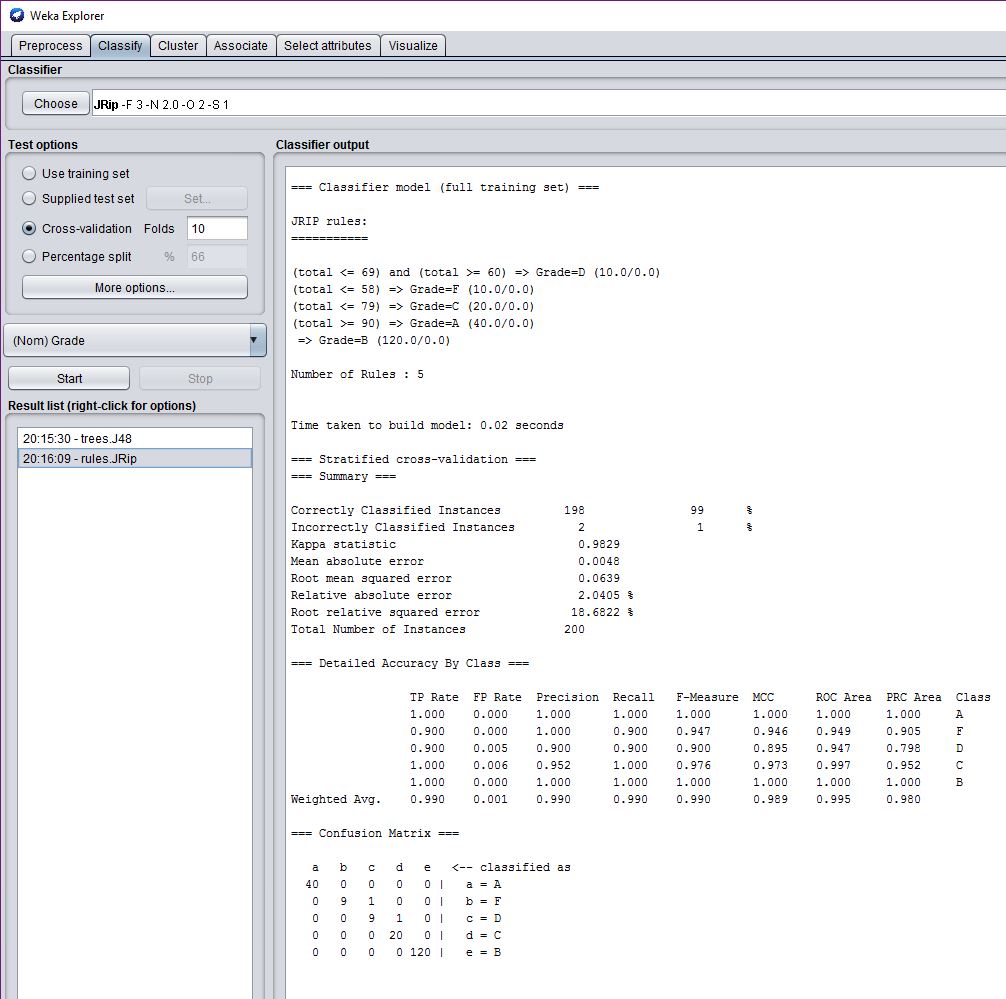


Figure 10 : Result by JRip algorithm

With help of algorithm J48 i.e. decision tree we can visualize the logic of grades as follows in the weka. This logic for the grades is implemented into the ‘FutureGrade’ website which is similar to the model used from JRip.

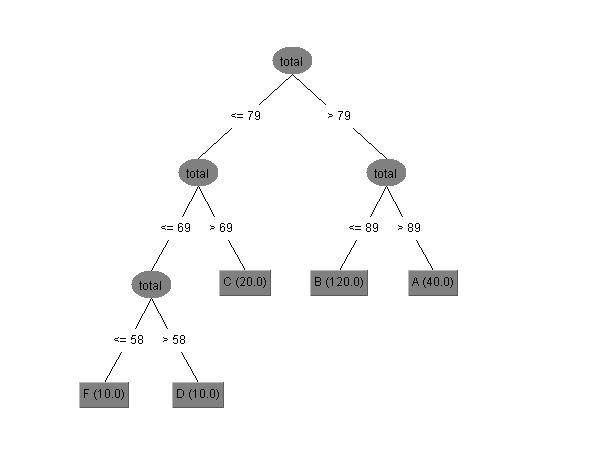
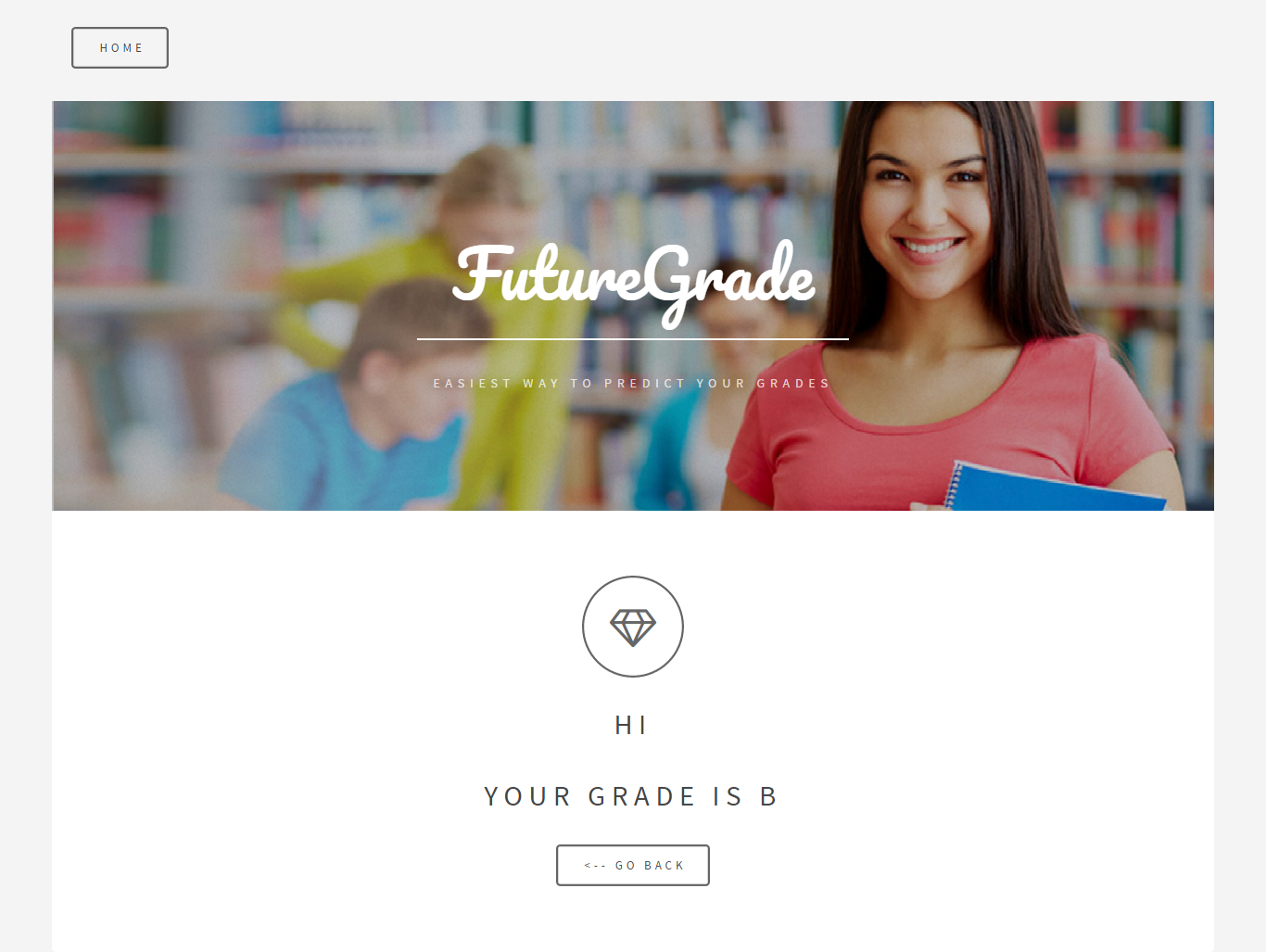


Figure 11 : Decision Tree Logic Tree of the training data

With help of this extracted pattern by applying the data minig we can predict the grade of student based on his entered attribute (Marks ) values . There are total 5 level Predicted grades : A,B,C,D,F.



**Predicted Grade**

**Conclusions, Implications and Recommendations**

**Conclusion**

Classification is one of the most interesting and important topic of data mining techniques. Most of the researchers in this field are using classification algorithm of data mining for knowledge discovery from the dataset. In this project, we have implemented the educational data mining by predicting the student’s final grades to help improve the quality of the education imparted to them. We tested the accuracies of different classification algorithms like Naïve Bayes, J48 , JRip in prediction of student’s grades in for two specific courses ‘Data Structure’ and ‘Computer Organization and Architecture’ based on their marks in the various exams throughout the semester. We tested the models to determine algorithms accuracy and interpreted that JRip algorithm gives 99% of accuracy but J48 i.e. Decision tree algorithm gives the decision tree login which is the extracted pattern used for prediction. We achieved the goal of our project i.e. to developed the model using data mining to predict the students grades by extracting the pattern from available training set. Hence, we can predict the future problems and find solution on them. Moreover, this experiment can be applied to more than another course, such as programming courses, to analyze student’s behavior and predict their performance.

**Implications and Recommendations**

For future work, the experiment can be extended with more distinctive attributes to get more accurate results, useful to improve the students learning outcomes. Also, experiments could be done using other data mining algorithms to get a broader approach, and more valuable and accurate outputs. Some different software may be utilized while at the same time various factors will be used. Also, introducing the advance algorithms, optimizing the system architecture, gathering more training samples, and using temporal information in the sequential data will help advancing the education field. We might, of course, encounter some problems, such as data collection and its processing as EDM field is still evolving. But we hope that this review will be able to shed some useful insights for researchers and educators in order for educational data mining to become a mature area.

**Reference**

Barahate Sachin & Shelake Vijay, (2012)

R. B. Sachin and M. S. Vijay, "A Survey and Future Vision of Data Mining in Educational Field," 2012 Second International Conference on Advanced Computing & Communication Technologies, Rohtak, Haryana, 2012, pp. 96-100.

Educationaldatamining, (2016) Educational Data Mining

http://www.educationaldatamining.org/

Mashael a. Al-barrak & mona s. Al-razgan, (2015) “Predicting student’s performance through Classification: a case study”

Mukesh Kumar, Prof (Dr.) A. J. Singh, (2016) “Predicting Students’ Performance Using Classification Techniques in Data Mining”

Oracle, (2016) Classification

https://docs.oracle.com/database/121/DMCON/classify.htm#DMCON004

Wikipedia, (2016) Data Mining

https://en.wikipedia.org/wiki/Data\_mining