Analysis Of Results of District Rural Schools Using Mining Techniques of in the Improvement of their results

[1]Prof Arif Hakeem, [2]Dr. Tryambak Hiwarkar
[1]Research Scholar in SSSUTMS, Sehore
[2]Professor, Department of Computer Science and Engineering, SSSUTMS, Sehore

Abstract

The main objective of primary education system is to provide quality education to its students. One way to achieve quality in education system is by discovering knowledge for students in a particular course, conversion of traditional classroom teaching system, detection of means used in examination, detection of abnormal values in the result sheets of the students, prediction about students' performance and so on. The knowledge is hidden among their result data set and it is extractable through data mining techniques. Present paper is designed to justify the capabilities of data mining techniques in context of rural education by offering a data mining model for rural education system in the Schools. In this research, the classification task is used to evaluate student's performance and as there are many approaches that are used for data classification, the decision tree method is used here.

By this task we extract knowledge that describes students' performance in examination. It helps earlier in identifying the dropouts and students who need special attention and allow the teacher to provide appropriate advising/counseling.

Keywords-Educational Data Mining (EDM); Classification; Knowledge Discovery in Database (KDD); ID3 Algorithm.

I. INTRODUCTION

The arrival of statistics generation in diverse fields has lead the big volumes of records garage in various formats like data, documents, documents, pix, sound, videos, clinical records and many new information codecs. The information accumulated from special programs require proper method of extracting expertise from huge repositories for better choice making. understanding discovery in databases (KDD), often known as facts mining, ambitions at the discovery of useful data from huge collections of statistics [1]. The predominant functions of information mining are applying diverse strategies and algorithms in an effort to find out and extract patterns of stored information [2]. records mining and information discovery programs have were given a wealthy attention because of its significance in decision making and it has turn out to be an vital issue in diverse organizations. statistics mining strategies have been added into new fields of records, Databases, system learning, pattern Reorganization, synthetic Intelligence and Computation abilities and so on.

There are increasing research interests in the usage of data mining in schooling. This new rising discipline, known as instructional records Mining, concerns with developing techniques that discover information from data originating from academic environments [3]. Instructional records Mining uses many techniques consisting of selection trees, Neural Networks, Naïve Bayes, ok- Nearest neighbor, and many others.

The use of those strategies many sorts of information can be determined inclusive of association guidelines, classifications and clustering. The found information can be used for prediction concerning enrolment of college students in a selected course, alienation of conventional lecture room teaching version, detection of unfair method used in on line exam, detection of unusual values in the result sheets of the scholars, prediction about college students‟ overall performance and so forth.

The main goal of this paper is to apply facts mining methodologies to study student's overall performance in the rural schools. Statistics mining offers many responsibilities that could be used to look at the pupil overall performance. in this studies, the category challenge is used to evaluate pupil's performance and as there are numerous tactics that are used for records class, the choice tree technique is used here. Statistics like Attendance, elegance test, and homework information had been accrued from the scholars, to are expecting the performance on the end of the session. This paper investigates the accuracy of decision tree techniques for predicting scholar performance.

II. DATA MINING DEFINITION AND TECHNIQUES

Data mining, also popularly known as Knowledge Discovery in Database, refers to extracting or "mining" knowledge from large amounts of data. Data mining techniques are used to operate on large volumes of data to discover hidden patterns and relationships helpful in decision making. While data mining and knowledge discovery in database are frequently treated as synonyms, data mining is actually part of the knowledge discovery process. The sequences of steps identified in extracting knowledge from data are shown in Figure 1.

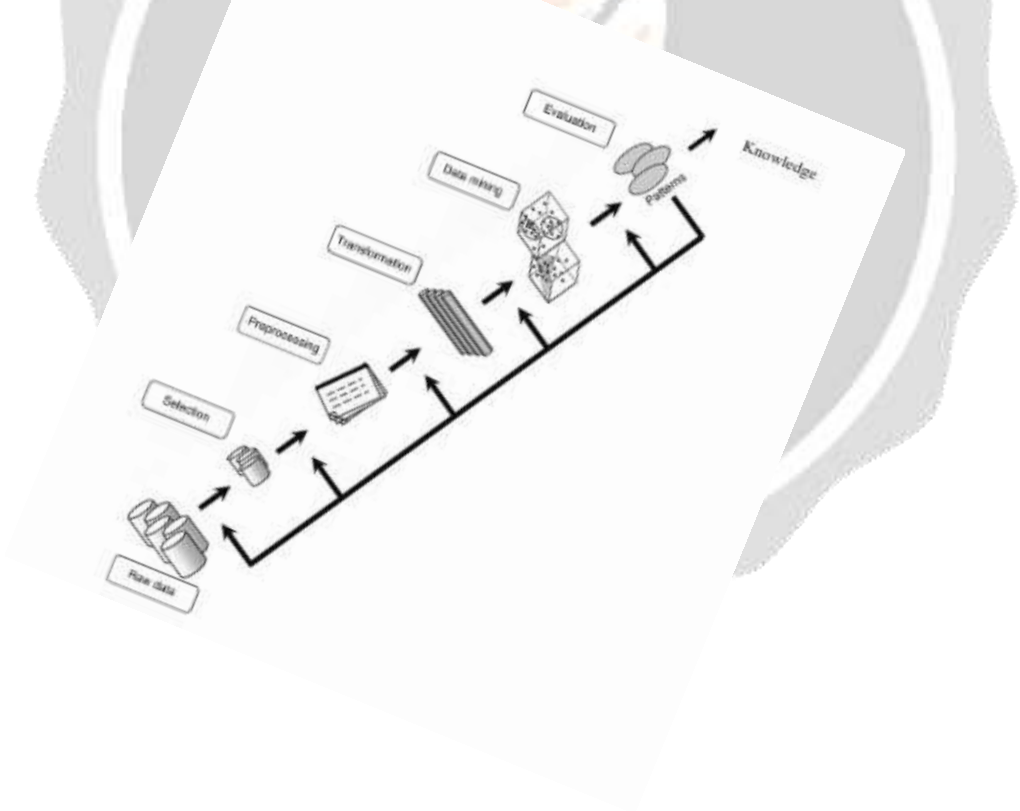

Figure 1: The steps of extracting knowledge from data

Various algorithms and strategies like category, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rule, Decision trees, Genetic set of rules, Nearest Neighbor approach and so on., are

used for understanding discovery from databases. These techniques and strategies in datamining needs to have better understanding.

A. Classification

Classification is the maximum generally implemented data mining approach, which employs a hard and fast of pre-classified examples to develop a model that may classify the population of facts at massive. This approach frequently employs decision tree or neural network-based classification algorithms. The data classification process involves learning and classification. In Learning the training data are analyzed by classification algorithm. In classification test data are used to estimate the accuracy of the classification rules. If the accuracy is acceptable the rules can be applied to the new data tuples. The classifier-training algorithm uses these pre-classified examples to determine the set of parameters required for proper discrimination. The algorithm then encodes those parameters right into a model known as a classifier.

B. Clustering

Clustering may be stated as identification of similar training of objects. by using clustering strategies we can similarly discover dense and sparse regions in item area and might discover normal distribution sample and correlations amongst statistics attributes. category approach can additionally be used for powerful way of distinguishing organizations or instructions of item but it turns into steeply-priced so clustering can be used as preprocessing approach for characteristic subset choice and class.

C. Predication

Regression technique can be adapted for predication. Regression evaluation can be used to version the relationship among one or more independent variables and based variables. In records mining impartial variables are attributes already regarded and response variables are what we need to are expecting. unfortunately, many actual-international problems are now not really prediction. therefore, more complicated techniques (e.g., logistic regression, selection timber, or neural nets) may be essential to forecast future values. The same model kinds can regularly be used for both regression and type. for example, the CART (class and Regression trees) choice tree set of rules may be used to build both class bushes (to categorize categorical reaction variables) and regression timber (to forecast continuous reaction variables). Neural networks too can create both type and regression fashions.

D. Association rule

Association and correlation is usually to find frequent item set findings among large data sets. his type of finding helps businesses to make certain decisions, such as catalogue design, cross marketing and customer shopping behavior analysis. Association Rule algorithms need to be able to generate rules with confidence values less than one. However the number of possible Association Rules for a given dataset is generally very large and a high proportion of the rules are usually of little (if any) value.

E. Neural Networks

Neural network is a set of connected input/output devices and every connection has a weight gift with it. At some stage in the gaining knowledge of section, network learns by using adjusting weights for you to be able to expect the correct magnificence labels of the input tuples. Neural networks have the great capability to derive that means from complex or obscure information and may be used to extract patterns and locate tendencies which

are too complex to be noticed by means of either human beings or other laptop techniques. These are well applicable for non-stop valued inputs and outputs. Neural networks are great at figuring out styles or tendencies in records and nicely appropriate for prediction or forecasting needs.

F. Decision Trees

Decision tree is tree-fashioned structures that constitute units of selections. These decisions generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID).

G. Nearest Neighbor Method

A technique that classifies each record in a dataset based on a combination of the classes of the k record(s) most similar to it in a historical dataset (where k is greater than or equal to 1). Sometimes called the k-nearest neighbor technique

III. RELATED WORK

Data mining in rural education is a recent research field and this area of research is gaining popularity because of its potentials to educational institutes. Data mining maybe utilized in educational area to decorate our knowledge of learning technique to consciousness on identifying, extracting and comparing variables associated with the learning procedure of college students as defined by means of Alaa El-Halees [4]. Mining in instructional environment is referred to as academic information mining.

Han and Kamber [3] describes facts mining software that allow the customers to research records from extraordinary dimensions, categorize it and summarize the relationships that are diagnosed at some point of the mining manner.

Pandey and Pal [5] carried out have a look at on the pupil overall performance based with the aid of choosing six hundred college students from special faculties of Dr. R. M. L. Awadh college, Faizabad, India. by Bayes class on class, language and heritage qualification, it was observed that whether or not new comer college students will performer or not.

Hijazi and Naqvi [6] carried out as take a look at at the pupil overall performance by using deciding on a sample of 300 college students (225 males, 75 females) from a group of schools affiliated to Punjab college of Pakistan. The hypothesis that was stated as "student's attitude toward attendance in magnificence, hours spent in look at on every day foundation after college, college students' circle of relatives earnings, students' mom's age and mother's schooling are significantly related with scholar performance" became framed. by way of simple linear regression analysis, it was observed that the factors like mother‟s education and student‟s own family profits were especially correlated with the pupil instructional performance.

Khan [7] carried out a overall performance have a look at on 400 students comprising 200 boys and two hundred girls decided on from the senior secondary school of Aligarh Muslim university, Aligarh, India with a major objective to set up the prognostic fee of various measures of cognition, personality and demographic variables for fulfillment at higher secondary level in technology flow. the choice turned into based on cluster sampling approach in which the whole populace of interest was divided into agencies, or clusters, and a random pattern of those clusters became selected for similarly analyses. It was found that girls with excessive socio-economic reputation had notably higher academic success in technology movement and boys with low socio- economic fame had pretty better academic achievement in standard.

Galit [8] gave a case take a look at that use college students data to research their learning conduct to expect the outcomes and to warn students at threat before their final exams.

Al-Radaideh, et al [9] implemented a selection tree model to are expecting the final grade of college students who studied the C++ route in Yarmouk university, Jordan in the year 2005. 3 special type strategies particularly ID3, C4.five, and the NaïveBayes have been used. The final results in their results indicated that decision Tree version had higher prediction than other models.

Ayesha, Mustafa, Sattar and Khan [11] describe the usage of okay-approach clustering set of rules to predict pupil‟s learning activities. The information generated after the implementation of facts mining approach can be useful for teacher as well as for college kids.

Bray [12], in his observe on non-public tutoring and its implications, observed that the share of students receiving non-public tutoring in India become distinctly better than in Malaysia, Singapore, Japan, China and Sri Lanka. It turned into additionally found that there was an enhancement of academic performance with the intensity of personal tutoring and this transformation of intensity of private tutoring relies upon on the collective aspect namely socio- economic situations.

Bhardwaj and pal [13] conducted have a look at at the scholar overall performance primarily based by deciding on 300 hundred students from five distinct institutions. by Bayesian category method on 17 attribute, it was observed that the factors like students‟ grade in senior secondary exam, living vicinity, medium of teaching, mom's qualification, students other dependency, own family annual earnings and scholar's family reputation were exceptionally correlated with the student academic performance.

IV. DATA MINING PROCESS

In present day's educational system, a student's performance is determined by the internal assessment and examination. The internal assessment is carried out by the teacher based upon students' performance in educational activities such as class test, assignments, general proficiency, attendance and lab work. The final examination is one that is scored by the student in examination. Each student has to get minimum marks to pass final examination.

A. Data Preparations

The information set used on this have a look at became received from various rural schools of sehore district, near Bhopal (Madhya Pradesh) on the sampling method of result of primary and secondary schools from session 2007 to 2010. To begin with the data is 50 schools. in this step information saved in exceptional tables was joined in a single table after joining process errors were eliminated.

B. Data selection and transformation

In this step only those fields were selected which were required for data mining. A few derived variables were selected. While some of the information for the variables was extracted from the database. All the predictor and response variables which were derived from the database are given in Table I for reference.

TABLE I. STUDENT RELATED VARIABLES

Variable	Description	Possible Values
PEM	Previous Exam Marks	{First > 60% Second >45 & <60% Third >36 & <45% Fail < 36%}
CTG	Class Test Grade	{Poor , Average, Good}

CP	Class Performance	{Poor , Average, Good}
HW	Home Work	{Yes, No}
GP	General Proficiency	{Yes, No}
ATT	Attendance	{Poor , Average, Good}
LW	Lab Work	{Yes, No}
ESM	End Semester Marks	{First > 60% , Second >45 & <60% ,Third >36 & <45%, Fail < 36%}

The domain values for some of the variables were defined for the present investigation as follows:

- **PEM-** Previous Examination Marks/Grade obtained in primary and secondary course. It is split into five class values: *First – >60%, Second – >45% and <60%, Third – >36% and < 45%, Fail < 40%*.

- **CTG –** Class Test Grade obtained. Here in each two class tests are conducted and average of two class test are used to calculate marks. CTG is split into three classes: *Poor – < 40%, Average – > 40% and <60%, Good – >60%.*

- **CP – Class** Performance obtained. In each class are organized to check the performance of students. Student performance is evaluated into three classes: *Poor – Presentation and communication skill is low, Average – Either presentation is fine or Communication skill is fine, Good – Both presentation and Communication skill is fine.*

- **HWP – Home Work** Performance. In each class assignments are given to students by each teacher. Home Work performance is divided into two classes: *Yes– student submitted assignment, No – Student not submitted assignment.*

- **GP -** General Proficiency performance. Like , in each class general proficiency tests are organized. General Proficiency test is divided into two classes: *Yes – student participated in general proficiency, No – Student not participated in general proficiency.*

- **ATT –** Attendance of Student. Minimum 70% attendance is compulsory to participate in End of Examination. But even through in special cases low attendance students also participate in Final Examination on genuine reason. Attendance is divided into three classes: *Poor - <60%, Average - > 60% and <80%, Good - >80%.*

- **LW –** Lab Work. Lab work is divided into two classes: *Yes – student completed lab work, No – student not completed lab work.*

- **FM -** Final Marks obtained in Examination and it is declared as response variable. It is split into five class values: First – *>60% , Second – >45% and <60%, Third – >36% and < 45%, Fail < 40%*.

C. Decision Tree

A decision tree is a tree wherein each department node represents a desire among a number of options, and every leaf node represents a choice.

Decision tree are usually used for gaining statistics for the pe of choice making. decision tree begins with a root node on which it's miles for users to take moves. From this node, customers cut up every node recursively in step with decision tree getting to know algorithm.

The very last result is selection tree wherein each branch represents a probable scenario of selection and its final results. The three widely used decision tree learning algorithms are: ID3, ASSISTANT and C4.5.

D. The ID3 Decision Tree

ID3 is a simple decision tree learning algorithm developed by Ross Quinlan [14]. The basic concept of ID3 set of rules is to construct the decision tree with the aid of using a top-down, grasping search via the given units to check every characteristic at every tree node. if you want to pick out the characteristic this is most useful for classifying a given units, we introduce a metric - records benefit.

To locate an ideal way to categorise a getting to know set, what we need to do is to reduce the questions asked (i.e. minimizing the intensity of the tree). therefore, we want a few characteristic that may degree which questions offer the most balanced splitting. The records gain metric is the sort of function.

E. Measuring Impurity

Given a data table that contains attributes and class of the attributes, we can measure homogeneity (or heterogeneity) of the table based on the classes. We say a table is pure or homogenous if it contains only a single class. If a data table contains several classes, then we say that the table is impure or heterogeneous. There are several indices to measure degree of impurity quantitatively. Most well known indices to measure degree of impurity are entropy, gini index, and classification error.

Entropy = $\sum_j - p_j \log_2 p_j$

Entropy of a pure table (consist of single class) is zero because the probability is 1 and log (1) = 0. Entropy reaches maximum value when all classes in the table have equal probability.

Gini Index = $1 - \sum_j p_j^2$

Gini index of a pure table consist of single class is zero because the probability is 1 and $1-1^2 = 0$. Similar to Entropy, Gini index also reaches maximum value when all classes in the table have equal probability.

Classification Error $= 1 - \max [p_j]$

Similar to Entropy and Gini Index, Classification error index of a pure table (consist of single class) is zero because the probability is 1 and 1-max (1) = 0. The value of classification error index is always between 0 and 1. In fact the maximum Gini index for a given number of classes is always equal to the maximum of classification error index because for a number of classes n, we set probability is equal to p=1/n

and maximum Gini index happens at $1 - n(1/n^2) = 1-1/n^2$ while maximum classification error index also happens at 1-max{1/n}= 1-1/n

F. Splitting Criteria

To determine the best attribute for a particular node in the tree we use the measure called Information Gain. The information gain, Gain (S, A) of an attribute A, relative to a collection of examples S, is defined as

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

Where *Values (A)* is the set of all possible values for attribute *A*, and *S* is the subset of *S* for which attribute *A* has value *v* (i.e., $S_v = \{s \in S \mid A(s) = v\}$). The first term in the equation for *Gain* is just the entropy of the original

$\frac{|S_v|}{|S|}$

collection *S* and the second term is the expected value of the entropy after S is partitioned using attribute *A*. The expected entropy described by this second term is simply the sum of the entropies of each subset , weighted by

$$Split\ Information\ (S, A) = -\sum_{i=1}^{n} \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|}$$

the fraction of examples that belong to *Gain (S, A)* is therefore the expected reduction in entropy caused by knowing the value of attribute *A*

And

$$Gain\ Ratio(S, A) = \frac{Gain(S, A)}{Split\ Information(S, A)}$$

The process of selecting a new attribute and partitioning the training examples is now repeated for each non terminal descendant node. Attributes that have been incorporated higher in the tree are excluded, so that any given attribute can appear at most once along any path through the tree. This process continues for each new leaf node until either of two conditions is met:

1. Every attribute has already been included along this path through the tree, or
2. The training examples associated with this leaf node all have the same target attribute value (i.e., their entropy is zero).

G. The ID3Algoritm

ID3 (Examples, Target_Attribute, Attributes)

- Create a root node for the tree
- If all examples are positive, Return the single-node tree Root, with label = +.
- If all examples are negative, Return the single-node tree Root, with label = -.
- If number of predicting attributes is empty, then Return the single node tree Root, with label = most common value of the target attribute in the examples.
- Otherwise Begin
  - A = The Attribute that best classifies examples.
  - Decision Tree attribute for Root = A.
    - For each possible value, *v* of A,
      - Add a new tree branch below Root, corresponding to the test A = *v i*
      - Let Examples(*v*) be the subset of examples that have the value *vi* for A*i*
      - If Examples(*v*) is empty
      - Then below this new *I* branch add a leaf node with label = most common

target value in the examples

- Else below this new branch add the subtree ID3 (Examples(*v*), Target_Attribute, Attributes – {A})
- End
- Return Root

V. RESULTS AND DISCUSSION

The data set of 30 students used in this study was obtained from Govt. HSS Excellence, Sehore (Madhya Pradesh) for Primary to Secondary students of session 2007 to 2014.

TABLE II DATA SET

S.No.	PEM	CTG	CP	HW	GP	ATT	LW	ESM
	Fir	Goo	Goo	Ye	Ye	Goo	Ye	Fir
	Fir	Goo	Ave	Ye	N	Goo	Ye	Fir
	Fir	Goo	Ave	No	N	Ave	No	Fir
	Fir	Ave	Goo	No	N	Goo	Ye	Fir
	Fir	Ave	Ave	No	Ye	Goo	Ye	Fir
	Fir	Poor	Ave	No	N	Ave	Ye	Fir
	Fir	Poor	Ave	No	N	Poor	Ye	Sec
	Fir	Ave	Poor	Ye	Ye	Ave	No	Fir
	Fir	Poor	Poor	No	N	Poor	No	Thi
	Fir	Ave	Ave	Ye	Ye	Goo	No	Fir
	Sec	Goo	Goo	Ye	Ye	Goo	Ye	Fir
	Sec	Goo	Ave	Ye	Ye	Goo	Ye	Fir
	Sec	Goo	Ave	Ye	N	Goo	No	Fir
	Sec	Ave	Goo	Ye	Ye	Goo	No	Fir
	Sec	Goo	Ave	Ye	Ye	Ave	Ye	Fir
	Sec	Goo	Ave	Ye	Ye	Poor	Ye	Sec
	Sec	Ave	Ave	Ye	Ye	Goo	Ye	Sec
	Sec	Ave	Ave	Ye	Ye	Poor	Ye	Sec
	Sec	Poor	Ave	No	Ye	Goo	Ye	Sec
	Sec	Poor	Poor	Ye	Ye	Poor	Ye	Thi
	Sec	Poor	Poor	No	N	Poor	Ye	Fai
	Thi	Goo	Goo	Ye	Ye	Goo	Ye	Fir
	Thi	Ave	Goo	Ye	Ye	Goo	Ye	Sec
	Thi	Goo	Ave	Ye	Ye	Goo	Ye	Sec
	Thi	Goo	Goo	Ye	Ye	Ave	Ye	Sec
	Fai	Ave	Goo	Ye	Ye	Ave	Ye	Thi
	Fai	Poor	Poor	Ye	Ye	Ave	No	Fai
	Fai	Goo	Poor	No	Ye	Poor	Ye	Fai
	Fai	Poor	Poor	No	N	Poor	Ye	Fai
	Fai	Poor	Goo	No	N	Poor	No	Fai

To work out the information gain for A relative to S, we first need to calculate the entropy of S. Here S is a set of 30 examples are 14 "*first*", 08 "*second*", 03 "*third*" and 05 "*fail*"..

Entropy (S)

$= -Pfirst \log_2(Pfirst) - Psecond \log_2(P\text{second})$

$-Pthird \log_2(Pthird) - Pfail \log_2(Pfail)$

$= -\left(\frac{14}{30}\right)\log_2\left(\frac{14}{30}\right) - \left(\frac{8}{30}\right)\log_2\left(\frac{8}{30}\right) - \left(\frac{3}{30}\right)\log_2\left(\frac{3}{30}\right) - \left(\frac{5}{30}\right)\log_2\left(\frac{5}{30}\right)$

= -(0.466)(-1.099)-(0.266)(-1.099)-(0.1)(-3.321)-(0.166)(-2.584)

= 0.512+0.292+0.332+0.428

= **1.564**

To determine the best attribute for a particular node in the tree we use the measure called Information Gain. The information gain, Gain (S, A) of an attribute A, relative to a collection of examples S,

$Gain(S, PEM) \square Entropy(S) - (|S_{First}|/|S|)\ Entropy(S_{First}) - (|S_{Second}|/|S|)\ Entropy(S_{Second}) - (|S_{third}|/|S|)\ Entropy(S_{third}) - (|S_{fail}|/|S|)\ Entropy(S_{fail})$

TABLE III. GAIN VALUES

Gain	Value
Gain(S, PEM)	0.187855
Gain(S, CTG)	0.167716
Gain(S, CP)	0.119113
Gain(S, HW)	0.071175
Gain(S,GP)	0.014303
Gain(S,ATT)	0.147131
Gain(S,LW)	0.147642
Gain(S,ESM)	0.1476425

PEM has the highest gain, therefore it is used as the root node as shown in figure 2.

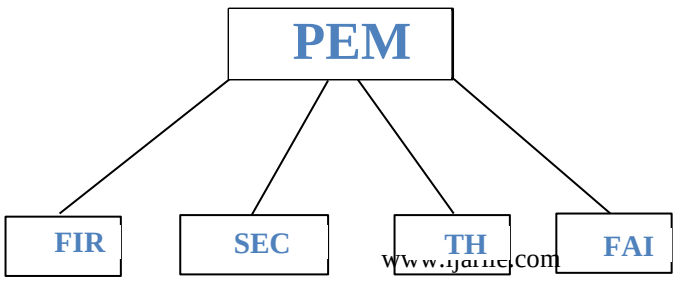

Figure 2. PEM as root node

Gain Ratio can be used for attribute selection, shown in table IV

TABLE IV. Gain Ratio

Gain Ratio	Value
Gain Ratio (S, PEM)	0.416158
Gain Ratio (S, CTG)	0.355674
Gain Ratio (S, CP)	0.22977
Gain Ratio (S, HW)	0.125289
Gain Ratio (S, GP)	0.022887
Gain Ratio (S, ATT)	0.298968
Gain Ratio (S, LW)	0.30032

This process goes on until all data classified perfectly or run out of attributes. The knowledge represented by decision tree can be extracted and represented in the form of IF-THEN rules.

RULE SET TO GENERATED DECISION TREE

I. IF PEM = „First‟ AND ATT = „Good‟ AND CTG = „Go„Average‟ THEN ESM = First
II. IF PEM = „First‟ AND CTG = „Good‟ AND ATT = OR „Average‟ THEN ESM = „First‟
III. IF PEM = „Second‟ AND ATT = „Good‟ AND ASS = THEN ESM = „First‟
IV. IF PEM = „Second‟ AND CTG = „Average‟ AND LW THEN ESM = „Second‟
V. IF PEM = „Third‟ AND CTG = „Good‟ OR „Average‟ATT = “Good‟ OR „Average‟ THEN PEM = „Second‟
VI. IF PEM = „Third‟ AND ASS = „No‟ AND ATT = „AvTHEN PEM = „Third‟
VII. IF PEM = „Fail‟ AND CTG = „Poor‟ AND ATT = THEN PEM = „Fail‟

One type regulations may be generated for every path from each terminal node to root node. Pruning approach changed into achieved by way of eliminating nodes with less than desired quantity of objects. IF- THEN guidelines can be simpler to understand from above mentioned points.

CONCLUSION

In this paper, the category venture is used on pupil database to predict the school students division on the foundation of preceding database. As there are numerous techniques which are used for records classification, the decision tree approach is used right here. Data's like Attendance, Class Test, Seminar and undertaking marks have been gathered from the student's preceding database, to expect the overall performance at the cease of the semester.

This observes will assist the students and the teachers to improve the division of the student. This study will even help to pick out those school students which wanted unique attention to reduce fail ratio and taking appropriate motion for the subsequent final exam.

REFERENCES

[1] Heikki, Mannila, Data mining: machine learning, statistics, and databases, IEEE, 1996.

[2] U. Fayadd, Piatesky, G. Shapiro, and P. Smyth, From data mining to knowledge discovery in databases, AAAI Press / The MIT Press, Massachusetts Institute Of Technology. ISBN 0–262 56097–6, 1996.

[3] J. Han and M. Kamber, "Data Mining: Concepts and Techniques," Morgan Kaufmann, 2000.

[4] Alaa el-Halees, "Mining students data to analyze e-Learning behavior: A Case Study", 2009..

[5] U . K. Pandey, and S. Pal, "Data Mining: A prediction of performer or underperformer using classification", (IJCSIT) International Journal of Computer Science and Information Technology, Vol. 2(2), pp.686-690, ISSN:0975-9646, 2011.

[6] S. T. Hijazi, and R. S. M. M. Naqvi, "Factors affecting student's performance: A Case of Private Colleges", Bangladesh e-Journal of
Sociology, Vol. 3, No. 1, 2006.

[7] Z. N. Khan, "Scholastic achievement of higher secondary students in science stream", Journal of Social Sciences, Vol. 1, No. 2005..

[8] Galit.et.al, "Examining online learning processes based on log files analysis: a case study". Research, Reflection and Innovations in Integrating ICT in Education 2007.

[9] Q. A. AI-Radaideh, E. W. AI-Shawakfa, and M. I. AI-Najjar, "Mining student data using decision trees", International Arab Conference on Information Technology(ACIT'2006), Yarmouk University, Jordan, 2006.

[10] U. K. Pandey, and S. Pal, "A Data mining view on class room teaching language", (IJCSI) International Journal of Computer Science Issue, Vol. 8, Issue 2, pp. 277-282, ISSN:1694-0814, 2011.

[11] Shaeela Ayesha, Tasleem Mustafa, Ahsan Raza Sattar, M. Inayat Khan, "Data mining model for rural education system", Europen Journal of Scientific Research, Vol.43, No.1, pp.24-29, 2010.

[12] M. Bray, The shadow education system: private tutoring and its implications for planners, (2nd ed.), UNESCO, PARIS, France, 2007.

[13] B.K. Bharadwaj and S. Pal. "Data Mining: A prediction for performance improvement using classification", International Journal of Computer Science and Information Security (IJCSIS), Vol. 9, No. 4, pp. 136-140, 2011.

[14] J. R. Quinlan, "Introduction of decision tree: Machine learn", 1: pp. 86-106, 1986.

[15] Vashishta, S. (2011). Efficient Retrieval of Text for Biomedical Domain using Data Mining Algorithm. *IJACSA - International Journal of Advanced Computer Science and Applications, 2*(4), 77-80.

[16] Kumar, V. (2011). An Empirical Study of the Applications of Data Mining Techniques in Higher Education. *IJACSA - International Journal of Advanced Computer Science and Applications, 2*(3), 80-84. Retrieved from http://ijacsa.thesai.org.

AUTHORS PROFILE

Arif Hakeem is Assistant Professor in the Department of Computer Science & Engineering, NRI College Bhopal He obtained his BE degree from L.N.C.T. University Bhopal (1998) and M.Tech. in Computer Science & Engineering from Bangalore University, Karnataka. He is currently doing research in Data Mining in the Department of Computer Science & Engineering, SOE , SSSUTMS under the guidance of DR Trimbak Hiwarkar. He has published one international paper.