ISSN 1991- 8690 الترقيم الدولي 8690 - 1991

Data Mining Approach for Predicting Learner's Achievement

Naofal Mohamad Hassin Azeez

Computer Department-Computer Science and Mathematical College

E.mail: nawofel_aziz@yahoo.com

Abstract:

Student achievement variables that may be included into student database can be classified into three main categories, student variables. Instructor variables and general variables. This paper presents a new machine-learning model for extracting knowledge From student attributes in a given database. This knowledge can be used for determining the relative importance and effectiveness of student's attributes for the prediction of their college academic achievement, and the relationship between these attributes and their achievement. The model includes three main algorithms namely: preprocessing of database, attribute selection and rule extraction algorithm. Preprocessing of database aims to alleviate the dimensionality of the given database. It is performed according to (i) Detecting memo attributes and abstracting their field values into minimum abstraction level, (ii) Detecting the attributes, which have repeated values (including sparse values), and dropping them from database and (iii) Using fuzzification for transferring the attributes of continuous values into linguistic terms. This transformation leads to reducing the search space. Attribute selection algorithm selects the most relevant attributes set by the calculations of an evaluation function. The resulted set of attributes is passed to rule extraction algorithm for extracting an accurate and comprehensible set of rules.

Keywords: Student Achievement Variables, Attribute Selection, Dimensionality Reduction, Rule Extraction, Knowledge Acquisition.

طريقة تنقيب البيانات للتنبؤ بمستوى تحصيل المتعلم

نوفل محمد حسين عزيز

قسم الحاسبات – كلية علوم الحاسوب والرياضيات – جامعة ذي قار

الخلاصة:

إن متغيرات تحصيل الطالب التي يتم تضمينها في قاعدة بيانات الطالب يمكن تصنيفها لثلاث فئات رئيسية هي: المتغيرات الخاصة بالطالب، المتغيرات الخاصة بالمعلم، وأخيرًا المتغيرات العامة. يقدم هذا البحث نموذجًا جديدًا لتعلم الالة وذلك من أجل استخلاص المعرفة من خلال سمات الطالب في إحدى قواعد البيانات المعطاة. ويمكن استخدام هذه المعرفة في تحديد الأهمية النسبية لسمات الطلاب وفعاليتها في التنبؤ بتحصيلهم الأكاديمي في كليتهم. ويتضمن هذا النموذج ثلاث خوارزميات رئيسة وهي: ما قبل معالجة البيانات لقاعدة البيانات، اختيار أفضل السمات، وأخيرًا خوارزمية استخلاص القواعد. وتهدف المعالجة القبلية لقاعدة البيانات إلى اختزال الأبعاد غير الملائمة في قاعدة البيانات المعطاة. ويتم أداء هذه العملية وفقًا إلى (1) تحديد

العناصر كثيرة القيم وتلخيصها إلى أدنى مستوى (2) تحديد العناصر ذات القيم المتكررة (بما في ذلك القيم المتناثرة) وإسقاطها من قاعدة البيانات (3) تحويل العناصر ذات القيم المتصلة إلى قيم منفصلة في صورة حدود لغوية. ويؤدي هذا التحويل إلى تقليل أبعاد البحث. أما خوارزمية اختيار العناصر فيتم من خلال اختيار أكثر العناصر ملائمة وذلك عن طريق حساب دالة التقييم. وأخيرًا تمر المجموعة الناتجة من العناصر إلى خوارزمية استخلاص القواعد وذلك من أجل استخراج مجموعة من القواعد الدقيقة والمفهومة.

1- Introduction:

Student achievement encompasses student ability and performance. It is multidimensional intricately related to human growth and cognitive, emotional, social, and physical development. It does not relate to a single instance, but occurs across time and levels, through a student's life, Merriam Webster defined achievement as "the quality and quantity of a student's work'[1].

There are several attempts to find relationship among academic achievement and other individual variables such as, mental capacity, emotional intelligent, gender, and other variables [2]. Self evaluations of one's abilities not only statistically predicted one's preferred way of doing things, including learning approaches, modes of thinking, and thinking styles, but also predicted the degree of intensity of one's interest in different types of careers, one's cognitive – developmental levels, as well as one's personality traits [3]. The relative importance and effectiveness of student's attributes on the predication of his / her college academic achievement, and the relationship between these attributes and their achievement are active fields of research [8][14].

Student databases which include the effective attributes are rich sources of knowledge, however they need further analysis to be transformed it into useful knowledge. Unfortunately these databases are so large that human analysis to extract useful knowledge is difficult even with the capabilities of modern database systems [15] . The need of tools to tap this knowledge and bring it to a specific level of abstraction where it can be used in decision-making and expert system design is necessary. This has led to the emergence of data mining[16]. The goal of data mining is to discover knowledge that has not only high predictive accuracy but also is comprehensible to user. This goal can be achieved by using high level knowledge representation [17]. Popular machine learning techniques for extracting If THEN rules from databases are ID3 [18]. , [19] and is successor, C4.5 [20], multi-layer feed-forward artificial neural networks (ANNs) [21][22] and evolution-based genetic algorithms (Gas) [23][24]. However, they suffer from the following problems; (i) ID3 and C4.5 lack backtracking and may ignore some attributes, so they make less accurate prediction. (ii) ANNs suffer from the curse of dimensionality because the number of input nodes increased exponentially with the input values of attributes. (iii) Gas suffer from difficulty of input data representation. Other rule induction algorithms were based on complex statistical evaluation function, which make the heuristic search computationally burdensome specially in case of large noisy databases. Also some of these algorithms can't deal with databases of continuous attributes.

This paper describes the student attributes which may affect his/here academic achievement. It also present algorithms for selecting the most relevant attributes, and extracting useful knowledge from them. The paper is organized as follows. First it briefly describes the student variables (attributes). Next an automated knowledge extraction model for extracting rules from databases is presented. Finally the model is applied on student database and the paper is concluded. The proposed framework for predicting learner's achievement from a data mining technique is shown in figure (1).

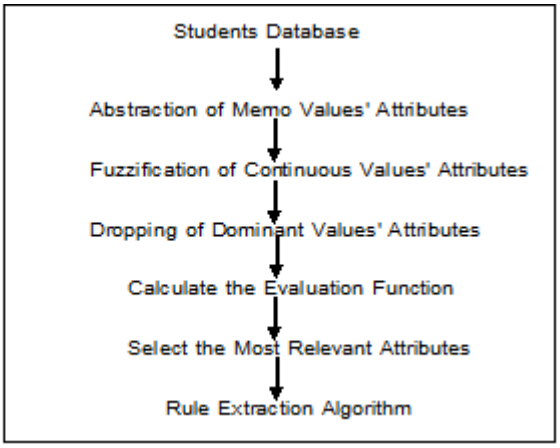

Figure 1: The proposed framework for predicting learner's achievement

2. Student Achievement Variables

The academic achievement studies according to achievement can be classified into three main categories: student's variables, instructor's variables and other general variables. The following section presents some of these studies and the main related works.

2.1. Student's Variables

A study to evaluate the process of selecting for engineering and other sciences courses at a tertiary institution through some predictor variables was presented [2]. The variables were the grade results for science, English and mathematics, the general scholastic aptitude test senior, the senior attitude test and sixteen personality factor questionnaire.

Data Mining Approach fore predicting Learner's Achievement. The relationship between self-identity and academic persistence and achievement in a counter stereotypical domain was a study conducted by cover [3]. It revealed that the higher self-concept and self-schema, the more positive self-description and the better the academic achievement. The study also showed that self-identity improves through social interaction communication with others, which would enhance achievement.

Prediction of student achievement based on development study and investigating different students' types of censoring variables influencing student achievement was appropriate for realizing student achievement [4]. The variables were the number of credit hours and grade point average (GPA).

A report that illustrates the use of hierarchical linear models (HLM) with national assessment of educational progress (NAEP) data to identify student achievement was introduced [5]. The report focused on the methodology. The multilevel nature of data and student-level outcomes such as gender and race/ethnicity were used to predict function for school-level factors.

Several types of HLM analysis were conducted on 1990 NAEP data to predict achievement outcomes in mathematics and geometry, predicting average achievement between school, the gender gap, and the race/ethnicity gap.

The validity of Miller analogies test (MAT) for predicting achievement of graduate students in education was investigated [6]. The extent of age and gender differences in prediction of graduate student performance from MAT scores were examined. Results from this study indicated that MAT scores were significantly correlated with students' cumulative grade point averages and with their grade performance in five specific graduate courses.

Differences in predicting achievement by sex on the lowa tests of basic skills (ITBS) from the verbal, quantitative, and nonverbal scores on the cognitive abilities test was examined [7].

2.2 Instructor's Variables

Many studies examine the ways in which instructor variables such as qualifications and other inputs are related to student achievement. Instructor effects appear to be additive and cumulative. Instructors' competence variables which have been examined for their relationship to student achievement include: general academic ability and intelligence, verbal ability , years of teaching experience, measures of subject matter and teaching knowledge, certification status, and teaching behaviors. Medley et.

Concluded that there is little or no relationship between instructors' measured intelligence and their student's achievement [8].

Murnane concluded that instructors' verbal ability is related to student achievement and this relationship may be differentially strong for instructors of different types of students [9].

Kreitzer found small, statistically insignificant relationship between instructors' subject matter knowledge and their student's achievement [10].

Byrne summarized the results of thirty studies relating instructors' subject matter knowledge to student achievement [11]. The instructor knowledge measure were either a subject knowledge test or number of college courses taken within the subject area. This was supported by Monk's in a study of mathematics and science achievement [12]. Brophy linked student learning to variables such as instructor clarity, Enthusiasm, task-oriented behavior, variability of lesson approaches, and student opportunity to learn criterion material [13]. Instructors' abilities to structure material, ask higher order question, use student ideas, and probe student comments have also been found to be important variables in what students learn.

2.3. General Variables

There are additive variables, such as, class sizes and pupil loads, planning time, opportunities to plan and problem solve with colleagues, curricular supports including appropriate materials and equipment, and library media programs [14].

As mentioned in the previous studies the student's variables are the most effective on the student's achievements. The following section concentrates on a novel machine learning model to extract knowledge about student achievements where student variables are given in database.

3. Knowledge Extraction for predicting student Achievement

The use of student's variables (attributes) given in student database to extract knowledge about the student achievement is the purpose of the following sections. The automated knowledge acquisition model developed in the following sections consists of three main parts. Namely; preprocessing of database, attribute selection and a rule extraction.

3.1. Preprocessing of Database

Field values in the real world database may take different shapes, model values, missing values, unique values (s), and constant values fields. The unique value fields have a different value for each record. The fields, which have normal values, contain information and are important in acquiring knowledge. These fields determine each record exactly and have no predictive value. One value fields (undervalued) does not contain information that helps different between different files. They should, therefore, be ignored for purposes of data extraction. As general, if 85-99 percent of the values in fields are identical then these fields will be useless [15]. Some fields may contain missing values (kind of dirty data), so these fields must be ignored. The main problem occurs in the fields that contain multi-terms (complete paragraph) such as the patient complaint in medical database and the favorite subject in student database. These fields contain valuable information but they may include data that has the same meaning with different expressions. Pruning or abstracting this kind of data is necessary in order to reduce their values and consequently transforming them into columns with normal values. The following techniques are used to perform these aims.

3.1.1. Text Attribute Abstraction

The learning algorithms assume that the attributes in database contain single symbolic or numeric value. However, attributes of text (memo) values may exist in real database. These attributes must be abstracted to the minimum level. Information retrieval system, which users vector space representation to the memo data, is one approach to the memo representation [16]. However, it fails with memo of high dimensions. Other statistical measures for weighting terms in the memo attributes were introduced (27). However, two or more terms with completely different meaning may have the same weight, which yields to ambiguity and difficulty in the final interpretation. Pruning the field data from undesired words may be another approach. Unfortunately, these words are unlimited and their drop may lost the sentence meaning. A new approach to information extraction based on capturing the relevant terminology in specific domain is presented here. It depends on building a domain field dictionary (DFD) with the help of domain expert and matching the memo field with the terms of the DFD. The domain field dictionary includes the possible relevant expression sets si;

DFDset = U^{S}_{i} …………………………..…(1)

The matching process between the memo search term, STj, and the DFDset has one of the following conditions:

- $ST_j \neq S_i$ and ST_j ¢ S_k
- $ST_j \neq S_i$ and $ST_j \leq S_k$ ………..(2)
- $ST_j \neq S_i$ and ST_j £ S_k
- $ST_j \neq S_i$ and $ST_j \leq S_k$

Where:

$ST_j = S_i$, indicates that the memo search term is identically exist in DFD set

$ST_j \leq S_k$, indicates that the memo search term is exist as a part of Si in DFDset

```
for j = 1 to J DO
STj = { };
LSTj = { };
FFTj = { };
STj = {first word in field}j

Then
LSTj = STj
Else
If STj ≠ Si and STj ¢ Sk
Then
Drop STj
LSTj = { }
Else
If STj = Si and STj ≤ Sk
Then

Do While not end of memo fieldj
Abstraction_fun;
If STj = Si and STj ¢ Sk

LSTj = STj
Else
If STj ≠ Si and STj ≤ Sk
Then
LSTj = { }
STj = STj + Next word in fieldj
Call Abstraction_fun;
End If
End If
End If
End If
FFTj = FFTj U LSTj
STj = { Next word in fieldj }
```

Loop

J: Total number of records.

LTS_j: Last term in the j^{th} record

FV_j: Field value Set, in the j^{th} record

3.1.2. Fuzzification of Continuous Attribute

If the database contains continuous attributes (real values), a search based on all the values of the thread is possible to extract the rules leads to cumbersome calculations and takes time. This problem can be solved by using fuzzification process, reduction of search space. A membership function describes the fuzzy subset of universe of discourse U (0,1), that represents the degree to which and belongs to the set V. A fuzzy linguistic variable, V, is an attribute whose domain contains exact values, which are labeled for the fuzzy subsets (28). Therefore, continuous attributes can be converted into scientific term such as; short (S), medium (M) and long (L). A non-overlapping rectangular membership function may be used, and the bounds of each scientific term can be determined by using the horizontal histogram of real values (29).

3.1.3 Dominant Values Attribute Reduction

Any attribute in given database may have n values; {v1, v2,…, vk…,vn}. Assuming that the missing values in a given column will take the same value such as "null". These values can be used for determining the attribute type by calculating their probabilities of existence {p1,p2,…,pk…pn}. If pk equals 1.0 then the attribute belongs to one value attributes and If pk in range 0.85 to 1.0 then type of attribute is almost one value. Such attributes must be stripped from the database because they have worthless information.

For j = 1 to j
(J : number of columns)
For K = 1 to n

$$(P_K)_j = \frac{No.of\ idintical\ values\ of\ type\ V\ K}{Totl\ number\ of\ values} \quad \text{.....(3)}$$

Find max {P1, P2, …., Pk, …. Pn } $_j$
IF max $\{P_k\}_j \geq £$
THEN
(Drop feature from the original table)
ELSE Keep the attribute in the database
Next K
Next j
Where;
£ : is the threshold level for attribute dropped.

3.2. Attribute Selection

Preprocessing stage of database constructs new universal attribute set, Uatt_new (Uatt_new ≤ Uatt). This set is now suitable for extracting the most relevant attributes. The most relevant attributes are determined according to an evaluation function.

3.2.1. Evaluation Function

An evaluation function, (IS), represented in (30) was used to determine whether the attribute may be included in the most relevant attributes set or not. This function is given by:-

$$Es = \frac{n*CAT}{\sqrt{n+n\,(n-1)*C\,AA}} \quad \text{.....(4)}$$

Where;
n :Number of attribute in test set.
CAA : Average of attribute – attribute correlation.
CAT : Average of attribute –target correlations.
P (tjlai) : the conditional probability that target has value tj when an attribute has a value ai
This function depends on the engagement among the attribute set (CAA) and the attribute to the corresponding target (CAT), it succeeds in special cases. In the real database, the evaluation function fails if any attribute has only one value (CAT=0.0 & CAA=undefined value). So this paper introduces a unit step function, U, to modify the evaluation function as follows:-
$Es_{\_mod} = U \,.\, E_s$

3.2.2. Search Strategy

The number of Es calculation to extract Uatt_most from Uatt_new which has N attributes is 2N-1. The search space increases with N. this is burdensome. A proposed methodology to extract Uatt_most and reduce the search space is presented. The methodology is based on constructing N groups. The ith group (Gi) includes a number of sets equal to N-i+1. The total number of sets required for the Es calculation will be $\sum_{I=1}^{i=N} N - I + 1$. This reduces the search space linearly. The following algorithm explains the search strategy.

U_{att_new} = {a1, a2, …… aN}
Test_Set + { }
Temp_Set (ij) = ( )
ES = 0.0
For I = 1 to N
Construct G_i
For j = 1 to N-I +1
Temp_Set (i-j) = Test Set U j^{th} attribute of U_{att_new}
Calculate E_s (I, j)
If E_s (I, j) ≥ E_s
ThEN
Max_E_s = E_s = E_s (I, j)
E_s = E_s (I, j)
ELSE
Next j
Test_Set = U_{att_most}
$U_{att\_most} = U_{att\_most}$ – Test_Set
Next i

The final and the most crucial stage is the rule extraction algorithm. The proposed algorithm divides the database into N predictive attributes and one target

attribute (class). Each attribute is denoted by Ai (i = 1,2, …., N). The different possible values, mi, of the attribute Ai is vij (j = 1,2, …., mi). The target attribute has K classes. Each class is denoted by class k (K = 1,2, …., K). The level of search depth is labeled as LeveI L ($1 \leq L \leq N-1$). The Steps of the proposed algorithm are shown in figure (5).

For K1 to K , (number of classes)
For I to N ,(number of attributes)
For h to M_i , (number of values in each attriute)
L = 1 , (Supporting Level)
S = $\{v_{ij}\}$ (Test_set)
Do until the end of all attribute
Do While L < N-1
P = p (S Class $_K$)
IF p =1
THEN
Get rule from values of S and Level L
Drop the values used from attributes
Check conjunction of all remaining
Values in set S }
L = L + 1
End Do
End Do
Next j
Next K
Begin Refinement of the extracted rules
{Neglect rules which have supporting lecel less than
Specific threshold value
Remove redundant rules
Summarize rules
Defuzzify the continuos attribute values
}
End

The proposed algorithm is efficient for extracting accurate and comprehensible set of rules from a given database fir the following reasons; (i) It extracts only rules that have antecedent(s) satisfy 100 % of the conditional probability within certain class, consequently the output rules are accurate. (ii) It controls the number of antecedents(s) in each extracted rule according to stopping criterion. The stopping criterion prevents number of antecedent in test_set to exceed N-1. Therefore the extracted rules are comprehensible for the user. (iii) the proposed algorithm extracts only a set of rules which satisfy a specific threshold supporting level (number of records in a given database covered by the extracted rules). The increase of the supporting reduces the number of output rules and increases their quality.

4. Application and results

A student database table is used for the implementation. It includes 9 predictive attributes and one target attribute. The predictive attribute are student ID, scholar achievement, prior culture, mental capacity, disembedding ability, residence, gender, favorite subject and family income. The target attribute is the grade (final student academic achievement). A sample of this database is shown is table (1). The preprocessing stage performs the following steps; (i) text abstraction algorithm abstracts the attribute "favorite subject" using domain field dictionary, while the other attributes are not changed, (ii) the algorithm of dominant values drops the student ID, residence, family income attributes. (iii) the scholar achievement, the mental capacity and disembedding ability attributes are fuzzified into nominal values by the fuzzification algorithm. These values are good (G), very good (VG) and excellent (EX) for scholar achievement attribute, high, medium and low for the mental capacity attribute and independent, intermediate and dependent for disembedding ability attribute. The attribute selection algorithm determines the most relevant attribute set as: { scholar Achievement, prior culture, Mental capacity, Disembedding ability, Favorite subject}, which have the largest evaluation function value, Es = 0.8031. Table (2) shows a sample of the database, which contains the most relevant attributes and their values.Rule extraction algorithm can be applied on this database for extracting a set of comprehensible rules at different levels and threshold supporting level > 3.5 %

Table 1: A sample of student's database

Student ID	Scholar Achievement	Prior Culture	Mental Capacity	Dis – Embedding Ability	Residence	Gender	Favorite Subject	Favorite Subject	Family Income	Grade
1801	315.5	S-Science	5	12	village	Female	I Like Language and Specially English which is My favorite Subject	English		Good
1803	330	S-Machematics	4	10	village	male	Physics is My Favorite Subject	Physics		Good
———	———	———	———	———	———	———	———	———	———	———
1813	309	S-Science	5	12	village	male	I Like Chemical	Chemical		Good
1816	329.5	S-Machematics	4	12	village		Mathematics is My favorite Subject	Mathematics	2000	VG
———	———	———	———	———	———	———	———	———	———	———
1841	313	Literary	5	12	village	female	The favorite subject is English	English	350	Pass
1843	338.5	S-Machematics	6	15	village	female	I Like Mathematics	Mathematics	600	VG
———	———	———	———	———	———	———	———	———	———	———
1916	319.5	Literary	6	15	City	male	The Favorite Subject is English	English		Pass
1918	293.5	Literary	5	12	village	male	English is My favorite Subject	English	230	Pass

Table 2: The most relevant attributes

Scholar Achievement	Prior Culture	Mental Capacity	Disembedding Ability	Favorite Subject	Grade
VG	S-Science	M	INDEPENDENT	English	Good
VG	S-Mathematics	M	INDEPENDENT	Physics	Good
EX	S-Mathematics	H	INDEPENDENT	Mathematics	VG
———	———	———	———	———	———
VG	S-Science	M	INDEPENDENT	Chemical	Good
Good	Literary	L	INDEPENDENT	English	Pass
G	S-Mathematics	H	INDEPENDENT	Mathematics	VG
EX	S-Mathematics	H	INDEPENDENT	Mathematics	VG
———	———	———	———	———	———
VG	Literary	M	INDEPENDENT	Geography	Good
Good	Literary	L	INDEPENDENT	English	Pass
VG	S-Science	M	INDEPENDENT	Mathematics	Good
VG	S-Mathematics	M	INDEPENDENT	English	Good
Good	S-Science	L	INDEPENDENT	Mathematics	Pass
VG	S-Mathematics	H	INDEPENDENT	Mathematics	VG

Level 1

1- IF Mental Capacity is Medium THEN Grade is Good (Supporting level = 67.27 %)
2- IF Mental Capacity is High THEN Grade is VG (Supporting Level = 22.73 %).
3- IF Mental Caoacity is Low Then Grade is Pass (Supporting Level = 10.0 %)
4- IF Scholar Achievement is Good THEN Grade is Pass (Supporting Level = 10.0 %)

Level 2:

1- IF Prior Culture I S-Mathematics and Favorite Subject is Mathematics THEN Grade is VG (Supporting Level 22. 73 %)>
2- IF Scholar Achievement is VG and Favorite Subject is Physics THEN Grade is Good (Supporting Level = 10.91 %)
3- IF Scholar Achievement is EX and Prior Culture is S-Mathematics THEN Grade THEN Grade is VG (Supporting Level = 3.64 %)

Level 3:

IF Scholar Achievement is VG and Prior culture is Literary and Disembodying ability is Independent THEN Grade is Good (Supporting Level 19>1 %) The extracted rules of level 4 and 5 have a supporting level less than the threshold supporting level. Therefore they don't appear.

5. Conclusion

Early databases were designed for data manipulation and helping decision-maker. Thousands of records have been kept in these databases. However, these databases were not prepared for automated knowledge extraction. These databases need careful treatment to be prepared for machine learning algorithm. The preprocessing algorithm presented in this paper makes fine pruning of the attribute used in the data tables. This pruning is preformed through systematic sequences of operations starts from recognizing memo terms. Generalizing their attributes, fuzzified the continuous values and detecting missing values. The less information attributes are dropped as preliminary stage of attribute reduction. The second algorithm is the attribute selection which depends on the examination of the correlation between attribute-to-attribute-to-target. This examination is necessary in order to extract the most relevant attributes. These attributes form the rich source of knowledge in the given database.

A set of accurate and comprehensible rules is obtained. The machine learning algorithm extracts only the rules which have the antecedent(s) satisfy 100% of the conditional probability within certain class. It controls the number of antecedent(s) in each extracted rule according to stopping criterion. Additionally, the algorithm uses the threshold supporting level to control the quality of extracted rules. The future work is to consider the fuzzy nature of the values in the given attributes and to extract fuzzy rules from the given database. Also to utilize the genetic algorithm for both dimensionality reduction and rule extraction to introduce the intelligent agents in this domain.

References

[1] Fu-Yun Yu and Hsin-Ju Jessy Yu, "Incorporating e-mail into the learning process: its impact on student academic achievement and attitudes", Computers & Education, volum 38, Issues 1-3, January-April 2002, pages 117-126.

[2] R van Eeden, M de Beer & C H Coetzee,"Cognitive ability, learning potential, and personality traits as predictors of academic achievement by engineering and other science and technology students", south African Journal of Higher Education, volum 15(1), 2011.

[3] Coover, G.E., & Murphy.," the communicated self: Exploring the interaction between self and social context", Human communication Rrsearch,26(1), 125-147,2000.

[4] Faxian Yang,"Using survival analysis to analyze and predict students" achievement from their status of developmental study",paper presented at the 40thAnnual AIR forum (May 21-24,2000) in Cincinnati, Ohio . 2000.

[5] Susan stone," correlates of change in student reported parent involvement in schooling: A new

look at the National Education Longitudinal study of 1988", American Journal of Orthopsychiatry, volume 76, issue4, Oct. 2006, pages 518-530

[6] Helen Demetriou, paul Goalen and Jean pudduck," Academic performance, transfer, transition and friendship: listening to the student voice", international Journal of Educational research, volume 33,Issue 4, 2000, pages 425-441.

[7] Peter C. scales, peter L. Benson, Eugene C. Roehlkepartain, Jr., Arturo sesma and Manfred van Dulmen,"The rple of developmental assets in predicting academic achievement: A longitudinal study", Journal of Adolescence, volume 29, issue 5, Oct. 2006,pages 691-708.

[8]Gerard van de watering and Janine van der Rijt, "teachers' and student' perceptons of assessments: A review and a study into the ability and accuaracy of estimating the difficulty levels of assessments items", Educational Research Review, Volume 1, Issue 2, 2006, pages 133-147.

[9] Murnane, R.J.," Do effective teachers have common characteristics: Interpreting the quantitative research evidence", paper presented at the National Research council conference on teacher Quality in science and Mathematics, Washington, D.C, June 1985.

[10] Maria Teresa Tatto, Sylvia schmelkes, Maria Del Ref ugugio Guevara and Medardo Tapia., "Implementing reform amidst resistance: the regulation of teacher education and work in Mexico",. international Journal of Educational Research, Volume 45, issue 5, 2006, pages 267-278.

[12] Monk, D. H. and king, J.A ," Multilevel teacher resource effects in pupil performance in secondary mathematic and science: the case of teacher subject matter preparation",Pp. 29-58 in R.G. Ehrenberg (Ed.), choices and consequences: contemporary policy issues in education. Ithaca, NY:ILR press, 1994.

[13] Marie-Christine Opdenakker and Jan Van Damme.,"Teacher characteristics and teaching styles as effectiveness enhancing factors of classroom practice.". teaching and teacher Education, volume 22, Issue 1, January 2006, pages 1-21.

[14] Lance, Keith Curry, and David V. Loertscher. Powering Achievement : School Library Media programs Make a Difference: The Evidence Mounts. 2nd ed. San Jose, Calif: Hi willow Research and publishing, 2002.

[15] Stelios H. Zanakis and Irma Becerra-Fernandez, " Competitiveness of nations: A knowledge discovery examination:, European Journal of Operational Research, Volume 166, Issue 1, 1 October 2005, pages 185-211.

[16] Duen-Ren Liu and chih-Kun Ke, " knowledge support for problem-solving in a production process: A hybrid of knowledge discovery and case-based reasoning", Expert systems with Applications, Volume 33,Issue 1,July 2007, pages 147-161.

[17] Haining Yao and Letha Etzkorn " Automated conversion between different knowledge representation formats", knowledge-Based systems, volume 19, Issue 6, Oct. 2006, pages 404-412.

[18] Shihchieh Chou and Change-Ling Hsu, "MMDT: a multi-valued and multi-labeled decision tree classifier for data mining" , Expert Systems with Applications, Volume 28, Issue 4,May 2005, pages 799-812.

[19] Omer Akgobek, Yavuz Selim Aydin, Ercan Oztemel and Mehmet sabih Aksoy, " A new algorithm for automatic knowledge acquisition in inductive learning", knowledge-Based systems, Volume 19, Issue 6,Oct. 2006, pages 388-395.

[20] Guoqing Chen, Hongyan Liu, Lan Yu, Qiang Wei and Xing Zhang "A new approach to classification based on association rule mining", Decision support systems, systems, Volume 42, Issue 2, Nov. 2006, pages 674-689

[21] Guobin Ou and Yi Lu Murphey, , "Multi-class pattern classification using neural networks", pattern Recognition, volume 40, Issue 1, Jan. 2007, pages 4-18.

[22] Greer B. Kingston, Holger R. Maier and Martin F. Lambert "A probabilistic method for assisting knowledge extraction from artificial neural networks used for hydrological prediction". Mathematical and computer Modeling. Volume 44, Issues 5-6, sept. 2006, pages 499-512.

[23] S. Dehuri and R. Mall. "predictive and comprehensible rule discovery using a multi-objective genetic algorithm", knowledge-Based Systems, volume 19,Issue 6, Oct. 2006, pages 413-421.

[24] M. kaya and R. Alhajj, " Genetic algorithm based framework for mining fuzzy association rules",

fuzzy sets and systems, Volume 152,Issue 3, 16 June 2005, pages 586-601.

[25] Michael J.A. Berry and Gordan Linoof "Mastering Data Mining the Art and Science of Customer Relationship Management" Wiley computer publishing, John Wiley & Sons Inc. 2000.

[26] Thomas Rolleke , Theodora Tsikrika and Gabriella Kazai, "A general matrix framework for modeling information Retrieval", Information processing & Management, Volume 42, Issue 1, Jan. 2006, pages 4-30.

[27] Working of Learning from Text and the Web, conference on Automated Learning and Discovery CONALD-1998.

[28] Vassilios Petridis, Vassilis G. Kaburlasos, " Clustering and classification in Structured Data Domains Using Fuzzy Lattice Neurocomputing (FLN)", IEEE Transaction on Knowledge and Data Engineering, Mar./Apr., Vol. 13, No. 2, 2001.

[29] J.H. Wang, Wen-Jeng Liu and Lina-Da Lin, " Histogram-Based Fuzzy Filter for Image Restoration", IEEE Trans on Systems, Man, and Cybernetics, Vol.32, No.2, PP. 230-238, Apr.2002.

[30] R.M. Hogarth. "Methods For Aggregating Opinions", In H.Jungermann And G.De Zeeuw, Editors, Decision Making And Change In Human Affairs. D. Reidel publishing, Dordrecht-Holland, 1977.

[31] Li-fang Zhang "predicting cognitive development, intellectual styles, and personality traits from self-rated abilities" Learning and Individual Differences, Volume 15, Pages 67-88, 2005.