ISSN 2047-3338

Investigation of Talent Components in Educational Performance Prediction, A Data Mining Approach

Shahpar. Zahra[1], Khatibi Bardsiri. Vahid[2], and Khatami. Saideh[3]

***Abstract*—The present study aims at investigating and predicting students' performance based on educational talent components. All Birjand's 3rd-grade high school male and female gifted students are the population of the study. 195 individual were selected as the sample of the study through available sampling approach. Job Strong test results were used for assessing talent administered homogenously among the 1st-garade high-school students. Data mining techniques are used for studying and prediction applying three models of multiple linear regression, step-wise multivariable regression, and decision-making tree with cart algorithm. The results of step-wise regression model also indicate that among talent components, information sufficiency, space imagination, and mechanical reasoning are significant predictors for educational performance.**

***Index Terms*— Data Mining, Educational Performance, Job Strong Test and Talent**

I. INTRODUCTION

HUMAN being owns different mental aspects influencing different life aspects as well as each others' . For this, psychologists have always taken the recognition and understanding of the quality and quantity of these relationships into account. Education, nowadays, takes the most important part of each individual's life. The quality and quantity of education play an important role in personal life. Therefore, psychologists try to identify the predictive factors of educational improvement [1]. They have generally made use of statistical approaches and attempted to discover the best predictive model of educational performance through available information.

Educational development, as a dependent variable, is not under the influence of one factor. It is influenced by different factors including educational talent, cognitive factors like general intelligence, educational self-efficacy, self-control strategies, class structure, educational motivation, learners' capabilities, teacher training, and learners' motivation [2].

[1]PhD-student, Islamic Azad University, Kerman Branch, Iran, (Email:zahrashahpar@yahoo.com)

[2]Assistant Professor, Islamic Azad University, Kerman Branch, Iran, (Email: khatibi78@yahoo.com)

[3]Master of psychology, (Email: saideh_khatami@yahoo.com)

Given the interwoven nature of these factors and their mutual reaction toward each other, it seems less possible to distinguish each one's role. Among all, personal-educational factors with social-cognitive nature have the most influence on education development [2].

On the other side, high-speed data mining techniques have developed in different areas to discover efficient models through piles of data and create predictive models. Data mining refers to finding new relations, models, and approaches through exploring piles of saved data and/or applying artificial intelligence and statistical-mathematical techniques. Therefore, the present study aims at investigating the relationship between talent components in educational performance and data mining techniques. For this, first, the concept of talent is defined and contrasted with intelligence and skills, then a literature of educational performance prediction is provided, and finally the concept of data mining and applied models in the present study are elaborated. Finally, after explaining research population, instrument, and methodology, we will discuss the results and provide conclusion.

II. RESEARCH THEORETICAL BASIS

A. Definition of Talent

Talent is a Latin term meaning some amount of money. It also means tendency, willingness, and wish. Contemporarily, its primary use is for having natural and mental capabilities. As stated by Oxford Dictionary, this term has been used in its contemporary sense since 1600. Having talent means having good skills in doing some things and tasks [3]. Other definitions of talent are as follows

Hood and Johnson define talent as getting some specific skill or knowledge [3].

Kerno thinks of talent as the level of readiness for learning and good performance in specific situation [3].

Littre refers to talent as natural readiness for determined task and activity [4].

Shafiabadi believes that talent is an innate capability helping learning and predicting the level of individual learning in different fields [5].

Talent enjoys different provided definitions. Different statements of talent definition are arranged in the range of general talent (for instance high intelligent capability or high IQ score) and specific talent (for instance development in a specific course like math). By the way, approximately all definitions refer to remarkable superiority (distinction) and/or development (achievement) in one field [5].

B. Difference among talent, intelligence, development, and skill

It is difficult to make a distinction between intelligence and talent. Some believe that these two are synonym and/or with overlap of meaning. For example, in 1972, national definition of talent was provided in Marland's report and then repeated in America's general law. No clear distinction between talent and intelligence is made in this definition [4]. The term intelligent and talented, when used for students, children, and/or youth refers to indicative evidence of students, children, and youth's high performance in different fields including mind, creativity, art, management, and/or in specific area. To develop completely, they require something not being provided in schools [5].

On the other hand, some believe that talent and intelligence are different components. For some, talent refers to mental, art, creativity, and sport superiority. Being intelligent, in this concept, refers to the general capability of individual. Talent, relative to intelligence, is a concept more used and less criticized. It seems that talent carries hard attempt through time while intelligence is an innate bestow. Talent, in this regard, can be defined as performance criterion while intelligence is a personal trait [3].

Intelligence is referred to as the most general ability because it is in every activity requiring cognition. General intelligence is dependent to growth, while talent (like the ability of acquiring music, drawing skills or speed of action) is not dependent to individual's age and growth [4].

Like the overlap existing between intelligence and talent, this overlap exists between skill and ability, as far as talent is concerned. There is a delicate distinction between these three terms. What an individual can presently do is ability. Skill is what s/he achieved before. Talent refers to the fact that how individual can learn easily and rapidly in future [5]. Skill and talent are different. Skill is an ability that individual gained before for accomplishing tasks successfully, while talent is an innate ability helping acquisition [3].

Intelligence and talent tests are different. Intelligence tests focus on measuring the level of current abilities. Talent tests are designed for predicting individuals' ability for enjoying educational or job experience. Talent tests are mostly used for selecting individuals for educational programs, jobs, or for helping her/him to understand her/his potentials, and making job planning and decision making easier. Therefore, a talent test do not show individual's intelligence level. Difference between individual of the same age group is low in general intelligence tests, while talent test are less dependent on age. Therefore, given the distinction between talent and intelligence, separate tests are applied estimating individual's ability correctly [5]. Talent identification techniques and approaches depend on the approved theoretical basis of this concept. For this, approaches not only are in a wide range of variety, but they also change during different periods. Strong test applied in the present paper is the newest test applied in Iran for talent identification.

C. Data mining concept

Data mining is a communicative bridge among statistic science, computer science, artificial intelligence, modeling, machine acquisition, and visual replaying. Data mining is a complex procedure of identifying patterns and correct models, as new and beneficial in a wide range so as to be understandable for human. It is a procedure that should be defined as a project, and it should be implemented and controlled through pre-defined and programmed steps [6].

Data mining is applicable in various fields including computer, IT, networking, banking and financing, production, medical, customer relation management, sociology, psychology, and education [6].

Data mining algorithms include predictive algorithms including regression, data base technology systems like dependence principles and techniques based on learning like neural network and decision-making tree. Each of these algorithms are applicable regarding data mining ultimate goal and the kind of data modeling [6]. The present study makes use of three models of multiple linear regression, stepwise multivariable regression, and decision-making tree.

Multiple linear regression is one of the most applicable approaches in regression analysis in which dependent variable is a linear synthesis of coefficients. This approach is used for predicting dependent variable. Its equation is stated as follows [6]:

$$Y_i = B_0 + B_1 X_1 + E_{1,} \qquad i=1,\dots,N \tag{1}$$

This equation is a regression with N point, X1 dependent variable, linear B1 coefficients, and E1 error. I footnote shows the number of each observation (each pair of X1,Y1). Having these points, we attain the model. In this approach, the amount of error is ei= yi-yi. The most common approach attaining parameters is the approach of the least squares. In this approach, parameters are attained lessening the function:

$$SSE = \sum_{i=1}^{N} e_i^2 \tag{2}$$

In order to study the influence of some independent variables on dependent variable, we make use of step-wise multivariable regression. It shows that which independent variable can predict dependent variable in the best way, what the share of each variable is, and how they can totally predict [7].

Decision-making tree is one of the strong and common instruments for categorization and prediction. Decision-making tree produces indices and explains its prediction according to some principles. It consists of three main

components including root, internal node, and loaf. Decision-making tree's algorithms are of different kinds. The present study makes use of cart decision-making tree as a comparison algorithm model. This model introduces a non-circling graph like a tree with double categorization for introducing an identifying and categorizing model [7].

III. LITERATURE REVIEW

Background researches of educational performance prediction have been carried based on different criteria and according to statistical approaches. Some of these studies are as follows:

Delkhon (2012) in the prediction of educational performance from creativity and emotional intelligence among elementary students of Oroumieh county made use of two variables of creativity and emotional intelligence and applied descriptive inductive statistics (correlation and step-wise regression) to study and predict 5th grade elementary students' educational performance. He showed that creativity and educational performance have a significant correlation i.e. educational performance is predictable by creativity, while there is no significant correlation between educational performance and emotional intelligence i.e. educational performance is not predictable by emotional intelligence [2].

Bakhshi et al (2012) in the pattern of educational development prediction: the role of critical thinking and strategies of self-controlling learning applying Pearson's correlational statistical tests, multiple regression, the analysis of multivariable variance, and covariance analysis showed that the share of self-controlling learning, critical thinking, and average score have relationship with students' educational development. The coefficients of two variables of self-controlling learning and diploma average score were significant and critical thinking was responsible for no significant variance in educational development scores [8].

Vahedi (2012) in educational development prediction of Physics students: the role of background knowledge, gender, the understanding and attitude toward physics learning applying regression analysis showed that the score of chemistry and gender were the most important predictors of educational performance in physics course [9].

In (2012), Yadollahi et al, in predicting educational development according to critical thinking and sociology components through predictive statistics including mean, standard variation, the lowest and the highest range, and inductive statistics including step-wise regression showed that there was no significant relationship between sub scales of critical thinking and educational success, and there was only significant relationship between sub scales of interpretation and educational success. As indicated by the regression coefficient of interpretation prediction variable, interpretation prediction variable can well explain the relevant changes of educational success [10].

Ashuri et al (2010) in the investigation of personal features, resource management strategies and motivational strategies in low-ability students' educational development prediction applying step-wise regression came to the conclusion that acceptance, accountability, attempt orderings, contentment, skill-oriented objective, performance avoidance objective, and self-efficacy have significant positive relationship with educational development, while mental disorder and exam anxiety have significant negative relationship with educational development. Self-efficacy, accountability, skill-oriented objective, and mental disorder can well predict 45% of educational development variance, and self-efficacy predicts the most of educational development [11].

In 2006, Zahrakar, in the investigation of relationship between emotional intelligence components and educational performance studied the correlational coefficient between emotional intelligent components and educational performance. For this, all students of Islamic Azad university of Islamshahr were selected as population of the study, and step-wise regression analysis was used. The results indicated that emotional intelligence components, optimism components, satisfaction, stress tolerance, problem solving and controlling are significant predicators for educational performance [12].

Sarmadi et al (2011) in the investigation of correlated factors with third grade secondary students' educational development according to the results of TIMSS-R and developing track analysis pattern for studying the influence of each of factors on educational development studied the relevant factors and their direct and indirect influence through correlational matrix and studied the pattern considering Hyniko index. They showed that educational activities of teacher in class and school have the most direct and indirect influence, respectively. However, the most explanation of educational development variable variance was reported for teacher's educational experience and traits [13].

Atashrouz (2007) in educational development prediction through the level of affective dependence identified the level of relationship between educational development through being dependent on mother, father, and others with the same age and other dimensions including trust, relationship, being far from parents and other groups with the same age. Methodologically, this research was correlational and the results indicated that there is a significant relationship between being dependent on mother, father, and others with the same age and educational development. Being dependent to mother showed the most predictability of educational development [1].

IV. METHODOLOGY

A. Research hypotheses

According to the theoretical basis of the study, research hypotheses are developed as follows in Table 1.

B. Population and sampling

The population of the study consists of all male and female students of Birjand's third-grade high school students. 195 individuals (110 females, 85 males) were selected as samples

through available sampling approach.

C. Instruments

Job strong questionnaire was applied in this study to collect the required data of talent components. This questionnaire was translated in Isfahan in 2005, and its reliability and validity were confirmed. It is known as test of job-education willingness. This questionnaire consists of 76 statements organized in 6 parts. It can be applied for high school students, college students, and adults. Therefore, it can be used for the range of age of 50. This instrument is not usually used for students younger than 13 or 14, since willingness patterns do not develop in this age to be identified. Currently, this test is used for first-grade high school group students and administered nationally homogenously with talent identification test. These tests comprise of 11 sub tests including 1- vocabulary meaning 2- speech reasoning 3- information sufficiency 4-mathematic principles 5- digital reasoning 6- abstract reasoning 7- logical reasoning 8- space imagination 9- figure turning 10-segment synthesis 11- mechanical reasoning. The objective of running this test is to identify students' talents in different job-educational areas.

The present study, to investigate students' educational performance, made use of final average scores of tests held nationally.

Table 1: Research Hypotheses

1- There is a relationship between talent components of vocabulary meaning and educational performance.
2- There is a relationship between talent components of speech reasoning and educational performance
3- There is a relationship between talent components of information sufficiency and educational performance
4- There is a relationship between talent components of mathematic principles and educational performance
5- There is a relationship between talent components of digital reasoning and educational performance.
6- There is a relationship between talent components of abstract reasoning and educational performance.
7- There is a relationship between talent components of logical reasoning and educational performance.
8- There is a relationship between talent components of space imagination and educational performance.
9- There is a relationship between talent components of figure turning and educational performance.
10- There is a relationship between talent components of segment synthesis and educational performance.
11- There is a relationship between talent components of mechanical reasoning and educational performance.
12- Talent components can be applied in predicting educational performance through applying multiple linear regression.
13- Talent components can be applied in predicting educational performance through applying step -wise multivariable regression.
14- Talent components can be applied in predicting educational performance through decision-making tree model

As mentioned before, in this paper, we made use of data mining techniques, multiple linear regression, step-wise multivariable regression and decision-making tree with cart algorithm in Matlab software. As indicated by Fig. 1, after collecting and arranging data, all independent variable (talent components) are normalized and arranged in the range of 0-1. Correlation between independent and dependent variables (educational performance) with correlational coefficient of 0.01 is then investigated. Next, three data mining techniques are applied for prediction. It should be noted that in order to assess the results of running models, 3-fold standard approach was used. Moreover, to improve the accuracy, models were run 30 times and results were reported based on the average of these 30 times.

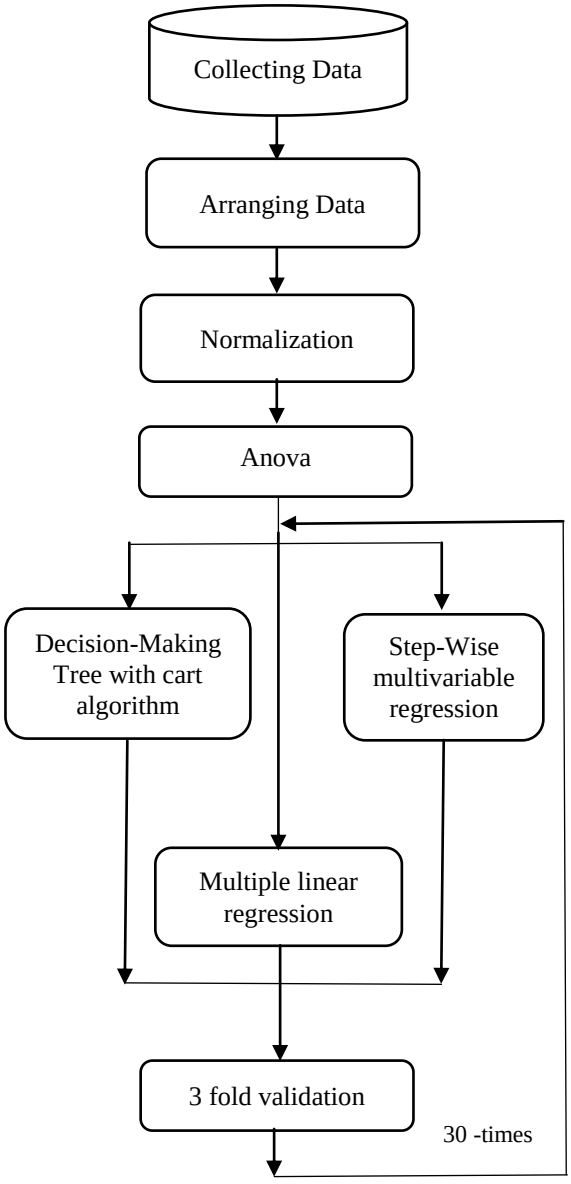

Fig. 1: Methodology

V. RESULTS

In this section, first we describe the obtained data of the research through mean and standard deviation. The relationship between talent components and educational performance is then investigated through correlational coefficient and significance test. Finally, the level of predictability of talent's different components of educational performance is studied through analysis of linear regression, step-wise multivariable regression, and cart regression.

The mean and standard deviation of population of talent components and educational performance is provided in Table 2.

Table 2: The mean and standard deviation of population of talent components and educational performance

Variables	Statistical Indices	
	Mean	Standard Deviation
Vocabulary meaning	100.8467	15.4778
Speech reasoning	84.8830	24.2502
Information sufficiency	115.4623	14.7675
Mathematic principles	99.9174	90.6324
Digital reasoning	113.0042	13.5320
Abstract reasoning	105.1459	13.9041
Logical reasoning	109.6708	12.9688
Space imagination	107.3022	22.1146
Figure turning	110.2230	13.8396
Segment synthesis	101.8881	13.1081
Mechanical reasoning	104.2054	10.6514
Educational performance	18.51334	1.3732

As indicated by the results of Table 3, the obtained results indicated that students' educational performance have significant relationship with 11 components of talent.

In order to determine the relative influence of each dependent variable on relevant changes of educational performance, as mentioned in methodology section, multiple linear regression was run 30 times among which only 16 times predicting variables enters the model. As indicated by the results of Table 4, among all independent variables, information sufficiency, space imagination, and mechanical reasoning are high predictors of educational performance. Step-wise multivariable regression has the predictability of educational performance with error average of 1.0060 and standard deviation of 0.9634.

In order to determine the predictability rate of talent components of educational performance, in addition to step-wise multivariable regression, two other models of multiple linear regression and decision-making tree with cart algorithm are also run 30 times in Matlab software. As indicated by Table 5, results showed that these models predict educational performance based on talent components with error average of 2.3290 and 1.2800 and standard deviation of 5.4850 and 1.2568, respectively.

Table 3: the results of correlation coefficient significance between talent components and educational performance

Variables	Statistical Indices	
	Correlation coefficient	Significance level
Vocabulary meaning and Educational performance	2.2123e-259	0.01
Speech reasoning and Educational performance	1.6225e-259	0.01
Information sufficiency and Educational performance	1.0567e-258	0.01
Mathematic principles and Educational performance	4.4890e-264	0.01
Digital reasoning and Educational performance	7.3586e-259	0.01
Abstract reasoning and and Educational performance	1.2681e-259	0.01
Logical reasoning and Educational performance	6.9784e-259	0.01
Space imagination and Educational performance	3.2025e-259	0.01
Figure turning and Educational performance	3.9242e-257	0.01
Segment synthesis and Educational performance	9.9428e-261	0.01
Mechanical reasoning and Educational performance	1.1025e-259	0.01

Table 4: the results of step-wise multivariable regression of talent components on educational performance

	Predicting variables entered into the model	Mean error	Error Standard Deviation
1	Mechanical reasoning	1.0096	0.9664
2	Mechanical reasoning and space imagination	0.9907	.09575
3	Mechanical reasoning and information sufficiency	1.0203	0.9722
4	Mechanical reasoning and information sufficiency	1.0065	0.9639
5	Mechanical reasoning	1.0066	0.9505
6	Mechanical reasoning , digital reasoning and abstract reasoning	1.0610	0.9703
7	Mechanical reasoning and information sufficiency	1.0301	0.9856

8	Mechanical reasoning	0.9988	0.9770
9	Mechanical reasoning and information sufficiency	1.0261	0.9595
10	Mechanical reasoning and information sufficiency	1.0299	0.9722
11	Information sufficiency	1.0106	0.9424
12	Mechanical reasoning and space imagination	1.0369	0.9913
13	Space imagination	1.0115	0.9846
14	Mechanical reasoning and abstract reasoning	1.0171	0.9831
15	Information sufficiency and logical reasoning	1.0292	0.9686
16	Mechanical reasoning and information sufficiency	1.0054	0.9525
17	Mean	1.005597	0.963487

Table 5: Error mean and error standard deviation in educational performance prediction according to talent components

	Data mining approach	Mean error	Error Standard Deviation
1	Step-Wise multivariable regression	2.328913	5.485017
2	Multiple linear regression	1.005597	0.963487
3	Decision-making tree with cart algorithm	1.279587	1.256783

VI. CONCLUSION

As indicated by the results of the study, correlational coefficient in the level of $p<0.01$ is significant. Therefore, we conclude that there is a correlation between educational performance and all components of talent including vocabulary meaning, speech reasoning, information sufficiency, mathematic principles, digital reasoning, abstract reasoning, logical reasoning, space imagination, figure turning, segment synthesis, and mechanical reasoning. In other words, the more an individual enjoys talent component, the higher s/he enjoys educational performance.

As indicated by the results of table 4, among all talent components, information sufficiency, space imagination, and mechanical reasoning are important and significant predictors of educational performance.

The results of table 5 also indicated that talent components are strong predictors in educational performance. Among data mining techniques, step-wise multivariable regression and decision-making tree are respectively the best models for educational performance prediction with the least error of approximately 1 score.

REFERENCES

[1] B. Atashrouz, S. Pakdaman, and A. Asgari, "educational development prediction through the level of affective dependence," *Family research journal*, 4th year, vol. 14, pp. 193-203, 2007.

[2] S. Delkhon, and M. Dehghani, "The prediction of educational performance from creativity and emotional intelligence among elementary students of Oroumieh county," *5th national education conference*, Iran, 2012.

[3] T. F. Craft, and E. Zomkeh, *Talent discovery and growth.* Translated by M. T. Taher, Tehran, 2006.

[4] A. Hoseini, "creativity and its relationship with personal features of intelligence and mental health," *Roshd journal*, vol. 66, 2010.

[5] M. Ayati, "Educational materials of Tarhe Shahab (Plan of identifying and leading top talents)," Talent National Foundation, Iran, 2nd ed, 2011.

[6] J. Hosseinkhanin, *Data mining. an approach toward concepts and algorithms*. Yegan, 2012.

[7] J. Shahrabi, *Data mining 2*. 2nd ed. Tehran. Jahad Daneshgahi, Amirkabir Industrial Unit, 2010.

[8] M. Bakhshi, and M. R. Ahanchian, "Educational development prediction pattern: the role of critical thinking and delf-control learning strategies," *Iranian journal of education in medical science*, 2n year, vol. 13, pp.153-156, 2012.

[9] S. Vahedi, and M. Yari, "Physics educational development: the role of background information, age learning understanding, and attitude toward physics," *14th conference of Iran physics education; and 4th conference of physic and laboratory*, Farhangian University, Tehran, 2012.

[10] A. Yadollahi, L. Fanni, A. Mirzazadeh, and A. Soltani, "Predicting educational development according to critical thinking and sociology," *Journal of educational development in medical science*, 6th year, vol.11, Peps 72-80, 2012.

[11] M. Ashuri, "The investigation of personal features, resource management strategies and motivational strategies in low-ability students' educational development prediction," *Journal of low-ability studies*, 1st year, No.1, 2010.

[12] K. Zahrakar, "The investigation of relationship between emotional intelligence components and educational performance," *Journal of applied psychology*, 2nd year, vol.5, pp. 89-98, 2006.

[13] M. R. Sarmadi, M. H. Sayf, S. Talebi, and S. Abedi, "in the investigation of correlated factors with third grade secondary students' educational development according to the results of TIMSS-R and developing track analysis pattern for studying the influence of each of factors on educational development," *New educational approaches, psychology and education College of Isfahan University*, 5th year, 1st vol., pp. 1-30, 2011