Academic performance prediction algorithm based on fuzzy data mining

Anil Kumar Tiwari[1], G. Ramakrishna[2], Lokesh Kumar Sharma[3], Sunil Kumar Kashyap[4]
[1,2]Department of Computer Science and Engineering, K L University, Vijayawada, India
[3]National Institute of Occupational Health, Ahmadabad, Gujarat, India
[4]Department of Mathematics, School of Advanced Sciences, Vellore Institute of Technology University, Vellore,
Tamil Nadu, India

Article Info	ABSTRACT
	This paper presents an algorithm for prediction of academic performance of students by fuzzy data mining. The fuzzy-trace concept applied to predict the academic performance of the students. An algorithm is proposed in this paper lies with this idea. The fuzzy academic set is generated from the student's academic data. This is analyzed by the fuzzy-matrix set. The prediction academic data is referred as the management of data or data mining. Data mining is the science of analyzing the data for obtaining more information than the current information. The hidden information appears by this technique.

Corresponding Author:

Anil Kumar Tiwari,
Department of Computer Science and Engineering,
K L University,
Vijayawada, India.
Email: anil1969_rpr@yahoo.com

1.   INTRODUCTION

Every academic institute has the various data of the students. These have been used for the limited purpose. But by the data mining technique various hidden information can be found. These information is discrete than the existed information. Fuzzy set introduce by Zadeh [1] in 1965. The characterization of the set defined as the fuzzy set. The elements of the fuzzy set lies with the values under the range of [0,1]. This set theory is established by grading or membership function. Shafer et al [2] presented an idea on hyper trees in 1988.The local computation is applied in this paper to study the hyper trees. The uncertainty theory is developed in 1992 by Shenoy [3]. The uncertainty is studied in this paper for the managing the expert systems. The uncertainty is defined dynamically by using the fuzzy set. In 1993, an uncertainty is extended with vagueness by Gebhardt et al [4]. The compact computation is performed in this paper by the fuzzy rule. The common results are discussed for adjoining this two distinct elements.

Kruse et al [5] generalized the fuzzy systems. The application of fuzzy systems is presented in this paper. In 1996, Fayyad et al [6] proposed the new theory of data mining. The knowledge management and data mining are adjoined together with the fuzzy logic. Borgelt et al [7] applied the probabilistic and possibilistic theory for learning systems in 1997. The learning method is developed for measuring the performance by the combinatorial techniques. Berthold et al [8] presented the fuzzy graphs in 1999. The several social illustration is midelled by the fuzzy graphs. The analysis of the examples also presented as model in this paper. In 1999, Bezdek et al [9] presented some algorithms for pattern recognition and image processing by the fuzzy techniques. The fuzzy model and the formulation of the data of are presented by the algorithm performs

efficiently. In 2010, Romero et al [10] reviewed the educational data mining. The new results and techniques are established in this paper.

Chalaris et al [11] presented the data mining technique for improving the educational quality. The new knowledge management system is established by the fuzzy theory. The advancement of the knowledge system is studied in this paper. In 2012, Krapn et al [12] introducing the e-learning system for the students. The grouping of data of the students is studied by the data mining technique. The educational data is analyzed by the fuzzy theory in this paper. In 2010, Delen [13] studied the student's retention management system over the machine learning system. The analysis of educational data by machine learning technique is presented in this paper. Same year, Buldua et al [14] presented some applications of data mining for the academic data. The academic data is analyzed and formulated for predicting the performance of the students. Recently, Kaura et al [15] developed a method to improve the performance of the slow learner students. The prediction technique is presented for the slow learners by the data mining technique. The education sector is used the proposed technique with the improved results. Campagini et al [16] designed some data models to improve the performance of the students by the data mining techniques. The career of the students is predicted by this technique. Shukora et al [17] applied the data mining technique for the effective learning. The examination based on the online system is studied. The effectiveness of the online learning is modeled by data mining technique. The result and performance analysis by data mining technique is conducted in this paper. The student's academic performance correspondence to the elements of the academic set is focused in this research. The data mining technique is applied on the academic data to map with the intelligence parameters. Alfiania et al [18] studied in this paper in 2015.

2. FUZZY SET [1]

Let X be a space of points (objects), with a generic element of X denoted by x. Thus, $X = \{x\}$. The fuzzy set A in X is characterized by a membership (characteristic) function $f_A(x)$ which associates with each point in X a real number in the interval [0,1] with the value of $f_A(x)$ at x representing the grade of membership of x in A. Next, the academic data is defined.

3. ACADEMIC DATA

This is information of academic performance of students. The following matrix is represented the academic data as:

$$S = \begin{bmatrix} s_{11} & . & . & . & s_{1n} \\ . & . & . & & . \\ . & . & . & . & . \\ . & . & . & & . \\ s_{m1} & . & . & . & s_{mn} \end{bmatrix}_{m \times n} ,$$

Where,

$$students : i = 1,2,...,m.$$

$$performance : j = 1,2,...,n.$$

The fuzzy set is defined over the academic data in next.

4. FUZZY ACADEMIC SET

Let S be a set of academic performance of students, then A is the fuzzy set characterized by "performer" defined as below:

$$f_A(s_{11}) = 0,$$

Or,
$$f_A(s_{12}) = 0,$$
Or,
$$f_A(s_{1n}) = 0.2,$$
Or,
$$f_A(s_{21}) = 0.3,$$
Or,
$$f_A(s_{m1}) = 0.8,$$
Or,
$$f_A(s_{mn}) = 1.$$

Hence,
The fuzzy set matrix is represented by,

$$A = \begin{bmatrix} 0 & . & . & . & 0.2 \\ . & . & . & & . \\ . & . & . & . & . \\ . & . & . & & . \\ 0.8 & . & . & . & 1 \end{bmatrix}_{m \times n}.$$

Next, we study the prediction of academic performance of students by fuzzy.

5. THEORY OF PREDICTION OF ACADEMIC PERFORMANCE OF STUDENTS BY FUZZY

Let, The matrix of expected performance be,

$$B = \begin{bmatrix} 1 & . & . & . & 0.8 \\ . & 1 & . & & . \\ . & . & 1 & . & . \\ . & . & . & 1 & . \\ 0.8 & . & . & . & 1 \end{bmatrix}_{m \times n}.$$

The statement of the problem is, "how to find the followings?"

Or,
$$f_B(s_{11}) = 1,$$
Or,
$$f_B(s_{11}) = 1,$$
Or,
$$f_B(s_{1n}) = 1,$$
Or,
$$f_B(s_{m1}) = 0.8,$$
Or,
$$f_B(s_{mn}) = 1.$$
Now, we set the theory for achieving the above.
By contradiction,
$$f_B(s_{mn}) \neq 1.$$
Or,

$f_B(s_{mn}) < 1.$
Or,
$f_B(s_{mn}) = 0,$
Or,
$f_B(s_{mn}) = 0.1,$
Or,
$f_B(s_{mn}) = 0.2,$
Or,
$f_B(s_{mn}) = .9.$
Then,
Performer = P, where P;
$P = A \cup B,$
Or,
$P = \max\{f_A(s), f_B(s)\}.$
Or,
$P = f_A \vee f_B.$
Or,

$$P = \begin{bmatrix} f_A(s_{11}) & & \\ & & \\ & & f_A(s_{mn}) \end{bmatrix} \vee \begin{bmatrix} f_B(s_{11}) & & \\ & & \\ & & f_B(s_{mn}) \end{bmatrix}.$$

Or,

$$P = \begin{bmatrix} f_A(s_{11}) \vee f_B(s_{11}) & & \\ & & \\ & & f_A(s_{mn}) \vee f_B(s_{mn}) \end{bmatrix}.$$

Or,

$$P = \begin{bmatrix} f_C(s_{11}) & & \\ & & \\ & & f_C(s_{mn}) \end{bmatrix}.$$

Or,
The trace of P is defined by,

$$Trace(P) = f_C(s_{11}) + ... + f_C(s_{mn}).$$
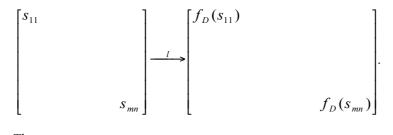
Or,

$$Trace(P) = f_D(s_{mn}).$$

Or,

$$Trace(P) = n.$$

Where n is a numeric value. The defuzzification of this value gives the linguistic variable (l) which will be the prediction defined as,

$$Trace(P) = l.$$

This result maps to the individual student:

$$\begin{bmatrix} s_{11} & & \\ & & \\ & & \\ & & s_{mn} \end{bmatrix} \xrightarrow{\ l\ } \begin{bmatrix} f_D(s_{11}) & & \\ & & \\ & & \\ & & f_D(s_{mn}) \end{bmatrix}.$$

Then,
$$f_D(s_{11}) = f(s_{11}),$$
.
.
.
$$f_D(s_{mn}) = l(s_{mn}).$$

Now, the l-function will be analyzed by fuzzy cognitive map. The prediction after the time (t) presents by the following map:

If,
$$f_D(s_{mn}) = 1,$$
Then,
$$l(s_{mn}) = 1.$$
But,

$$\begin{bmatrix} l & 0 & 0 & 0 & 0 \\ 0 & l & 0 & 0 & 0 \\ 0 & 0 & l & 0 & 0 \\ 0 & 0 & 0 & l & 0 \\ 0 & 0 & 0 & 0 & l \end{bmatrix} \begin{bmatrix} p_1(t) \\ \\ \\ \\ p_5(t) \end{bmatrix} = 1,$$

Then,

$$\begin{bmatrix} s_1(t+1) \\ \\ \\ \\ s_5(t+1) \end{bmatrix} = f(\begin{bmatrix} & s_{11} & & s_{1n} \\ s_{21} & & & \\ & & & \\ & & & \\ & & & \end{bmatrix} \begin{bmatrix} s_1(t) \\ \\ \\ \\ s_n(t) \end{bmatrix} + \begin{bmatrix} l & 0 & 0 & 0 & 0 \\ 0 & l & 0 & 0 & 0 \\ 0 & 0 & l & 0 & 0 \\ 0 & 0 & 0 & l & 0 \\ 0 & 0 & 0 & 0 & l \end{bmatrix} \begin{bmatrix} p_1(t) \\ \\ \\ \\ p_5(t) \end{bmatrix}).$$

Or,

$$\begin{bmatrix} s_1(t+1) \\ \\ \\ \\ s_5(t+1) \end{bmatrix} = f(\begin{bmatrix} & 0.6 & & s_{1n} \\ 0.7 & & & \\ & & & \\ & & & \\ & & & s_n(t) \end{bmatrix}\begin{bmatrix} s_1(t) \\ \\ \\ \\ s_n(t) \end{bmatrix} + \begin{bmatrix} l & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} p_1 \\ \\ \\ \\ p_5(t) \end{bmatrix}).$$

Or,

$$\begin{bmatrix} s_1(t+1) \\ \\ \\ \\ s_5(t+1) \end{bmatrix} = f(\begin{bmatrix} & 0 & & s_{1n} \\ 0 & & & \\ & & & \\ & & & \\ & & & s_n(t) \end{bmatrix}\begin{bmatrix} s_1(t) \\ \\ \\ \\ s_n(t) \end{bmatrix} + \begin{bmatrix} l & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} p_1(t) \\ \\ \\ \\ p_5(t) \end{bmatrix}).$$

Or,

$$\begin{bmatrix} s_1(t+1) \\ \\ \\ \\ s_5(t+1) \end{bmatrix} = f(\begin{bmatrix} & 1 & & s_{1n} \\ 1 & & & \\ & & & \\ & & & \\ & & & s_n(t) \end{bmatrix}\begin{bmatrix} s_1(t) \\ \\ \\ \\ s_n(t) \end{bmatrix} + \begin{bmatrix} l & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} p_1(t) \\ \\ \\ \\ p_5(t) \end{bmatrix}).$$

Our objective is to predict the performance of student on the basis of existed data by an algorithm. This is presented in below:
–   Input the crisp data.
–   Set the linguistic variable.
–   Generate the fuzzy set.
–   Analyze by fuzzy operation.
–   Establish the fuzzy trace.
–   Execute (5).
–   Error reduction.
–   Find new crisp set.
–   Defuzzification.
–   Inverse map of fuzzy function.
–   Prediction.

6. CONCLUTION

The academic performance prediction is the modern challenge in the society. The objective of every academic training is to achieve the good result. Thus, the proposed method will be assisted for reaching to the goal. The academic data comprises with the numeric and linguistic both. Fuzzy lies with both thus this is applied to develop this algorithm.

REFERENCES

[1]   L. A. Zadeh, Fuzzy Sets, Information and Control, 8, 338-353, 1965.
[2]   G. Shafer and P.P. Shenoy. Local Computations in Hyper trees, Working paper 201. School of Business, University of Kansas, Lawrence, KS, USA 1988.
[3]   P.P. Shenoy. Valuation-based Systems: A Framework for Managing Uncertainty in Expert Systems. In: L.A. Zadeh and J. Kacprzyk, eds. Fuzzy Logic for the Management of Uncertainty, 83{104. J. Wiley & Sons, New York, NY, USA 1992.
[4]   J. Gebhardt and R. Kruse. The Context Model | An Integrating View of Vagueness and Uncertainty. *Int. Journal of Approximate Reasoning* 9:283{314. North-Hollan, Amsterdam, Netherlands 1993.
[5]   R. Kruse, J. Gebhardt, and F. Klawonn. Foundations of Fuzzy Systems. J. Wiley & Sons, Chichester, United Kingdom 1994.

[6]   U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, eds. Advances in Know ledge Discovery and Data Mining. MIT Press, Menlo Park, CA, USA 1996.

[7]   C. Borgelt and R. Kruse. Evaluation Measures for Learning Probabilistic and Possibilistic Networks. Proc. 6th IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE'97, Barcelona, Spain), 669{676. IEEE Press, Piscataway, NJ, USA 1997.

[8]   M. Berthold and K.-P. Huber. Constructing Fuzzy Graphs from Examples. *Int. J. of Intel ligent Data Analysis*, 3(1), 1999. Electronic journal (http://www.elsevier.com/locate/ida).s

[9]   J.C. Bezdek, J.M. Keller, R. Krishnapuram, and N. Pal. Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. The Handbooks on Fuzzy Sets. Kluwer, Dordrecht, Netherlands 1999.

[10]   Crist´obal Romero, Member, IEEE, and Sebasti´an Ventura, Senior Member, IEEE, "Educational Data Mining: A Review of the State of the Art " VOL. 40, NO. 6, November 2010.

[11]   Manolis Chalaris, Stefanos Gritzalis, Manolis Maragoudakis, Cleo Sgouropoulou and Anastasios Tsolakidis," Improving Quality of Educational Processes Providing New Knowledge using Data Mining Techniques*", Science Direct - Procedia - Social and Behavioral Sciences* 147; 390-397, 2014.

[12]   Krpan, Slavomir Stankov, "Educational Data Mining for Grouping Students in E-learning System", Proceedings of the ITI 2012 34th Int. Conf. on Information Technology Interfaces, June 25- 28, 2012, Cavtat, Croatia.

[13]   Dursun Delen, "A Comparative Analysis of Machine Learning Techniques for Student Retention Management", *Science Direct - Decision Support Systems* 49; 498-506, 2010.

[14]   Ali Buldua, Kerem Üçgüna, "Data Mining Application on Students' Data", *Science Direct - Procedia Social and Behavioral Sciences* 2; 5251-5259, 2010.

[15]   Parneet Kaura, Manpreet Singhb ,Gurpreet Singh Josanc "Classification and Prediction based Data Mining Algorithms to Predict Slow Learners in Education Sector", Science Direct Procedia Computer Science 57; 500-508, 2015 (ICRTC- 2015).

[16]   Renza Campagni, Donatella Merlini, Renzo Sprugnoli, Maria Cecilia Verri, "Data Mining Models for Student Careers", *Science Direct - Expert Systems with Applications* 42; 5508–5521, 2015.

[17]   Nurbiha A Shukora , Zaidatun Tasira, Henny Vander Meijden, "An Examination of Online Learning Effectiveness using Data Mining", *Science Direct - Procedia - Social and Behavioral Sciences* 172, 555-562, 2015.

[18]   Harwatia, Ardita Permata Alfiania, Febriana Ayu Wulandari," Mapping Student's Performance Based on Data Mining Approach", *Science Direct Agriculture and Agricultural Science Procedia* 3; 173-177, 2015.