Big Data Management by Fuzzy, Neural Network and Genetic Algorithm

Anil Kumar Tiwari, G. Ramakrishna, Lokesh Kumar Sharma, Sunil Kumar Kashyap

Abstract: ***This paper manages the academic data by the dynamic techniques. The data may have the infinite information. This infinite information transforms into the finite information by the dynamic algorithm. This dynamic algorithm consists fuzzy logic, neural and genetic algorithm. Thus the result lies the data analysis from Data Mining to Dynamic Data Mining. New techniques are introduced here for redefining the database and its analysis. The database Student's Academic Performances is selected for the generalization of the proposed method. It is all is studied over Fuzzy, Neural Network and Genetic Algorithm.***

Index Terms: Data Mining (DM), Dynamic Data Mining (DDM), Database (DB), Student's Academic Performance (SAP), Neural Network (NN), Genetic Algorithm (GA).

I. INTRODUCTION

"Several data in single data" or "Various information in single data" or … there may be several natural questions and answer of all these can be obtained by either fuzzy or neural network or may be another method and its optimization either by Genetic Algorithm or any method. This paper deals on the raised. 1965 is the year, when Fuzzy came in the existence. The measurement is the tool only for measuring the quantity but no idea is described about the quality. More precisely, the grading or the membership in the measurement, it is essential sometimes, when we expect more information from the current database. The Theory of Fuzzy Sets is introduced by Lotif A. Zadeh [7]. Since that year, there are several applications has been presented in different disciplines. Engineering, Medical, Biology, Management, Computer Science, Mathematical Science, etc. are some major sciences, where fuzzy set theory has been used frequently. Shao [5] designed a controller which was a self organized for the study of dynamic process in the year 1988. Again in the year 1988, Tanscheit et.al. [6] presented an application on robotics. They also presented a survey on fuzzy sets and applications in industrial applications in the year 1985. Lee [1] designed a control system based on fuzzy logic in the year 1990. The father of fuzzy sets, Zadeh [8] again presented a revolutionary theory on fuzzy logic in 1990. This is also became another landmark in the world of application sciences. Self [3] proposed a designing theory based on fuzzy logic in 1990. Li et.al. [2] shown a demonstration on fuzzy applications in servo systems in the year 1989. The similar type of this work had presented in the year 1985, by Schaff et.al. [4] for robotics controller.

Data mining lies with the following four conditions:

1. Correspondence from old data to new data.
2. Functional system of sequential data.
3. Unique final interpretation.
4. Transforming the information Local to Global.

Beyond these, dynamic data mining states on more other objectives which does not carry with the classical data mining. Finally the objective remains same, i.e. more and more information could be found from the database. Therefore, various methodologies are existed for getting this unique objective. Clustering, Fuzzy, Neural, Genetic, Graph, Tree, Associate Rule etc. are the some modern methodologies, which are basically based on, classical or classical type characteristics. The cycle of information is observed in the data. Its direction and dimension are the two major challenges in data mining. Although, in a cycle for the classification of the data the followings are required;

a. New Class Creation.
b. Class Elimination.
c. Class Movement.

Thus, we propose fuzzy, Neural and Genetic Algorithm corresponds to the (a), (b) & (c).

Key Methodology: A proposition is delivered which is the necessary and sufficient reason for FS-ANN-GA used together.

The fundamental Proposition on Fuzzy-Neural-Genetic Algorithm is stated below:

The Statement: Let the data of SAP be x, its fuzzy set be $A=(x,\mu_A(x))$, the ANN be $\sum w(A)$ and GA be $f(\sum w(A))$, then its equivalence data is y, which is smaller in size with the same characteristics.

Proof: Let the set be $X=\{x_1,...,x_n\}$. Either this is in ordered set or unordered set. If this is an ordered set then there is a relation. The elements of the set $X \in R$. The fuzzy set $A=[0,1] \in R$.

Let, the number of layers in ANN be *n(L).*

$n(L) \in R$ then it can be assumed that there may be at least one linguistic/numerical objective which is fixed for the expected outcome from the data. Let this be any relation. If this relation is defined as ordered set then it can be optimized by GA. This ordered set is weighted by the ANN.

Let, E be a subset of X, then there exists an element which belongs to the set X as $\beta \in X; x \in E.$

Revised Version Manuscript Received on March 30, 2017.

Anil Kumar Tiwari, Department of Computer Science, KL University, Vijayavada, Andhra Pradesh, India, E-mail: anil1969_rpr@yahoo.com

G. Ramakrishna, Department of Computer Science, KL University, Vijayavada, Andhra Pradesh, India, E-mail: ramakrishna_10@yahoo.com

Lokesh Kumar Sharma, National Institute of Occupational Health, Ahemdabad, India, E-mail: lksharmain@gmail.com

Sunil Kumar Kashyap, Department of Mathematics, School of Advanced Sciences, VIT University, Vellore, Tamil Nadu, India E-mail: 7sunilkumarkashyap@gmail.com

Then E is bounded above and β is an upper bound of E.

Let $\alpha \in X$, where α is an element which is synonym to that linguistic variable which is selected previously.

Hence, $\alpha = \sup E$

But not, $\alpha = \inf E$

By GA,

$$f(\alpha) = f_1(\alpha) + f_2(\alpha) + ... + f_n(\alpha),$$

where,

$$f_1(\alpha) = c_1 E$$

.

.

.

$$f_n(\alpha) = c_n E$$

$c = (c_1, c_2, ..., c_n)$ is a sequence.

Then, $\min f(\alpha)$ and $|f_{i-1} - f_i| \in R$.

And the fractional relation is defined by,

$$\left| \frac{f_1}{f_2} + \frac{f_2}{f_3} + ... + \frac{f_{n-1}}{f_n} \right| \in R$$

&

$$\left| \frac{c_1}{c_2} + \frac{c_2}{c_3} + ... + \frac{c_{n-1}}{c_n} \right| \in R$$

Hence.

$$y = \min f(x) = \sup E.$$

This completes the proof.

Proposed Results: The algorithm and illustration is presented in below:

Algorithm: The database is analyzed as per the following steps:

5.1.*Input*.

5.2.x.

5.3.$(x, \mu(x))$.

5.4.$\sum w_n$.

5.5.$f(\sum w_n)$.

5.6.$\sup E$.

5.7.$\min f(x)$.

5.8.y.

5.9.*Output*.

Next, its illustration is given.

Illustration:

6.1.*Input*.

6.2.$\{5, 7, 6\}$.

6.3.$\{(5, 0.5), (7, 0.7), (6, 0.6)\}$

6.4.$\{1.8\}$

6.5.$\{0.6\}$

6.6.$\{0.5\}$

6.7.5.

6.8.*Average*

6.9.*Output*

II. CONCLUSION

Zadeh [30,31] represented the set not only as per the aspect of membership but it is the relativity theory of set. It's a natural law. The classical set theory is the real mathematics but it is the actual mathematics. SAP is still challenge, because it is all for the truth. Truth cannot be classified but justified and verified. The current work is the one step towards this. The paper lies the data analysis over reality very closely.

ACKNOWLEDGMENT

We are very thankful to the reviewers for their valuable remarks.

REFERENCES

1. Lee C. C., Fuzzy logic in control systems: Fuzzy logic controller-Part I & II, IEEE Trans. Syst., Man, Cybern, SMC 20, 2, 1990, 404-435.
2. Li Y. F., Lan C. C., Development of fuzzy algorithms for servo systems, IEEE cont. sys. mag., 1989, 65-72.
3. Self K. L., Fuzzy logic design, IEEE spectrum,27, 1990, 42-44.
4. Scharf E. M., Mandic N. J., The application of a fuzzy controller to the control of a multi-degree-freedom robot arm, in industrial application of fuzzy control, M. Sugeno Ed. Amsterdam: North-Holand, 1985, 41-62.
5. Shao S., Fuzzy self-organizing controller and its application for dynamic processes, Fuzzy sets systems, 26, 1988, 151-164.
6. Tanscheit R., Schraf E. M., Experiments with the use of a rule based self-organizing controller for robotics applications, fuzzy sets systems, 26, 1988, 195-214.
7. Zadeh L. A., Fuzzy Sets, Information and control, 8, 1965, 338-353.
8. Zadeh L. A., Fuzzy Logic, IEEE Computer Magazine, 1988, 83-93.

AUTHORS PROFILE

Anil Kumar Tiwari received his graduation and post graduation degree from Pt. Ravishankar Shukla University Raipur, Chhattisgarh, India. His area of research is Data mining.

G. Ramakrishna is a renowned professor of computer science in India. Currently he is a senior professor in KL University, Vijaywada, Andhra Pradesh, India

Lokesh Kumar Sharma is a computer scientist in National Institute of Occupational Health, Ahmadabad, Gujarat, India. He completed his education from Pt. Ravishankar Shukla University Raipur, Chhattisgarh, India. Computer Architecture is the major area of his field of research.

Sunil Kumar Kashyap is a cryptographer. He is the life member of Cryptology Research Society of India, Indian Statistical Institute, Kolkata, India.