Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 45 (2015) 753 – 761

Procedia Computer Science

International Conference on Advanced Computing Technologies and Applications (ICACTA-2015)

Evolving fuzzy min-max neural network for outlier detection

Nilam Upasani[a], Hari Om[b]

[a]Research Student, Indian School of Mines, Dhanbad, India
[b]Assistant Professor, Indian School of Mines, Dhanbad, India

Abstract

Outlier detection is a complex task to perform because of the uncertainty involved in it. Fuzzy logic is more suitable for handling uncertainty. Many applications require real time outlier detection. Neural networks are good at real time operation, online adaption and efficient as they are massively parallel in nature. The hybridization of fuzzy and neural computing system is very promising, since they exactly tackle the situation associated with outliers. In this paper, a Fuzzy min-max neural network is used for outlier detection. In testing phase, a method is proposed for outlier detection which is based on majority voting. User has to define a threshold (t) and if the fuzzy membership value of test pattern in a hyper-box is below t then the pattern will be declared as an outlier with respect to the hyper-box. User should also define a parameter p which decides the percentage of hyper-boxes to be considered for voting a test pattern as an outlier. Experimentation is done on synthetic data and a standard database available on UCI Machine Learning Repository [19]. The proposed method has increased the recognition accuracy whereas the drawback is recall time increased as one more level of voting calculation with a serial time complexity of $O(k)$ is added in the testing phase.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Peer-review under responsibility of scientific committee of International Conference on Advanced Computing Technologies and Applications (ICACTA-2015).

Keywords: Outlier detection; fuzzy logic; neural network; data mining; FNN;FMN; Fuzzy Min-Max Neural Network;

1877-0509 © 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Peer-review under responsibility of scientific committee of International Conference on Advanced Computing Technologies and Applications (ICACTA-2015).
doi:10.1016/j.procs.2015.03.148

1. Introduction

Data mining is defined as “a process of nontrivial extraction of implicit, previously unknown and potentially useful information from the data stored in a database” by Fayyad [1]. In simple words it can be defined as a method of extracting knowledge from huge amount of data or knowledge discovery from databases. However it may mislead the decisions and adversely affect business management if the input data it is not clean and correct. Therefore data cleansing should be done as a pre-processing part for the successful knowledge discovery from databases (KDD). Data cleansing is a method to resolve incomplete, inaccurate, or unreasonable data and then to improve the quality of data by correcting the detected errors and omissions. However the definition of data cleansing varies according to the area in which the process is applied. Data cleansing plays a vital role in data warehousing, KDD, data/information quality management and considered as an integral part of these processes. Data cleansing process can be format checking, limits or completeness checking, subject area specialist may assess data, sensibleness checking, identifying outliers (geographic, statistical, environmental, or temporal) or other errors. Outliers may obstruct the process of data mining due to incorrect knowledge discovery which may badly affect the business profit. An outlier is an observation that deviates so much from other observations as to arouse suspicions that is was generated by a different mechanism (Hawkins, 1980) [2]. Outlier may occur due to data error, improper sensors, anomaly, noise, or typo error. Outlier detection is to find objects in the database which do not obey the laws valid for major part of the data. Outlier detection is an important pre-processing task [3] and has applications in diverse areas [4, 5, 6, 7, 8]. For any outlier detection technique, there are two ways to report the outliers which are scores (i.e. assigning outlier scores) and labels (just assigning label as normal or anomalous) [9].

Outlier detection methods are derived from three fields of computing which are statistics (proximity-based, parametric, non-parametric and semi-parametric), machine learning and neural networks (supervised and unsupervised)[10]. Statistical models are commonly appropriate to quantitative real-valued data sets and are less applicable for least quantitative ordinal data. Machine learning techniques works well on noisy data, do not suffer the Curse of Dimensionality, robust in nature, and have simple class boundaries. Machine Learning has mainly centred on categorical data whereas much outlier detection has only focused on continuous real-valued data attributes. Neural networks (NN’s) represent a computational approach to intelligence as contrasted with the traditional, more symbolic approaches [11]. NN’s can generalize well to unseen patterns and are good at adaptive learning, fault tolerance using redundant information coding, real time operation, managing complex nonlinear input-output relationships, self-organization, nonparametric i.e. they can work without any prior information and are efficient as they are massively parallel in nature. NN has the ability to learn by example which makes them very flexible and powerful. Most of the literature in this area says that neural networks are best at identifying patterns or trends in data. Hence they are mainly used for pattern recognition including, speech recognition, image recognition, prediction or forecasting including, sales forecasting, customer research, target marketing, industrial process control, data validation , risk management, SAR/Sonar data classification to stock market tracking, and so on [12]. Much of the pattern recognition task is now implemented using fuzzy logic as it can handle uncertainty and can be used to reason in imprecise situations [13]. The advantages of fuzzy logic are human type reasoning, can be represented in terms of If–THEN rules which can be used as expert knowledge, good at higher level decision making, etc. However fuzzy systems create their fuzzy rule base using human expert knowledge and hence, lack adaptation. Elicitation of fuzzy rules from experts is usually difficult. As well as they do not have the ability to construct models solely based on the target system's sample data. To overcome this, much of the research work is going on building fuzzy systems with learning and adaptation capabilities. Various methods have been proposed to automatically generate and adjust the fuzzy rules without human experts assistance; the neural fuzzy [14, 15, 16] and genetic fuzzy [17, 18] are two most successful approaches in this regard.

Based on this survey, in this paper, a neural fuzzy system proposed in [16] is used for training. In the testing phase, a method is proposed which is based on the voting done by majority of the hyper-boxes to detect outliers. Experimentation is done on synthetic data and a standard database available on UCI Machine Learning Repository [19]. The paper is organized as follows: section-2describes the survey done on Fuzzy min-max neural network (FMN) and its variants as well as its applications in various domains. Section-3 gives an overview of the FMN learning and recall phase. In section-4, majority voting based method for outlier detection is proposed. Section 5 describes about the synthetic dataset and a standard dataset used for the experimentation on FMN for outlier detection and the results are obtained. Conclusions are given in section-6 and references are cited at the end.

2. Literature survey

Many researchers today, believe that future systems will be of hybrid in nature that combines more than one approach to solve a given type of problem. As outlier itself is a fuzzy concept and required in many real time operations, we believe that the hybridization of fuzzy logic and neural network for outlier detection is a promising area of research. However, a very less work is done in this area. This survey discusses about FMN, the variants of FMN proposed by different researchers and its application in various domains including outlier detection. P. K. Simpson proposed a fuzzy min-max neural network (FMN) a classifier based on supervised learning [16]. FMN learns patterns in terms of fuzzy sets called hyper-boxes having two end points called min point and max point. A pattern class is a union of all the hyper-box fuzzy sets which belongs to the same class. In the training phase, FMN creates hyper-box fuzzy sets and learns the non-linear decision boundary between various pattern classes. In the recall phase FMN uses fuzzy membership to decide class of a test pattern. Jawarkar et al. have successfully used FMN for speaker identification [20]. Panicker et al. have used FMN for real time fault diagnosis [21]. Mohammadi et al. have used FMN which is also called as Fuzzy Hyper Sphere Neural Network (FHSNN), for real time detection of heart diseases [22]. Upasani and Potey have successfully used FMN classifier for face recognition [23]. Bogdan and Andrzej have done fusion of fuzzy min-max clustering and classification algorithms, named it as general fuzzy min max neural network (GFMN) [24]. The algorithm created with this hybridization can be used as pure clustering, pure classification, or hybrid clustering classification. They have also updated the fuzzy membership function of FMN, since the original fuzzy membership function defined in [16] is not properly convex. Nandedkar and Biswas have updated FMN by including compensatory neurons to handle the overlapping and containment situation and also to describe the gradation error in FMN [25]. Liu et al. have used FMN based on cancroids rather than min max points for faster recall [26]. Meneganti et al. have proposed a learning algorithm to remove 2 major drawbacks of the Simpson's model which are difficulty to finalize the threshold value and dependence of the classification performance on the presentation order of the input patterns [27-28]. They have used the modified algorithm for outlier detection and found that its performance is better than other well-known methods. Anas and Lim have modified FMN to maintain a small rule set in order to get high classification performance and used it for fault detection and classification [29].

3. FMN working

The FMN is a supervised learning algorithm proposed in [16] which creates decision boundaries by creating subsets of pattern space. It is an example of hyper-spherical attractor neural network [30]. FMN defines the pattern classes by a union of fuzzy set of hyper-boxes. Hyper-boxes are defined by min-max points which are updated and fine-tuned during learning. Min point of a hyper-box is the closest point to the origin and max point is the farthest point from the origin. Hyper-box has an associated membership function (also called as a discrimination or characteristic function) which is compliment of a sum of average amount of min point and max point violations. The fuzzy membership values range between 0 and 1, with 1 indicating full membership, 0 indicating no membership and values between 0 and 1 indicate partial membership. The FMN learning algorithm determines the min-max points during the single pass through the data, which leads to creation of new classes and refining the existing ones (without retraining). In learning process, pattern is selected from training set and the best match hyper-box for the same class is searched to accommodate the new training pattern. This process is called hyper-box expansion. If the expansion criteria are not satisfied then a new hyper-box is created from that pattern and added to the system. FMN allows overlap between hyper-boxes from same class but overlap between hyper-boxes belonging to different classes need to be removed. Thus, the newly created or expanded hyper-box is checked for the possibility of overlap with the hyper-boxes belonging to the other classes. If it indeed overlapped, then the contraction of hyper-box is done to remove the overlap. In the process of contraction both the overlapping hyper-boxes are minimally contracted just to remove the overlap. This process is repeated for all the patterns in the training set. The membership function for each hyper-box fuzzy set states the degree to which a pattern belongs (or fits) to a respective hyper-box. The fuzzy membership value can be used to give a fuzzy or a soft decision for any pattern, whereas a union operation can be applied on those fuzzy sets to get a crisp or a hard decision. The FMN has the advantages such as it has on-line learning ability, learning is done with a single pass through data, can learn any realistic and multidimensional data generally having great nonlinear decision boundaries and can be executed with a great speed on parallel hardware as it can be implemented using single precision arithmetic operations such as add, subtract and compare. Its recall time per pattern is less and hence can be used in real-time applications.

Classification accuracy is largely dependent on the size of hyper-box (θ). More is the hyper-box size (θ), less number of hyper-boxes is created and the classification accuracy reduces .To get good classification rate, θ should be reduced which will adversely affect the training and recall time. In the training phase, θ should be selected such that the classification accuracy is good as well as training and recall time is optimal.

4. Majority voting based method proposed in the recall phase to detect outliers

Consider n number of training samples is used to train FMN and k number of hyper-boxes is created in training phase. In the testing phase, a method is proposed for outlier detection, which is based on majority voting.

User has to define a threshold (t) and if the fuzzy membership value of test pattern in a hyper-box is below t then the pattern will be declared as an outlier with respect to that hyper-box only. User should also define a factor p which decides the percentage of hyper-boxes to be considered for voting a test pattern as an outlier. If p percentage of hyper-boxes declares any test pattern as an outlier then the pattern will be considered as an outlier. Value of parameters p and t should be selected such that $0 \le p <= 1 and 0 < t \le 1$. The parameters p and t need to be tuned such that all the test patterns from testing set are correctly recognized as outliers. Following algorithm is proposed to detect outliers, which is based on majority voting.

$$\left[\begin{array}{l}
\text{set value for t and } p \\
\text{count} = 0; \\
\textit{for} \quad i = 1 \quad \textit{to} \quad k \\
\quad \text{Step 1. Compute membership (z) of a test pattern in ith hyperbox.} \\
\quad \text{Step 2. If (z < t) then} \\
\qquad\qquad \text{count} = \text{count} + 1; \\
\qquad\quad \text{end} \\
\text{end} \\
\text{if count} \ge \lceil p \times k \rceil \text{, then} \\
\quad \text{declare the test pattern as an outlier.} \\
\text{end}
\end{array}\right]$$

The serial time complexity of this method is $O(k)$.

5. Simulation and results

Experimentation is done on synthetic data and a standard database available on UCI Machine Learning Repository.

5.1 Dataset description

A 2-Dimentional synthetic dataset and a standard UCI Machine Learning Repository dataset are used for the experimentation on FMN for outlier detection. The datasets are described in 5.1.1 and 5.1.2 section.

5.1.1 Synthetic dataset

2-Dimentional Synthetic data is created with nominal patterns and few outlier patterns. The patterns in nominal dataset belong to two different classes, class1 and class2. Nominal patterns are used to train the FMN and outlier patterns are applied in testing phase.

A 2-D Synthetic data is shown in Fig.1.

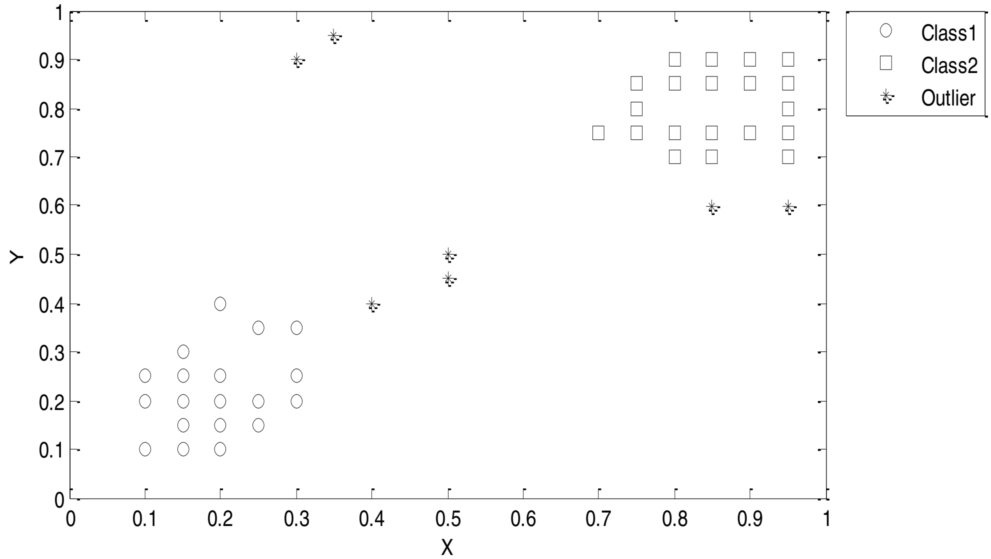

Fig. 1. Synthetic dataset

5.1.2 UCI Machine Learning Repository Dataset of Mechanical Analysis

This dataset [19] describes the Fault Diagnosis problem caused by vibration in electro-mechanical devices (mechanical pumps). Total 209 instances are included in the database. Each instance contains many components and each component is described with 8 features or attributes. First one is a dummy attribute used for numbering and can be ignored. The other 7 important attributes are: component number, support in the machine where measure was taken, frequency of the measure, measure (real), earlier measure (real), dir, rpm of the machine, and class label. Dir feature is described with 3 values viz. filter, type of the measure and direction. It is a 10 class problem out of which 6 are the basic classes. Class 1 depicts problems in the joint, Class 2 describes faulty bearings, Class 3 is mechanical loosening, Class 4 is about basement distortion, Class 5 describes unbalance situation and Class 6 is normal operating conditions. The other 4 classes which are combinations of these 6 basic classes are - Class 7 describes the shaft misalignment (includes class 1 and class 4), Class 8 is about problems in the pump (includes classes 2, 3 and 5), Class 9 depicts problems in the motor (includes classes 2, 3 and 5) and Class 10 is about the problems in the machine (includes all basic classes except Class 6 which is normal operating conditions). Total 9254 patterns are there in the dataset. This dataset is normalized and used to prepare Training and Testing set.

5.2 Experimental results of outlier detection using FMN

Experimentation is done on synthetic data and a standard database and results are obtained by varying the thresholds. The results are discussed in 5.2.1 and 5.2.2 section.

5.2.1 Results by varying thresholds for Synthetic dataset

FMN is trained with the nominal data by varying the size of hyper-box (θ). The size of hyper-box which has given 100% classification rate is finalized, with minimum number of hyper-boxes and optimal training time, the readings are as follows:
Hyper-box size (θ) = 0.05
Training time = 0.0620 seconds
No. of hyper-boxes created (k) = 25
After getting 100% classification, the algorithm proposed in section-4 is used to detect the outliers. Results which are shown in Table 1 are monitored by varying the value of threshold (t) and percentage of hyper boxes voting for outlier (p), such that the algorithm will correctly identify outliers.

Table 1. An Results illustrating the recognition rate, false -^{ve} and recall time for Synthetic dataset by varying threshold (t) and percentage voting (p)

Sr. No.	percentage of hyper-boxes voting for outlier(p)	Threshold (t)	Number of outliers detected	Recall time in seconds	False-^{ve}	%Recognition rate
1	0.6	0.8	0	0.01600	7	0 %
2	0.6	0.85	2	0.01600	5	28.57 %
3	0.6	0.89	5	0.01500	2	71.42 %
4	0.6	0.9	6	0.01500	1	85.71 %
5	0.6	0.91	7	0.01700	0	100 %
6	0.7	0.8	0	0.03100	7	0 %
7	0.7	0.85	2	0.01500	5	28.57 %
8	0.7	0.9	4	0.03200	3	57.14 %
9	0.7	0.91	5	0.01600	2	71.42 %
10	0.7	0.92	7	0.01600	0	100 %
11	0.8	0.8	0	0.03100	7	0 %
12	0.8	0.85	0	0.01600	7	0 %
13	0.8	0.9	4	0.01600	3	57.14 %
14	0.8	0.91	4	0.01600	3	57.14 %
15	0.8	0.92	5	0.01500	2	71.42 %
16	0.8	0.93	7	0.03200	0	100 %

5.2.2 Results by varying thresholds for Standard dataset

FMN is trained with 967 normal patterns and the optimum size of hyper-box is selected yielding 100% classification rate and the readings are:
Hyper-box size (θ) = 0.001
Training time = 3.688000 seconds
No. of hyper-boxes created (k) = 60

Then 8287 patterns are used for testing and the algorithm mentioned in section-4 is used to detect the outliers. Results which are shown in Table 2 are monitored by varying the value of threshold (t) and percentage of hyper boxes voting for outlier (p), such that the algorithm will correctly identify outliers.

Table 2. Results illustrating the recognition rate, false -^{ve} and recall time for Standard dataset by varying threshold (t) and percentage voting (p)

Sr. No.	percentage of hyper-boxes voting for outlier(p)	Threshold (t)	Number of outliers detected	Recall time in seconds	False -^{ve}	%Recognition rate
1	0.6	0.90938	1	36.0620	8286	0.012%
2	0.6	0.91	37	35.7970	8250	0.446 %
3	0.6	0.92	60	36.1250	8227	0.724%
4	0.6	0.925	242	37.1720	8045	2.92 %
5	0.6	0.93	8287	39.5630	0	100 %
6	0.7	0.909	0	35.3750	8287	0 %
7	0.7	0.91	0	34.9530	8287	0 %
8	0.7	0.92	2	35.5780	8285	0.024 %
9	0.7	0.925	10	35.7190	8277	0.12 %
10	0.7	0.93	8287	35.2340	0	100 %
11	0.8	0.909	0	35.6090	8287	0 %
12	0.8	0.91	0	35.3910	8287	0 %
13	0.8	0.92	2	35.7030	8285	0.024 %
14	0.8	0.925	9	35.7030	8278	0.108 %
15	0.8	0.93	4884	36.0620	3403	58.93 %
16	0.8	0.935	4974	36.3280	3313	60.02 %
17	0.8	0.94	5782	36.8440	2505	69.77%
18	0.8	0.95	8213	35.5310	74	99.10 %
19	0.8	0.955	8287	36.2190	0	100 %

It is observed that if the value of p is increased then threshold t should also be increased to get 100 % outlier detection. However if the threshold is increased then there is a possibility that few normal patterns are also recognized as outliers. Therefore the parameters p and t should be tuned such that 100% recognition is achieved without detecting normal patterns as outliers.

6. Conclusion and future directions

The modification proposed here in the testing phase of FMN is efficient and robust as it has successfully recognized all the outliers. Drawback is one more level of voting calculation with a serial time complexity of $O(k)$ is added in the testing phase. Therefore, the recall phase is more effortful compared to FMN and the recall time is increased. With the experimentation done in this paper on synthetic and standard UCI dataset, it is concluded that FMN with this additional method in the recall phase has good ability of classification and can be successfully used for outlier detection. Thus we recommend using FMN for outlier detection.

Limitation of this method is user has to tune the parameters p and t to get good recognition accuracy. The proposed method has increased the recognition accuracy at the cost of recall time. Future work can be done to retain the recognition rate without increasing the recall time.

References

1. Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy , R. G. R.: Advances in Knowledge Discovery and Data Mining. AAAI Press/ The MIT Press, Menlo Park, CA. (1996)
2. Hawkins, D. 1980. Identi¯cation of outliers. Monographs on Applied Probability and Statistics.
3. Silvia Cateni, Valentina Colla and Marco Vannucci, Outlier Detection Methods for Industrial Applications, Advances in Robotics, Automation and Control, Book edited by: Jesús Arámburo and Antonio Ramírez Treviño, ISBN 78-953-7619-16-9, pp. 472, October 2008, I-Tech, Vienna, Austria
4. Al-Zoubi, M. B, A. M. Kamel and M. J. Radhy Techniques for Image Enhancement and Noise Removal in the Spatial Domain, WSEAS Transactions on Computers, Vol. 5, No. 5, 2006, pp. 1047-1052.
5. Yang J. and Z. Zhao, A Fast Geometric Rectification of Remote Sensing Imagery Based on Feature Ground Control Point Database, WSEAS Transactions on Computers, Vol. 8, No. 2, 2009, pp. 195-204.
6. Bolton, R. and D. J. Hand, Statistical Fraud Detection: A Review, Statistical Science, Vol. 17, No. 3, 2002, pp. 235-255.
7. Lane, T. and C. E. Brodley. Temporal Sequence Learning and Data Reduction for Anomaly Detection, 1999. ACM Transactions on Information and System Security, Vol. 2, No. 3, 1999, pp. 295-331.
8. Chiu, A. and A. Fu, Enhancement on Local Outlier Detection. 7th International DatabaseEngineering and Application Symposium (IDEAS03), 2003, pp. 298-307.
9. Barnett, V. and T. Lewis, Outliers in Statistical Data, John Wiley, 1994.
10. Hodge, V. and J. Austin, A Survey of Outlier Detection Methodologies, Artificial Intelligence Review, Vol. 22, 2004, pp. 85–126.
11. R. J. Marks, II, "Intelligence: Computational versus artificial," IEEETrans. Neural Networks, vol. 4, 1993.
12. D. E. Rumelhart and J. L. McClelland, Parallel Distributed Processing:Explorations in the Microstructure of Cognition. Cambridge, MA: MITPress, 1986.
13. Peng Yang, Qingsheng Zhu and Xun Zhong, Subtractive Clustering Based RBF Neural Network Model for Outlier Detection, Journal Of Computers, Vol. 4, No. 8, August 2009, pp. 755-762
14. J.-S. R. Jang, "ANFIS: Adaptive-Network-based Fuzzy Inference Systems," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 23, No. 3, pp. 665-685, May 1993.
15. D. Nauck, R. Kruse, "NEFCLASS - A Neuro-Fuzzy approach for the classification of data," presented at the Symposium on applied Computing, Nashville, USA, 1995.
16. Patrick K. Simpson, Fuzzy Min-Max Neural Networks-Part 1: Classification, IEEE Trans. on Neural Network," Vol. 3, No.5, pp. 776-786, Sept. 1992.
17. H. Ishibuchi, T. Nakashima, T. Murata, "A fuzzy classifier system that generates fuzzy if–then rules for pattern classification problems," Proceedings of second IEEE international conference on evolutionary computation, Perth, Australia, November, pp. 759–64, 1995.
18. J. Liu, J. Kwok, "An extended genetic rule induction algorithm," Proceedings of the Congress on Evolutionary Computation Conference, 2000.
19. http://archive.ics.uci.edu/ml/machine-learning-databases/mechanical-analysis/older-version/mechanical-analysis.names
20. N. P. Jawarkar, R. S. Holambe and T. K. Basu, Use of fuzzy min-max neural network for speaker identification, Proc. IEEE Int. Conference on Recent Trends in Information Technology (ICRTIT 2011), MIT, Anna university, Chennai, Jun.-3-5, 2011.
21. S. S. Panicker, P. S. Dhabe, M. L. Dhore, Fault Diagnosis Using Fuzzy Min-Max Neural Network Classifier, CiiT International Journal of Artificial Intelligent Systems and Machine Learning, Issue Jul. 2010. http://www.ciitresearch.org/aimljuly2010.html.
22. M. Mohammadi, R. V. Pawar, P. S. Dhabe, Heart Diseases Detection Using Fuzzy Hyper Sphere Neural Network Classifier, CiiT International Journal of Artificial Intelligent Systems and Machine Learning, Issue Jul. 2010. http://www.ciitresearch.org/aimljuly2010.html.
23. Nilam Upasani and M. Potey, Generalized Ring averaging: a new method for left and right directional illumination invariant face recognition for frontal poses and their small variants, Int. Conf. and Workshop on Emerging Trends in Technology, Mumbai, ACM-ISBN 978-1-60558-812-4, 2010.
24. Bogdan G. and Andrzej B., General Fuzzy Min-Max Neural Network for Clustering and Classification, IEEE Trans. on Neural Networks, Vol.11, No.3, May 2000.
25. V. Nandedkar and P. K. Biswas, A fuzzy in max neural network classifier with compensatory neuron architecture, Proc. 17th Int. Conference on pattern recognition (ICPR), 2004.
26. Liu J, H. Xin, Yang J., A fuzzy min-max neural network classifier based on centroid, Proc. 30th Chinese Control conference, Yantai, China, pp. 2749-2763, Jul. 22-24, 2011.
27. M. Meneganti, F. Saviello, and R. Tagliaferri, Fuzzy neural networks for classification and detection of anomalies, Proc. 7th IFSA World Congress Prague, M. Mares et al. Eds. Prague, Czechoslovakia: Academia, pp. 561–566, 1997
28. M. Meneganti, F. S. Saviello, and R. Tagliaferri, Fuzzy neural networks for classification and detection of anomalies, IEEE Trans. on Neural Networks, Vol. 9, No.5, Sept. 1998.
29. Anas Quteishat, Chee Peng Lim, "A modified fuzzy min–max neural network with rule extraction and its application to fault detection and classification", Applied Soft Computing 8 (2008) 985–995, www.elsevier.com/locate/asoc
30. R. Kennedy, "The hyperspherical attractor neural network: concepts, theory and application", in proc. WNN-AIND 91, 1991, pp. 241-271.
31. M.Deepa, P. Revathy, Validation of Document Clustering based on Purity and Entropy measures, International Journal of Advanced Research in Computer and Communication Engineering Vol. 1, Issue 3, May 2012
32. E. Cox, Fuzzy Modeling And Genetic Algorithms For Data Mining And Exploration, Elsevier, 2005

33. Bezdek, J, L. Hall, and L. Clarke, Review of MR Image Segmentation Techniques Using Pattern Recognition, Medical Physics, Vol. 20, No. 4, 1993, pp. 1033–1048.
34. Rignot, E, R. Chellappa, and P. Dubois, Unsupervised Segmentation of Polarimetric SAR Data Using the Covariance Matrix, IEEE Trans. Geosci. Remote Sensing, Vol. 30, No. 4, 1992, pp. 697–705.
35. Al- Zoubi, M. B., A. Hudaib and B Al- Shboul A Proposed Fast Fuzzy C-Means Algorithm, WSEAS Transactions on Systems, Vol. 6, No. 6, 2007, pp. 1191-1195.
36. Maragos E. K and D. K. Despotis, The Evaluation of the Efficiency with Data Envelopment Analysis in Case of Missing Values: A fuzzy approach, WSEAS Transactions on Mathematics, Vol. 3, No. 3, 2004, pp. 656-663
37. Chiu, S. L. (1994). Fuzzy model identification based on cluster estimation. Journal of Intelligent and Fuzzy Systems, 2, 267–278.
38. Binu Thomas and Raju G, A Novel Fuzzy Clustering Method for Outlier Detection in Data Mining, International Journal of Recent Trends in Engineering, Vol. 1, No. 2, May 2009
39. Moh'd Belal Al-Zoubi, Ali Al-Dahoud, Abdelfatah A. Yahya, New Outlier Detection Method Based on Fuzzy Clustering, WSEAS transactions on information science and applications, Issue 5, Volume 7, May 2010, pp. 681 – 690.
40. Hawkins, S.; He, X.; Williams, G.J. & Baxter, R.A., Outlier detection using replicator neural networks. Proceedings of the 5th international conference on Knowledge Discovery and Data Warehousing, 2002.
41. D. H. Ackley, G. E. Hinton, and T. J. Sejinowski. A learning algorithm for boltzmann machines. Cognit. Sci., 9:147–169, 1985.
42. R. Hecht-Nielsen. Replicator neural networks for universal optimal source coding. Science, 269(1860-1863), 1995.
43. Williams, G.; Baxter, R.; He, H. & Hawkison,S., A comparative study of RNN for outlier detection in data mining, Proceedings of the IEEE International Conference on Data Mining, pp. 709–712, 9-12 December 2002, Australia.
44. Nag, A.K.; Mitra, A. & Mitra, S., Multiple outlier Detection in Multivariate Data Using Self-Organizing Maps Title, Computational Statistical, N.20, 2005, pp.245-264
45. Varun chandola, arindam banerjee and vipin kumar, outlier detection : a survey
46. Wang G., Jinxing Hao, Jian Ma, Lihua Huang, A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering. Expert Systems with Applications (2010), doi:10.1016/j.eswa.2010.02.102
47. Hawkins d. (1980). Identification of outliers. Chpaman and hall. London.
48. Knorr, E. and R. Ng, Algorithms for Mining Distance-based Outliers in Large Data Sets, Proc.the 24th International Conference on Very Large Databases (VLDB), 1998, pp. 392-403.
49. Knorr, E., R. Ng, and V. Tucakov, Distancebased Outliers: Algorithms and Applications. VLDB Journal, Vol. 8, No. 3-4, 2000, pp. 237- 253.
50. Rousseeuw, P. and A. Leroy, Robust Regression and Outlier Detection, 3rd ed., John Wiley & Sons, 1996.Chandola, V., Banerjee, A. and Kumar, V. (2009) Anomaly detection: A survey, ACM Computing Surveys (CSUR), ACM Digital Library,Vol. 41, Issue 3, Pp.1-58.
51. Ramaswami, S., R. Rastogi and K. Shim,Efficient Algorithm for Mining Outliers from Large Data Sets. Proc. ACM SIGMOD, 2000, pp. 427-438.
52. Angiulli, F. and C. Pizzuti, Outlier Mining in Large High-Dimensional Data Sets, IEEE Transactions on Knowledge and Data Engineering, Vol. 17 No. 2, 2005, pp. 203- 215.
53. Breunig, M., H. Kriegel, R. Ng and J.Sander, 2000. LOF: Identifying Density-Based Local Outliers. In Proceedings of ACM SIGMOD International Conference on Management of Data. ACM Press, 2000, pp. 93–104.
54. Papadimitriou, S., H. Kitawaga, P. Gibbons, and C. Faloutsos, LOCI: Fast Outlier Detection Using the Local Correlation Integral. Proc. Of the International Conference on Data Engineering, 2003, pp. 315-326.
55. Gath, I and A. Geva, Fuzzy Clustering for the Estimation of the Parameters of the Components of Mixtures of Normal Distribution, Pattern Recognition Letters, Vol. 9, 1989, pp. 77-86.
56. Cutsem, B and I. Gath, Detection of Outliers and Robust Estimation using Fuzzy Clustering, Computational Statistics & Data Analyses, Vol. 15, 1993, pp. 47-61.
57. Jiang, M., S. Tseng and C. Su, Twophase Clustering Process for Outlier Detection, Pattern Recognition Letters, Vol.22, 2001, pp. 691- 700.
58. M. I. Petrovskiy, “Outlier Detection Algorithms in Data Mining Systems”, Programming and Computer Software, Vol. 29, No. 4, pp. 228–237, 2003.