Interpretability and Accuracy Analysis of Fuzzy Rule Based System Designed for Abalone

Prabhash Chandra, Devendra Agarwal, Praveen Kumar Shukla

Abstract*: Fuzzy Rule Based Systems are playing vital role in the implementation of human decision making. The development of interpretable Fuzzy Rule Based Systems with improved accuracy is a crucial research aspect in fuzzy based systems. Mamdani type fuzzy rule based systems are used to implement the proposed model. In this manuscript a FRBS is implemented with Guaje Open-Access Java based software. The interpretability and accuracy assessments are recorded on the different experiments with various rule generation methods, like Fuzzy decision tree and Wang Mendel method. The results are found satisfactory and a trade-off is handled between interpretability and accuracy. The major concern of the experimentation is number and type of fuzzy partitions. K-means and Hierarchical Fuzzy Partitions are used in the experiments with three and five number of fuzzy partitions.

***Keywords*: Fuzzy Ruled Based System (FRBS), Fuzzy Logic, Interpretability, Accuracy.**

I. INTRODUCTION

The implementation of human decision process in machines faces the problems in terms of mathematical modeling of linguistic computing which is fundamental aspect of human decision making [1, 2]. The mathematical implementation and processing of linguistic information is done with the help of fuzzy logic. Fuzzy logic provides the mathematical framework to represent and process the uncertainty and imprecision in the real word information of the application domain [3. 4]. Fuzzy techniques are basically the extended version of binary logic and support the concept of multi-valued logic [5]. Fuzzy logic is also used to implement Knowledge Based System or FRBS in which the form of knowledge representation is if then rules [6, 7, 8]. These rules can be generated using several methods; i. e. Prototyping Methods, Wang Mendel Method [9], Fuzzy Decision Tree Methods [10].

Interpretability and accuracy [11, 12] are the prime features of any Fuzzy Rule Based System. Interpretability is concerned with the understanding of the functioning of any fuzzy system. The more interpretable systems are simpler and complex systems are less interpretable. On the other hand, accuracy is considered closeness between the real system and modeled system. Accuracy can be represented by calculating MSE (Mean Squared Error). The accuracy and interpretability have trade-off relations which is called Trade-off between Interpretability and Accuracy [13, 14]. The trade-off relation shows that the interpretable systems are less accurate and complex systems (less interpretable) are more accurate.

In this paper a FKBS has been developed for Abalone data set which is taken from UCI ML repository [15]. The manuscript consists of five sections. The introduction part is in Section I, Section II shows the related work. Section III discusses the proposed systems and experiment and result analysis is done in Section IV. Conclusion and future scope is summarized in Section V.

II. RELATED WORK

Fuzzy systems have strong capability of modeling complex real world applications along with the knowledge representation mechanism in interpretable form. An overview of interpretability and its quantification is well addressed in [16]. In this paper the designing and implementation of fuzzy models is also discussed. To address the issue of interpretability in fuzzy systems a special issue has been published in [17]. Fuzzy Systems are capable to deal with linguistic computation and its contribution towards the approximate reasoning [18, 19, 20]. Fuzzy logic is also applicable in modeling and control [21], machine learning and data mining [22]. The fuzzy information granulation is applied in the sentiment analysis which is also called opinion mining in [23] considering the aspect of interpretability. The objective is to recognize the emotions or the attitude of the people by using the concept of natural language processing. interpretability is addressed in the area of gradient learning, evolutionary learning, nonlinear modeling and control, identity verification in [25,37,38]. An interpretable FRBS is developed using noninvasive method for coronary artery diseases is developed in [26]. A fuzzy control with self structuring and with direct adaptive nature for computing a group of uncertain system with nonlinear nature is proposed in [27]. This concept is also leads to develop the interpretable fuzzy systems. More interpretable and accurate systems designed and implemented towards increasing the transparency of the system in [28]. Heuristic processing is utilized in mapping. A shared fuzzy rule based TSK type fuzzy classifier is developed in [29]. TSK Fuzzy systems are better for the interpretability and accuracy attainment. Similarly an accurate and fast rule creation method has developed based on Mamdani Fuzzy System in [30]. The quantification of interpretability is a difficult task due to its subjective nature. The assessment of interpretability has many indices in the literature.

Revised Manuscript Received on January 15, 2020

***** Correspondence Author

Prabhash Chandra*, Department of Computer Science & Engineering, School of Engineering, Babu Banarasi Das University, Lucknow, India. Email: pathakprabhash2@gmail.com

Dr. Devendra Agarwal, School of Computer Science & Engineering, Babu Banarasi Das University, Lucknow, India. Email: dev_bbd@yahoo.com

Dr. Praveen Kumar Shukla, Department of Information Technology, Babu Banarasi Das Northen India Institute of Technology, Lucknow, India. Email:drpraveenkumarshukla@gmail.com

The interpretability indices are well discussed in [31] for hierarchical fuzzy system. An interpretable fuzzy system for diabetes diagnosis is developed in [32] which were based on ontology decision making.

The interpretability and accuracy aspects of the knowledge base and fuzzy information granules are well discussed in [33]. In the double axis taxonomy for accuracy and interpretability is explained in detail. The importance of hierarchical fuzzy systems in interpretability is addressed in [34]. The interpretable fuzzy systems also have the core role in the designing of cognitive cities as discussed in [35].

III. PROPOSED SYSTEM

This is important to note here that the FRBS can be developed using two ways. In the first method, the FRBS is designed with the help of experts in the particular domain for which it is being designed. In this approach, the membership functions and rules are framed on the basis of the expert knowledge. On the other hand, the FRBS can be generated using the data sets available in the problem domain. Several methods are implemented using different programming languages to generate the Data Base and Rule Base automatically, i.e. Wang Mendel Method, Fuzzy Decision Tree Method, Fast Prototyping method etc. In this section, several other techniques and approaches are discovered which are dealing with the selection of rules in a huge number of fired rules, optimization of Data Base and Rule Base, assessing interpretability parameters. In the proposed system, second approach is used to implement the FRBS.

An FRBS is proposed to design based on the Abalone Data Set an UCI ML Repository [36]. Physical measurements are utilized for the prediction of abalone. The determination of age is done by cutting the shell using cone. This also includes its straining, number of rings inspected by the microscope. Sometimes weather pattern and location are important parameter for age calculation.

The data set attribute information is given below (Table II).

TABLE II: DATA SET INFORMATION

S. No.	Attribute Name	Data Type	Units of Measurement
1	Sex (SE)	Nominal	M, F, I*
2	Length (LEN)	Continuous	millimeter
3	Diameter (DIA)	Continuous	millimeter
4	Height (HEI)	Continuous	millimeter
5	Whole Weight (WW)	Continuous	gms[1]
6	Shucked Weight (SW)	Continuous	gms[1]
7	Viscera Weight (VW)	Continuous	gms[1]
8	Shell Weight (SW)	Continuous	gms[1]
9	Rings (RIN)	Integer	+/-1.5

[1]Grams
*Male, Female & Infant

The description of the data set is as follows:
1. Characteristics of data set: Multivariable
2. Total number of instances: 4177
3. Area of the data set: life
4. Type of attribute: Categorical, Real, Integer
5. Number of attributes: 8
6. Nature of data set: Classification

The diagram of the conceptual model is as follows:

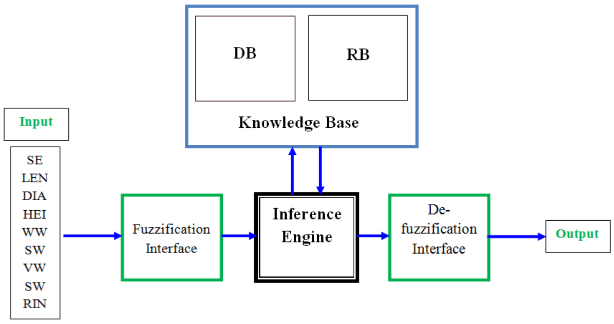

Fig. 1 Conceptual Model

The membership functions of all inputs are generated using default FRBS generation method which includes Wang and Mandel Method Fuzzy Decision Trees.

IV. EXPERIMENTS AND RESULT ANALYSIS

The implementation of the desired system is performed in Guaje a Java Based Open Access Software. For the verification of the system functionality and performance four experiments are done. The sample Membership Functions are shown below (Fig. 2-9)

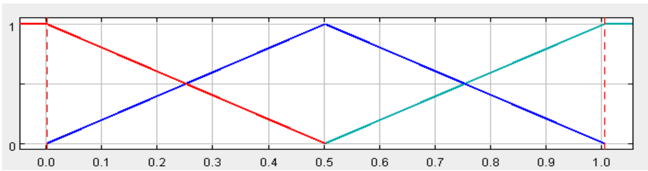

Fig. 2 Membership Function of Shell Weight (Three Partitions)

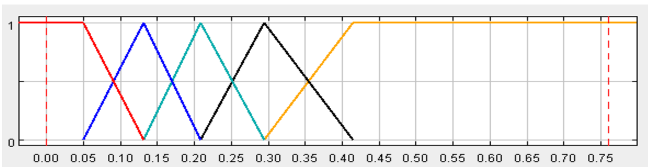

Fig. 3 Membership Function of Viscera Weight (Five Partitions)

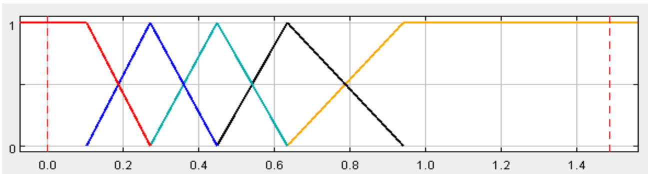

Fig. 4 Membership Function of Shucked Weight (Five Partitions)

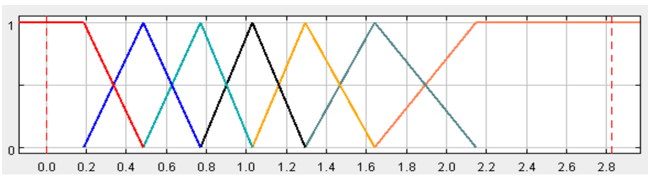

Fig. 5 Membership Function of Whole Weight (Five Partitions)

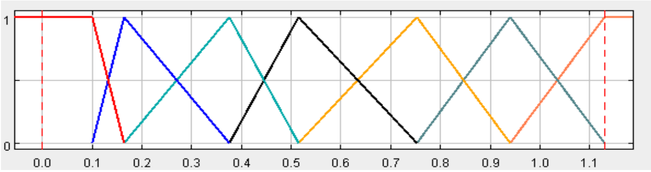

Fig. 6 Membership Function of Height (Five Partitions)

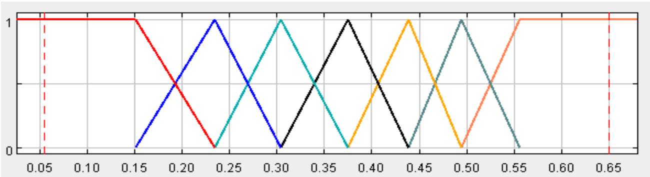

Fig. 7 Membership Function of Diameter (Five Partitions)

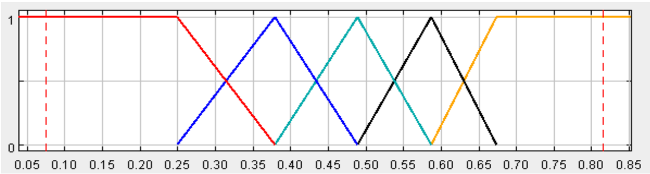

Fig. 8 Membership Function of Length (Five Partitions)

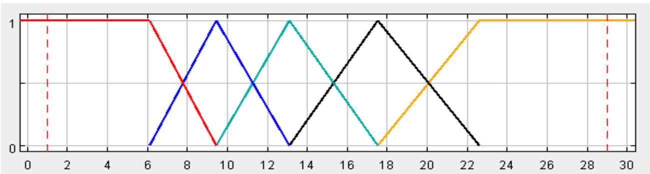

Fig. 9 Membership Function of Ring (Five Partitions)

The experimentation has been done for the performance evaluation of proposed system. Several parameters are identified for the evaluation of accuracy and interpretability parameters.

Accuracy Parameters:

1. **Performance Index (PI):** The PI is calculated using the Eqn. (1).

$$PI = \frac{1}{M}\sqrt{\sum_{i=1}^{M} \| O_i^I - O_i^O \|^2} \ \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Here, M=Number of Active Samples

O_i^I =Inferred Output Value

O_i^O =Observed Output Value

2. **Root Mean Square Error (RMSE):** The calculation of RMSE is given inv Eqn. (2).

$$RMSE = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(O_i^I - O_i^O)^2} \ \ldots\ldots\ldots\ldots\ldots (2)$$

Also, $RMSE = PI\sqrt{M}$

3. **Mean Squared Error (MSE):** The MSE is calculated using the Eqn. (3).

$$MSE = \frac{1}{2M}\sum_{i=1}^{M}(O_i^I - O_i^O)^2 \ \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

4. **Mean Absolute Error (MAE):** The calculation of MAE is given in Eqn. (4),

$$MAE = \frac{1}{M}\sum_{i=1}^{M}(O_i^I - O_i^O) \ \ldots\ldots\ldots\ldots\ldots (4)$$

Interpretability Parameters

1. **Nauck's Index (NI):** NI is calculated as follows,

$$NI = CMX \times PRT \times CRG \ \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Here, $CMX = \frac{\text{Total Number of Classes}}{\text{Total Number of Premises}}$ is called the complexity of the classifier.

PRT= Partition Index which is Average and Normalized on all the inputs

CRG= Coverage Degree which is Average and Normalized, of the Fuzzy Partition

2. **Number of Rules (NOR):** It is total number of rules fired in the Rule Base.
3. **Total Rule Length (TRL):** It is addition of the premises of all the rules.
4. **Average Rule Length (ARL):** It is calculated as follows,

$$ARL = \frac{TRL}{NOR} \ \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

5. **Theoretical Fired Rules (Avg.) (TFR):** It is estimation on the maximum number of fired rules on a data set.

Experiment -1

Type of Partition: K-means
Number of Labels: 05
Rule generation Method: Wang and Mendel
Total Number of Rules Fired: 72

Table III: Accuracy Parameters

PI	0.056
RMSE	3.557
MSE	6.434
MAE	2.862

Table IV: Interpretability Parameters

NI	0
NOR	552
TRL	4416
ARL	8
TFR	113.554

Experiment -2

Type of Partition: K-means
Number of Labels: 09
Rule generation Method: Wang and Mendel
Total Number of Rules Fired: 552

Table V: Accuracy Parameters

PI	0.048
RMSE	3.125
MSE	4.884
MAE	2.492

Table VI: Interpretability Parameters

NI	0
NOR	1404
TRL	11232
ARL	8
TFR	367.426

Experiment -3

Type of Partition: Regular
Number of Labels: 05
Rule generation Method: Wang and Mendel
Total Number of Rules Fired: 432

Table VII: Accuracy Parameters

PI	0.053
RMSE	3.45
MSE	5.95
MAE	2.613

Table VIII: Interpretability Parameters

NI	0
NOR	255
TRL	2040
ARL	8
TFR	55.229

Experiment -4

Type of Partition: K-means
Number of Labels: 05
Rule generation Method: Fuzzy decision Tree
Total Number of Rules Fired: 72
Fuzzy Decision Tree Parameters:
Maximum Tree Depth= 5
Minimum Significant Level= 0.2
Leaf Minimum Cardinality= 10
Tolerance Threshold= 0.1
Coverage Threshold= 0.9
Pruning= YES
Relative Performance Loss= 0.1

Table IX: Accuracy Parameters

PI	0.05
RMSE	3.251
MSE	5.283
MAE	2.37

Table X: Interpretability Parameters

NI	0.008
NOR	33
TRL	198
ARL	6
TFR	13

Table XI: Comparative Results Analysis

Parameters	Experiments			
	1	2	3	4
Accuracy Parameters				
PI	0.056	0.048	0.053	0.05
RMSE	3.557	3.125	3.45	3.251
MSE	6.434	4.884	5.95	5.283
MAE	2.862	2.492	2.613	2.37
Interpretability Parameters				
NI	0	0	0	0.008
NOR	552	1404	255	33
TRL	4416	11232	2040	198
ARL	8	8	8	6
TFR	113.554	367.426	55.229	13

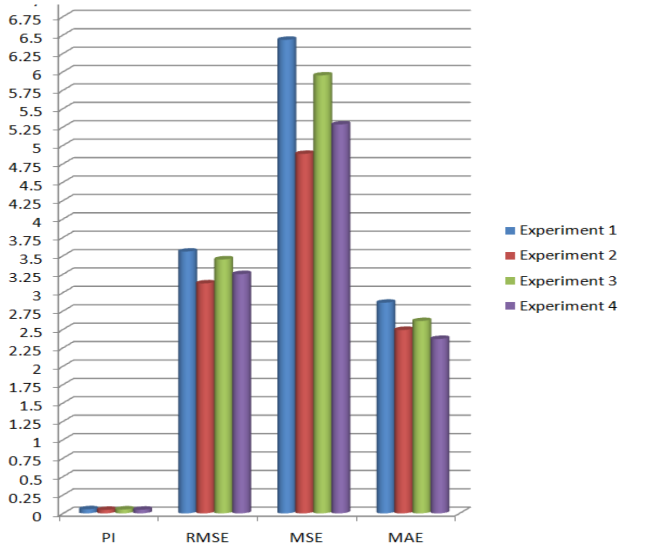

Fig. 10 Comparative Chart of Accuracy Results

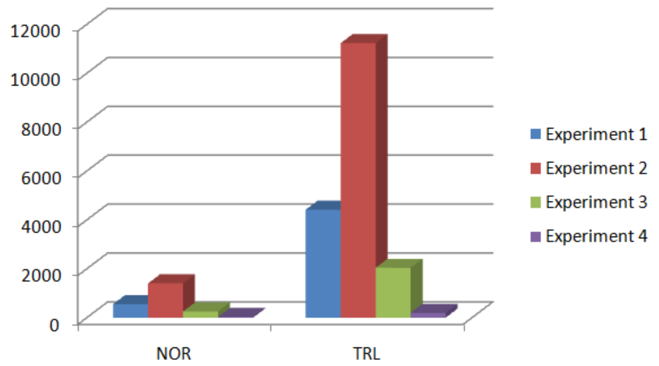

Fig. 11 Comparative Chart of Interpretability Issues

Following are the outcomes of the above result analysis;

1. The increment in number of partitions increases the accuracy which leads to increase in the complexity. However, it deteriorates the interpretability of the system.
2. K-means partitions are better than the regular partitions, for achieving better accuracy.
3. Increase in number of partitions will increase in Total Rule Length and Average Rule length losing the interpretability and improving accuracy.
4. Accuracy improvement enhances the system complexity. It leads to interpretability reduction of the system. Interpretability and accuracy have trade-off relation.
5. Fuzzy Decision Tree produce the Knowledge with higher accuracy compared to Wang Mendel Method but interpretability is better in case of Wang and Mendel method.

V. CONCLUSION AND FUTURE SCOPE

Fuzzy logic based systems are the best performing systems in the imprecise and uncertain environment. This paper also deals with design and implementation of FRBS for predicting the age of Abalone. The data set is extracted from the UCI ML Repository which has six input variables. The proposed system is implemented using Guaje Java Based Open Access Software. In the experimentation, three and five number of fuzzy partitions are used along with rule generation methods Wang Mendel and Fuzzy Decision Tree. After the comparison of results obtained from several experiments, several conclusion are drawn in which most valuable are outlined below,

1. Accuracy improvement is based on the increment in number of fuzzy partitions.
2. Fuzzy Decision Tree gives more accurate results compared to the Wang Mendel Method.
3. K-means partitions are better than the

The development of the FRBS using Type-2 and Interval type-2 fuzzy logic are the future interest of the authors.

REFERENCES

1. J. K. Klir, B. Yuan, (1995). Fuzzy Sets and Fuzzy Logic, *Prentice Hall India*.
2. L. Kuncheva (2000). Fuzzy classifier design, *Berlin: Germany: Springer-Verlag*.
3. F. O. Karray, C. D. Silva (2009). Soft Computing and Intelligent System Design: Theory, Tools and Applications, *Pearson*.
4. T. J. Ross (1995). Fuzzy Logic with Engineering Applications, *McGraw-Hill*.
5. P. K. Shukla, S. P. Tripathi (2014). A new approach for tuning interval type-2 fuzzy knowledge bases using genetic algorithms, *Journal of Uncertainty Analysis and Applications*, Vol. 2(4), pp. 1-15.
6. P. K. Shukla, S. P. Tripathi (2012) "On the design of interpretable evolutionary fuzzy systems (I-EFS) with improved accuracy", International Conference on Computing Sciences, pp. 11-14.
7. P. K. Shukla, S. P. Tripathi (2012). A review on the interpretability-accuracy trade-off in evolutionary multi-objective fuzzy systems (EMOFS), *Information*, Vol. 3, pp. 256-277.
8. P. K. Shukla, S. P. Tripathi (2013). Interpretability issues in evolutionary multi-objective fuzzy knowledge base systems, 7th International Conference on Bio-Inspired Computing: Theories and Applications (BICTA-2012).
9. A. Altay, O. Cinar (2016). Fuzzy Decision Tree, *Fuzzy Statistical Decision Making*, pp. 221-261.
10. J. Gou, F. Hou, W. Chen, C. Wang, W. Luo (2015). Improving Wang Mendel Method performance in fuzzy rules generation using the fuzzy c-means clustering algorithm, *Neuro Computing*, Vol. 151(3), pp. 1293-1204.
11. J. M. Alonso, O. Cordon, S. Guillaume, and L. Magdalena (2007). "Highly interpretable linguistic knowledge bases optimization: genetic tuning versus Solis-Wetts Looking for a good interpretability-accuracy trade-off", IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2007), London, July 23-26, pp. 1-6.
12. J. M. Alonso, L. Magdalena, G. G. Rodriguez (2009). Looking for a good fuzzy system interpretability index: an experimental approach, *International Journal of Approximate Reasoning*, vol. 51(1), pp. 115-134.
13. A. G. Di Nuovo, V. Catania (2009). Linguistic modifiers to improve the accuracy-interpretability trade-off in multi-objective genetic design of fuzzy rule based classifier systems. In Proceedings of 2009 9th International Conference on Intelligent Systems Design and Applications, Pisa, Italy, pp. 128–133.
14. P. K. Shukla, S. P. Tripathi (2014). Handling high dimensionality and interpretability-accuracy trade-off issues evolutionary multi-objective fuzzy classifiers, *International Journal of Scientific and Engineering Research*, Vol. 5(6), pp. 1-6.
15. P. K. Shukla, S. P. Tripathi (2011). A Survey on interpretability-accuracy (I-A) trade-off in evolutionary fuzzy systems, 5th International Conference on Genetic and Evolutionary Computing, pp. 97-101.
16. J. M. Alonso, C. Castillo, C. Mencar (2015). Interpretability of fuzzy systems: current research trends and prospects, Springer Handbook of Computational Intelligence, pp. 219-237.
17. J. M. Alonso, L. Magdalena, Editorial (2011). Special Issue on Interpretable Fuzzy Systems, *Information Sciences*, Vol. 181(20), pp. 4331-4339.
18. L. A. Zadeh (1975). The concept of linguistic variable and its application to approximate reasoning: Part I, *Information Sciences*, Vol. 8, pp. 199-249.
19. L. A. Zadeh (1975). The concept of linguistic variable and its application to approximate reasoning: Part II, *Information Sciences*, Vol. 8, pp. 199-249.
20. L. A. Zadeh (1975). The concept of linguistic variable and its application to approximate reasoning: Part III, *Information Sciences*, Vol. 9, pp. 43-80.
21. T. Takagi, M. Sugeno (1985). Fuzzy identification of systems and its applications to modeling and control, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 15, pp. 116-132.
22. E. Hullermeier (2005). Fuzzy methods in machine learning and data mining: study and prospects, *Fuzzy Sets and Systems*, Vol. 3, pp. 387-406.
23. H. Liu, M. Cocea (2017). Fuzzy information granulation towards interpretable sentiment analysis, *Granular Computing*, Vol. 2, No. 4, pp. 289-302.
24. G. Peters, R. Weber (2016). A framework for dynamic granular clustering, *Granular Computing*, Vol. 1, pp. 1-11.
25. K. Cpalka (2017). Design of interpretable fuzzy systems, *Studies of Computational Intelligence (SCI)*, vol. 684.
26. H. R. Marateb, S. Goudarzi (2015). A noninvasive method for coronary artery diseases diagnosis using a clinically-interpretable fuzzy rule-based system, *Journal of Research in Medical Sciences*, Vol. 20(3), pp. 214-223.
27. N. Wang, J. –C. Sun, Y. –C, Liu (2016). Direct adaptive Self-structuring fuzzy control with interpretable fuzzy rules for a class of nonlinear uncertain systems, *Neuro Computing*, Vol. 173(3), pp. 1640-1645.
28. T. Chen, C. Shang, P. Su, Q. Shen (2018). Induction of accurate and interpretable fuzzy rules from preliminary crisp representation, *Knowledge Based System*, Vol. 146, pp. 152-166.
29. Y. Zhang, H. Ishibuchi, S. Wang (2017). Deep Takegi-Sugeno-Kang Fuzzy classifier with shared linguistic fuzzy rules, *IEEE Transactions on Fuzzy Systems*, Vol. 26(3), pp. 1535-1549.
30. L.–C. Dutu, G. Mauris, P. Bolon (2017). A fast and accurate rule-base generation method for Mamdani fuzzy systems, *IEEE Transactions on Fuzzy Systems*, Vol. 26(2), pp. 715-733.
31. T. Razak, J. M. Garibaldi, C. Wanger, A. Pourabdollah (2017). Interpretability indices for hierarchical fuzzy systems", 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2017, Neplas, Italy.
32. S. E.–Sappagh, J. M. Alonso, F. Ali, J. H. –Jang (2018). An ontology based interpretable fuzzy decision support for diabetes diagnosis, *IEEE Access*, Vol. 6, pp. 37371-37395.
33. H. M. Ahmed, N. A. H. Isa (2017). Knowledge base to fuzzy information granule: a review from the interpretability-accuracy perspective, *Applied Soft Computing*, Vol. 54, pp. 121-140.
34. L. Magdalena (2018). Do hierarchical fuzzy systems really improve interpretability? *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations*, pp 16-26.
35. J. M. Alonso, C. Castillo, C. Mencar (2018). The role of interpretable fuzzy systems in designing cognitive cities, Studies in Systems, Decision and Control, Vol. 176, pp. 131-152.
36. URL: https://archive.ics.uci.edu/ml/index.php accessed on 23.08.2019
37. P. K. Shukla, S. P. Tripathi (2015). Interpretability and Accuracy Issues in Evolutionary Multi-objective Fuzzy Classifiers, *International Journal of Networking and Soft Computing*, Vol. 1(1), pp. 55-69.
38. P. Chandra, D. Agarwal, P. K. Shukla (2017). MOBI-CLASS: A Fuzzy Knowledge Based System for Mobile Handset Classification, 7th International Conference on Soft Computing for Problem Solving (SocPros-2017), IIT, Bhubaneswar.

AUTHORS PROFILE

Prabhash Chandra is a Research Scholar at the Department of Computer Science and Engineering at Babu Banarasi Das University, Lucknow, India. He is working in the area of Fuzzy Systems. He is also serving the Department of Computer Applications in the capacity of Associate Professor and Head of Department at Babu Banarasi Das University, Lucknow, India. He has 18 Years of academic experience.

Devendra Agarwal is currently working as Prof. & Head at Department of Computer Science, School of Engineering, BBD University, Lucknow. He has over 20 years of teaching & 5 years of industrial experience. He has over 20 research papers with 7 students pursuing Ph.D. under his guidance and one awarded Ph.D. to his credit. His area of research includes e-Commerce, Software Engineering, Fuzzy Logic and Data Mining. He has developed and implemented several software projects for Defense, Govt. Organizations, and Private Organizations and in last 17 years in academics he has developed various software's for Accounts, Payroll, Time Table, Library etc.

Dr. Praveen Kumar Shukla is presently working as a Professor and Head in the department of Information Technology, Babu Banarasi Das Northern India Institute of Technology. He is Ph. D. in Computer Science & Engineering from Dr. A P J Abdul Kalam Technical University, Lucknow. He is B. Tech in Information Technology and M. Tech. in Computer Science & Engineering. His research area includes Fuzzy Systems (Interval Type-2 Fuzzy Systems and Type-2 Fuzzy Systems), Evolutionary Algorithms (Genetic Algorithms), Genetic Fuzzy Systems, Multi-Objective Optimization using Evolutionary Algorithm. He is co-supervising Four Ph. D. Thesis at BBD University and Amity University and supervised 5 M. Tech. Projects at Department of CSE, BBD University. He has published 8 Papers in National Conferences, 12 Papers in International Conferences and 13 Papers in International Journals. He has also published a book on Cyber Security.