

Interpretability Assessment in Fuzzy Rule Based Systems

Arun Kumar Shukla, Akhilesh Yadav

Abstract— Rule based systems are basically knowledge base systems. They imitate the functionality of human decision making process in machines. The knowledge is stored in the Knowledge Base of the fuzzy rule based systems. The knowledge is expressed, manipulated and processed by using fuzzy logic. Fuzzy logic represents the human knowledge very well because its mathematical framework is very strong to deal with uncertainty and vagueness inherent with human knowledge. Interpretability is the subjective feature of the fuzzy rule based system that quantifies the understandability of the system functioning by a user. Due to its subjective nature, it is not easy to quantify the interpretability. No global index has been developed till now to deal with interpretability assessment. In this paper the authors have proposed a new interpretability assessment index. The experimental activities are carried out by using open access software “Guaje”.

Index Terms— Fuzzy Rule Based Systems, Fuzzy Logic, Interpretability, Knowledge Representation, Rule Base, Knowledge Base, Inference Engine.

1 INTRODUCTION

To integrate the advantages of the human decision making process into machine, the knowledge based intelligent system have been developed. The representation of knowledge in these systems is a complex task due to its subjective, imprecise and uncertain nature. Bivalent logic is ineffective to deal with these issues in representation knowledge. Fuzzy logic can be used as a knowledge representation theory with a strong mathematical background. The knowledge based systems developed using Fuzzy Logic are called Fuzzy Rule Based System.

Interpretability and accuracy are the important features of the fuzzy rule based system. Interpretability means the quantification of understandability of a fuzzy system by a user. This is a subjective feature and its assessment is not an easy task. On the other hand accuracy is the quantification of closeness between real system and modeled system.

Interpretability and accuracy are the contradictory features. This means the improvement in one can be done at the cost of other. This leads to a trade-off situation and is called “Interpretability-Accuracy Trade-Off”. Evolutionary Multiobjective Optimization is the best technique to deal with this trade-off.

In this paper, the assessment parameters of the interpretability of fuzzy system are identified and a system using a data set has been implementing with the help of “Guaje” open access software. A new interpretability index has been proposed.

2 INTERPRETABILITY OF FUZZY SYSTEMS

Several works has been carried out to demonstrate the qualification of interpretability. In [1] following criterias are identified to access interpretability:

- Arun Kumar Shukla is currently pursuing masters degree program in computer science & engineering in UP Technical University, Lucknow, India, PH-9793284778. E-mail: aruncoru@hotmail.com
- Mr. Akhilesh Yadav is currently Assistant Professor in Department of Computer Science and Engineering, Kanpur Institute of Technology, Kanpur PH-9452154313-. E-mail: akhillakh78@gmail.com

Fuzzy partitions; linguistic interpretability of fuzzy set, number of fuzzy set for each variable.

1. Number of fuzzy if then rules.
2. Number of Inputs.
3. Number of rules conditions.
4. Defuzzification and inference mechanisms

In [2] the interpretability of any fuzzy system is of two types;

1. Complexity Based Interpretability
2. Semantics Based Interpretability

Complexity Based Interpretability focuses on the parameters numbers of rules, labels, variable per rule etc.

A taxonomy has been proposed for interpretability assessment parameter in [3]. There are three classification of parameters; Knowledge Based Interpretability (KBI), Inference Engine Interpretability (IEI) and user knowledge Base Interpretability (UKBI).

Several other areas are found where the fuzzy system are applicable, like database [4,7], software engineering [5], control systems [6], etc.

During the design of fuzzy rule based systems, high dimensionality of data is a big problem. It creates the explosion in the number of rules generated and the interpretability and performance has been deteriorated due to this problem. A fuzzy clustering based solution has been presented in [8] by Shukla et al.

The Interpretability and accuracy parameters are studied in [9].

A review on the interpretability issues and interpretability-accuracy trade off has been carried out in [10,11]. Evolutionary algorithms are playing vital role in the designing & FRBS as on optimization problem. This is also reviewed in [11].

Interval type-2 fuzzy systems are much more capable to deal with imprecision and uncertainty. A DB tuning approach has been proposed and implemented “Lateral Displacement/Expansive Compressive (LDEC)’ using genetic algorithm in [12].

3 Proposed Interpretability Index

Several parameters are used to quantify the interpretability, i.e. number of rules, number of membership function, number of variable called gromelaity etc. A new interpretability index has been proposed here and implemented experimentally.

$$IAI = \frac{NOR \times NMF}{TRL \times ARL}$$

Here,

IAI=Interpretability Assessment Index

NOR=Number of Rules

NMF=Number of Membership Functions

TRL=Total Rule Length

ARL=Average Rule Length

The proposed index covers the database interpretability parameters.

4 Experiments and result analysis

A fuzzy rule based system has been practically implemented using the open access software 'Guaje'. This is java based framework to design and implement interpretable and accurate fuzzy rule based systems. This is basically the implementation of fuzzy modeling methodology 'Highly Interpretable Linguistic Knowledge (HILK)'.

4.1 Guaje Framework

GUAJE means 'Generating Understandable and Accurate fuzzy models in Java Environment'. This is basically the implementation of fuzzy modeling approach 'HILK (Highly Interpretable Linguistic Knowledge)'. The objective of HILK is to gain a good interpretability-accuracy trade-off that is a common framework for integrating expert and induced knowledge.

Different preexisting open source tools are combined to develop a computational environment for generating interpretable and accurate fuzzy systems. This software is basically an extended version of open access software tool KBCT (Knowledge Base Configuration Tool).

GUAJE combines the following six preexisting tools:

1. Knowledge Base Configuration Tool (KBCT)

KBCT is open source software tool for knowledge extraction and representation that integrates the expert and induced knowledge that is automatically extracted from data. During the integration of two kinds of knowledge, consistency analysis, simplification and optimization operations are carried out.

2. FisPro

FisPro is an open source software tool for developing fuzzy inference systems (FIS) used for reasoning tasks specifically in physical or biological systems. Many algorithms are incorporated to generate fuzzy partitions and rules directly from experimental data. FIS visualization methods are presented in a Java based user friendly GUI.

3. Xfuzzy

Xfuzzy is a free software environment for developing FIS. Different steps of designing the FIS are integrated in the method. This software is encoded in Java and its tools are based on a common specification language XFL3.

4. Ontology Rule Editor (ORE)

ORE is an open source software environment written in Java. This is used to define, manage and test inference rules on a model represented by specific ontology.

5. Weka

Weka is an open source tool with different types of algorithms for data mining purposes. Many classical algorithms J48 that correspond C4.5 algorithm are implemented.

6. Matlab Fuzzy Toolbox

This most commonly used commercial tool for developing fuzzy system application. It is fully integrated with functionalities provided by Matlab environment commonly applicable in engineering, educational and business applications.

4.2 Experimentation

In this experiment, the fuzzy rule based system is generated by using a data set "Liver Disorder" [13]. This is a multivariate data set. Total number of instances is 345 and total number of attributes is 7. This data set is created by BUPA Medical Research Ltd. First five variables are recorded by the blood test and the disorder may be found by excessive alcohol consumption. The attribute information is as follows:

1. mcv - mean corpuscular volume
2. alkphos-alkaline phosphates
3. sgpt- alanine aminotransferase
4. sgot- aspartate aminotransferase
5. gammagt gamma- glutamyl transpeptidase
6. Number of drinks sector field to split data into two sets.

The fuzzy membership functions are as follows:

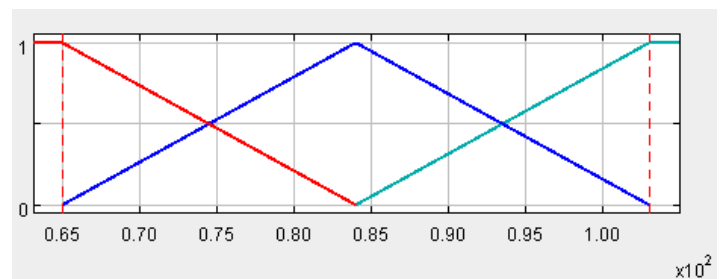


Fig. 1 MF of First Variable 1

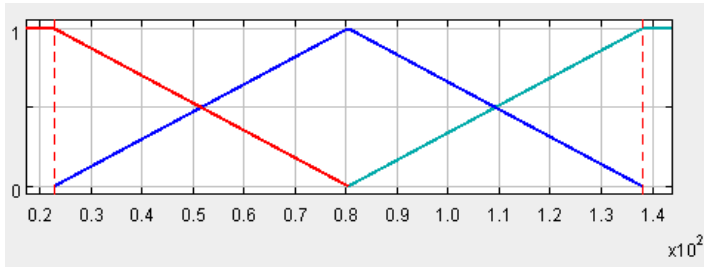


Fig. 2 MF of Variable 2

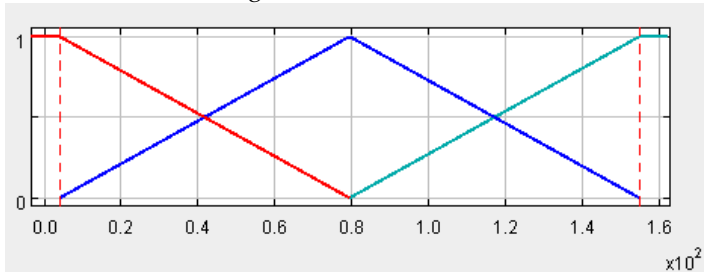


Fig. 3 MF of Variable 3

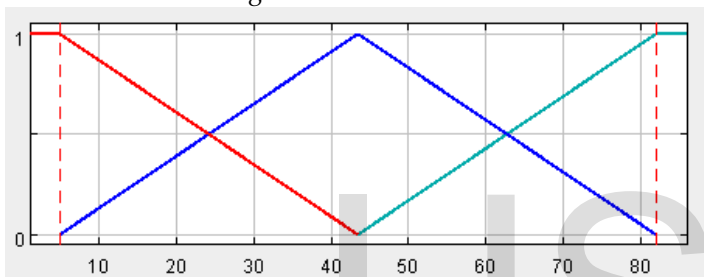


Fig. 4 MF of Variable 4



Fig. 5 MF of Variable 5

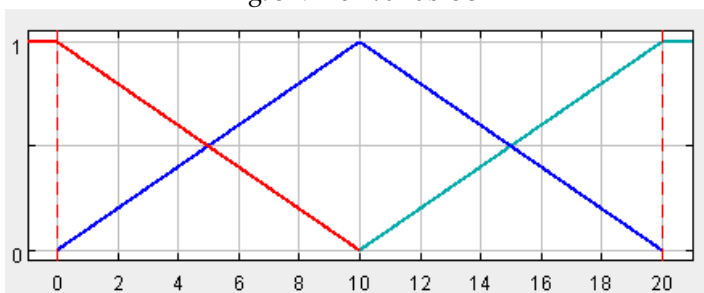


Fig. 6 MF of Variable 6

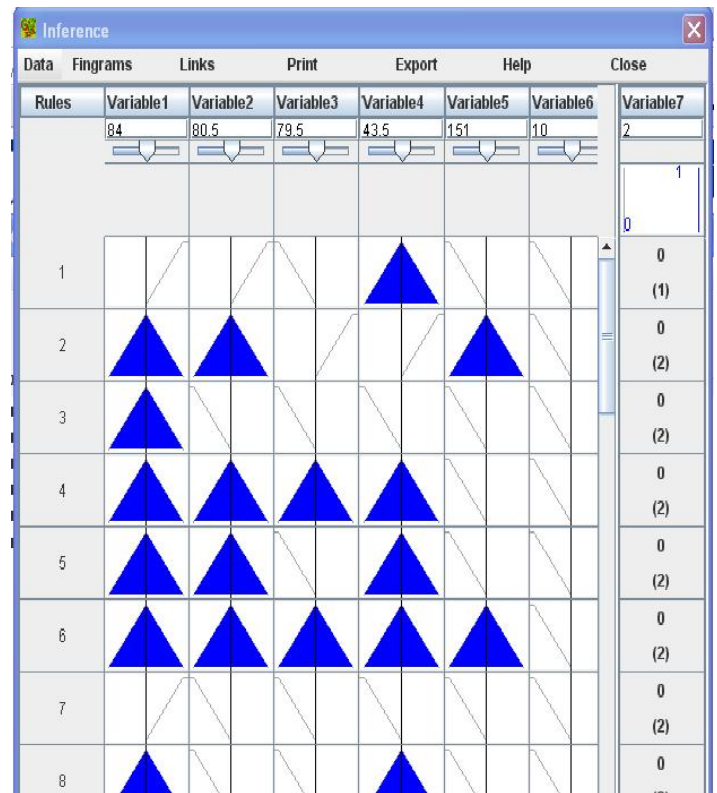


Fig. 7 Snapshot of inferencing process by Guaje open access software

The interpretability parameters are recorded as follows given below with this experiment.

TABLE I
 INTERPRETABILITY INDEX

S. No.	Interpretability Index	Value
1	Nauck Index	0.003
2	Number of Rules	48
3	Total Rule Length	288
4	Average Rule Length	6
5	Theoretical Fired Rules	15.25
6	Inferential Fired Rules	16.771

The recorded accuracy is accuracy is 91.3 %.
 The value of IAI is 0.1319.

4.3 Wang and Mendel Method

The Wang and Mendel method has been used for the rule base generation. The method is detailed as follows;
 Wang and Mendel's method is a simple and accurate data driven RB generation method proposed by Wang and Mendel (1992). Its functioning is based on the input-output data set $D = \{d_1, d_2, \dots, d_n\}$, $d_i = (x_1^i, x_2^i, \dots, x_n^i, y^i)$ that represents the behavior of the problem being solved. In this previous DB definition is considered that consists of input and output fuzzy partitions.

The inferencing process snapshot of the experimentation in fig. 7.

Following steps are considered for RB generation in this method.

1. Fuzzy partition of the input variable spaces

Expert information or normalization approaches are used to obtain fuzzy partition. The input variable spaces may be partitioned with equal (uniform/symmetric fuzzy partition) or unequal number. The MF are identified and assigned for each subspace.

2. Candidate rule set generation

A rule set is required to be generated that covers all the spaces of the input/output data pairs contained in D . Thus n candidate rules are generated. For rule generation, a $(n+1)$ dimensional real array that have n input and 1 output and setting each variable to linguistic label that covers every array component are used.

3. Assignment of importance degree to each rule

Consider the following rule,

$R_i = \text{if } x_1 \text{ is } A_1 \text{ and } \dots \text{ and } x_n \text{ is } A_n \text{ then } y \text{ is } B$

The above rule is generated from the example $d_i, i = 1, 2, \dots, n$. The importance degree is calculated as follows:

$$ID(R_i) = \mu_{A_1}(x_1^i) \dots \mu_{A_n}(x_n^i) \cdot \mu_B(y^i)$$

4. Development of final RB from candidate linguistic rule sets

The n candidate rules are first clustered into p different groups. In this each group is composed of all the candidate rules having same antecedent. It is noted that R_{ij} is the j^{th} rule of i^{th} group. The final RB is generated by selecting the candidate rules with highest importance degree in each group $j, j=1, \dots, p$.

5 Conclusion and future scope

Interpreability assessment of a fuzzy system is a very complex task due to its subjective nature. This is still an open problem and no global index has been proposed to assess it. In this paper the authors have presented a new interpretability index. A fuzzy rule based system has been implemented using "Guaje" open access software for the data set "Liver Disorder". The interpreability index IAI and other interpretability parameters are experimentally approximated.

In future, this work can be extended with the following improvements proposed below in points,

1. Use of improved versions of the fuzzy logic, like type-2 fuzzy logic. Type-2 fuzzy logic is computationally very expensive so its improved version 'Interval Type-2 Fuzzy Logic' is used which carries the properties of type-2 fuzzy logic with less computational cost.
2. High dimensional data sets are leading to explosion in number of rules generated. This situation deteriorates the interpretability and performance of the system. This is an open problem.
3. Interpretability and accuracy are the contradictory fea-

tures and leads to situation of 'interpretability-accuracy trade-off'. Evolutionary Multiobjective Optimization is the approach to deal with this issue. A big number of research lines are identified at the crossroad of evolutionary multiobjective optimization and fuzzy systems.

References

- [1] H. Ishibuchi, T. Yamamoto, Fuzzy rule selection by multi objective genetic local search algorithm and rule evaluation measure in data mining, Fuzzy Sets and Systems, Vol. 4, No. , pp. 59-80, Jan 2004.
- [2] M. J. Gacto, R. Alcalá, F. Herrera, Interpretability of linguistic fuzzy rule based systems; an overview of interpretability measure, Information Sciences ,doi; 10.1016/j.ins.2011.02
- [3] P.K. Shukla, S.P.Tripathi, Interpretability issues in Evaluating Multi Objective Fuzzy Knowledge Based System, BIC-TA 2012, Advance in Intelligent System and computing; Sprigen 2013
- [4] P.K. Shukla, M Darbari, V.K. Singh, S.P. Tripathi, A Survey on Fuzzy Techniques in object oriented database, International Journal of Scientific and Engineering Research, Vol. N011,2011.
- [5] S. Singh, P.K.Shukla, R.Ranjan, A. Kumar, Developing a genetic fuzzy system model for cost- benefit analysis, International Journal & Computers and Tehnology, Vol.13, no.4,2014.
- [6] Poonam, S.p. Tripathi , P. K. Shukla, Uncertainty Handling using fuzzy logic in rule based system, International Journal of Applied Sunce and Technology, Vol. 45, n0. 3,PP. 31-45.
- [7] L. Sahoo, P.K.Shukla Fuzzy Technique in object based modeling, International Journal on Information Sciences and Computing, Vol. 2, No. 1, July 2008.
- [8] P. K. Shukla, S. P. Tripathi, Handling high dimensionality and interpretability accuracy trade-off issues in evolutionary multi objective fuzzy classifiers, International journal & Scientific and Engineering research, Vol 5, No. 6, June-2014.
- [9] P.K. Shukla, S.P. Tripathi, On the design of interpretable evolutionary fuzzy system (I-EFS) with improved accuracy, 2012. International conference of Computing Science, 2012, pp-11-14
- [10] P.K. Shukla, S.P. Tripathi, A survey on interpretability Accuracy Trade-off in evolutionary fuzzy systems. IEEE International Conference on cognitive and Evolutionary computation (ICGEC-2011), Japan, 29 August -01 September ,2011 pp.97-101.
- [11] P.K. Shukla, S.P. Tripathi, A review on the interpretability-accuracy-trade-off in evolutionary multi objective fuzzy system (EMOFS), Information, vol 3,pp 256-277,2012
- [12] P. K. Shukla, S.P.Tripathi , A new approach for turring interval type 2 fuzzy knowledge base using genetic algorithm, Journal of Uncertainty analysis and Applications, Vol. 2, No. 4,pp 1-15.
- [13] In URL "http://archive.ics.uci.edu/ml/datasets/liver+Disorders"
- [14] P. K. Shukla, S. P. Tripathi, Handling high dimensionality and interpretability-accuracy trade-off issues in evolutionary multiobjective fuzzy classifiers, International Journal of Scientific and Engineeringg Research, Vol. 5, No. 6, pp. 2014.
- [15] J. M. Alonso, L. Magdalena, S. Guillaume, HILK: A new methodology for designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism, Intrnational Journal of Intelligent Systems, Vol. 23, No. 7, pp. 761-794, 2008.