A Genetic-Fuzzy Algorithm to Discover Fuzzy Classification Rules for Mixed Attributes Datasets

Dr. Saroj
Associate Professor
Deptt. of CSE, GJUS&T, Hisar, Haryana, India

Nishant Prabhat
M.Tech (CSE)
Deptt. of CSE, GJUS&T, Hisar, Haryana, India

ABSTRACT
Genetic Algorithms, being global search method, have been extensively applied for discovery of automated classification rules. Fuzzy Logic was integrated with genetic algorithms for discovery of Fuzzy Classification Rules (FCRs) which are more interpretable and cope better with pervasive uncertainty and vagueness in real world decision making situations. At one hand, most of the Genetic Algorithm approaches have been implemented for datasets with categorical attributes only, at the other Genetic-Fuzzy approaches have the limitation to deal only with continuous attributes. This paper proposes genetic-fuzzy approach for discovery of fuzzy decision rules from datasets containing both categorical as well as continuous attributes. The continuous attributes are normalized and fuzzified in pre-processing step. A novel match procedure is devised to take care of mixed attributes during the fitness computations of individual rules. A direct match is carried out for categorical attributes whereas a Mumdani style min-max method is employed for matching continuous attributes with the instances in the training dataset. The proposed approach is tested on various datasets containing purely continuous or purely categorical or a mix of both types of attributes. Appropriate encoding scheme, fitness function and genetic operators with the necessary constrained are designed. The results are compared with three other machine learning techniques and results are comparable in terms of predictive accuracy. Moreover, the rule sets discovered with the suggested approach are compact and more comprehensible.

Keywords
Data Mining, Genetic Algorithms, Fuzzy Classification Rules.

1. INTRODUCTION
Data Mining is an essential process where intelligent methods are applied in order to extract patterns that are helpful in decision making [4], [6]. The “Production Rules (PRs: *IF-THEN* rules)” are one of the most popular way of representing Classification Rules in data mining. In high level symbolic rule representation, a rule can be written in the following two forms:

Conjunctive Normal Form(CNF):- *If* $(A_1=x \lor A_1=y) \land (A_2 = a) \land (A_3=p \lor A_3=q)$ *Then* C_1

Disjunctive Normal Form (DNF):– *If* $(A_1=x \land A_2=b \land A_3=p) \lor (A_1=y \land A_2=a \land A_3=q)$ *Then* C_2

Where A_1, A_2, A_3 are attributes and x, y, a, b, p, q are attribute values.

The disadvantage of the rule based classification systems is that they involve sharp cutoffs for distribution and hence are unable to deal with uncertainty and vagueness imperative to decision making. Fuzzy logic is a precise logic of imprecision and approximate reasoning. More specifically, fuzzy logic has capability to reason in an environment of imprecision, uncertainty, incompleteness of information, conflicting information and partiality of truth—in short, in an environment of imperfect information [22]. Hence, Fuzzy Logic has been integrated with the classification task to discover Fuzzy Classification Rules (FCRs) [17]. FCRs are considered more comprehensible as these are based on linguistic variables instead of numeric ranges and this is precisely the reason that systems based on fuzzy if-then rules have been so extensively and successfully used in control problems [11].

1.1 Fuzzy Classification Rules (FCR)
A FCR is represented as below:

IF x_1 is A_1 AND x_2 is A_2……..AND x_n is A_n THEN y is B

where n denotes the number of attributes present in dataset, $x_1, x_2, \ldots x_n$ attributes in the antecedent part of rule, y is the target class attribute and $A_1, A_2, \ldots, A_n$ and B are associated fuzzy sets on the unit interval [0, 1] for the antecedent and consequent parts respectively [20], [22]. FCR are more natural in representation and easy to interpret. For example the fuzzy condition “salary=low” and “age=old” seems more natural than crisp condition “salary<=145.6 $” and “age>=59” [17].

1.2 Genetic Algorithms for Classification
GA uses the principle of selection and evolution to produce several optimal solutions for a given problem [5]. Application of Genetic Algorithms in Data Mining task is preferred due to large candidate rule space, particularly in case of larger datasets with a large number of attributes. Moreover, GAs copes better with attribute interaction than most greedy rule induction algorithms that are used for the discovery of classification rules [4]. Therefore, the application of evolutionary techniques for knowledge mining has received a renewed emphasis and several sequential and parallel improved implementations of Genetic Algorithms or Genetic Programming to the domain of rule mining have come up. These new implementations improve the efficiency or/and efficacy of the evolutionary algorithm applied through systematic research that include advance preprocessing techniques, reducing the redundancy in fitness evaluations, new encoding schemes and distributed and parallel mining techniques [4], [12], [15], [19]. However, while going through the related research matter, it is noticeable that many of these implementations work on datasets that are pure in the sense that they either have all the attributes as categorical or numeric/continuous. It is obvious that it is too much of a limitation as most of the real world datasets contain multiple types of data variables.

This paper proposes a genetic-fuzzy approach for the discovery of fuzzy decision rules from datasets containing categorical as well as continuous attributes. In this work we have fuzzified the continuous attributes using the traditional triangular fuzzy number and consequently designed a matching process that can take care of both types of attributes during computing the fitness of an individual in GA population. While a direct match is made for the categorical attribute, A Mumdani style (min-max) method is applied for matching fuzzified continuous attributes. We have used a binary encoding and the discovered fuzzy decision rules are in CNF form.

2. RELATED WORK

Evolutionary Algorithms have been applied in abundance covering all phases of knowledge discovery and various data mining tasks like association rule mining, clustering and discovering classification/prediction rules etc. Freitas (2003) has done an exemplary survey paper on application of EAs in data mining. As our research relates to discovery of Fuzzy Classification Rules from datasets that contain mixed attributes, in this section, we review the works that are relatively more relevant and help to contextualize this paper.

Any fuzzy controller required to have fuzzy if-then rules which were usually derived from human experts. Firstly, it was a tedious task to gather knowledge from an expert and secondly this knowledge might be biased with perspective of the expert. Therefore, many approaches were proposed for automatically generating fuzzy if-then rules from training datasets. Yaun & Zhuang (1996) and Janikow (1995) applied Genetic Algorithm for discovery of FCRs and results were found competitive to decision tree induction and the performance was better in terms of predictive accuracy of discovered rule set. Traditionally classification algorithms focused either on accuracy or interpretability [10], [21]. Roubos et al. (2001) proposed fuzzy classifier system that takes into account both accuracy as well as interpretability, so as to keep the discovered rule bases small and comprehensible [18]. Ishibuchi (1995, 1996, 1999) made a remarkable initial contribution for improving the performance of fuzzy classifier systems for pattern classification problems [7], [8], [9]. A hybrid algorithm of two fuzzy genetics-based machine learning approaches (i.e., Michigan and Pittsburgh) for designing fuzzy rule-based classification systems has also been proposed. A new method was also proposed to automatically learn the knowledge base (KB) by finding an appropriate database by means of a genetic algorithm while using a simple generation method to derive the rule base (RB) [2].

Advantage of using co-evolutionary approach is that fitness is evaluated across several rule sets rather than single fuzzy set. Nevertheless, this gain comes with increased processing time. Gene Expression Programming method also uses two populations. One for Fuzzy Classification Rules which is evolved by syntax genetic programming and the other one for membership function definitions which is evolved by mutation based evolutionary algorithm. These two populations co-evolve to better classify the underlying data set [1], [15]. Most of the approaches mentioned above consider smaller datasets of continuous attributes like Wine and Iris datasets from UCI machine learning repository for discovering FCRs. There is a clear need to devise approaches that can work with datasets containing mixed type of attributes.

Some of the recent works that employ evolutionary learning techniques to mine databases include discovery of Censored Production Rules (CPRs), Hierarchical Censored Production Rules with Fuzzy Hierarchy (FCRFHs) [19]. The discovered rules have the capability to exhibit variable precision logic with respect to certainty and specificity in decision making. Another example is Bioinformatics-oriented Hierarchical Evolutionary (BioHEL) that applies the iterative learning approach to mine large scale databases. The proposed attribute list rule representation of BIOHEL is specially suited to datasets with large number of attributes and instances, and this approach has been able to significantly improve the overall performance (e.g. run time, accuracy, rule set size) of rule learning system. Derivation of Fuzzy Rules based on heuristic method of probabilistic clustering from interval-valued data has been suggested by Dmitri in 2010 [24]. A Genetic Programming based method for performing classification on datasets with nominal attributes is developed by Loveared and Ciesielski [25]. Their system was able to match up other well known classification techniques in terms of predictive accuracy.

3. GENETIC-FUZZY APPROACH FOR DISCOVERY OF FCR FOR MIXED ATTRIBUTES DATASETS

A GA is a stochastic and adaptive search method that performs parallel search by maintaining a population of individual solution. In this paper we suggest a Michigan style GA where each chromosome in population represents a rule. Conjunction is used between attributes of rule and disjunction is used between sub-attributes. We use a deterministic Crowding GA which avoids convergence of GA population to a single best rule. In this GA the offspring replace the worst and most similar individual chosen from a small randomly selected population. The continuous attributes of the datasets are first normalized, and then fuzzified using simple triangular fuzzy function. A fixed length binary encoding scheme is mapped to represent rules in CNF form in the GA population. Crossover is restricted to the rules of same class and applied only at sites where an attribute ends and another one begins. As mutation is concerned, simple bit flip mutation served the purpose.

Most of the evolutionary based rule miner systems apply evolutionary process iteratively discovering a single rule at a time. These approaches run the algorithm several times to discover the whole rule set and are obviously computationally expensive. In contrast, our approach evolves a rule set in a single run of the genetic algorithm. However, the final population may contain rules whose fitness is computed on the basis of overlapping instances i.e. several rules may cover the same instances in dataset. Therefore, an exhaustive sequential coverage algorithm is applied to get the final fuzzy classifier system which is a subset of the final population. To apply sequential coverage algorithm, the final population is sorted in decreasing order of fitness. Starting with the best rule, the training instances that are covered by the rules are successively removed from the training set and these rules are inserted into the final classifier until the whole training set is covered or there are no more rules in the final population. Figure1 shows the evolutionary process for automated discovery of Fuzzy Classification Rules. The detailed steps are given below.

3.1 Normalization

First, step is to normalize continuous attributes so that attribute values fall within the specified range of 0 to 1.

$$A' = \frac{A - \min(A)}{\max(A) - \min(A)}$$

Where, A' is the normalized value, A is the actual value of the attribute, max (A) and min (A) are the maximum and minimum values present for the attribute under consideration. For instance, consider a mixed attributes dataset, the Weather dataset modified to have temperature and humidity as the two continuous attributes and outlook and windy as categorical attributes. Instances of this dataset are shown in Table1 and the normalized dataset is shown in Table 2.

Table1. The recoded Weather Dataset

Outlook	Temperature	Humidity	Windy	Play/ Not play
0	40	70	0	0
0	39	72	1	0
1	38	75	0	1
2	25	71	0	1
2	14	45	0	1
2	15	48	1	0
1	13	44	1	1
0	24	74	0	0
0	16	44	0	1
2	26	46	0	1
0	27	43	1	1
1	23	73	1	1
1	41	20	0	1
2	24	72	1	0

Outlook: 0=sunny, 1=overcast, 2=rainy; Temperature in celcius; Humidity in percentage; windy: 0=false, 1=true; Class: 0=not play, 1=play

Table 2. Normalized Weather dataset

Outlook	Tempera-ture	Humidity	Windy	Play/ Not play
0	0.96	0.91	0	0
0	0.93	0.95	1	0
1	0.89	1	0	1
2	0.43	0.93	0	1
2	0.04	0.45	0	1
2	0.07	0.51	1	0
1	0	0.44	1	1
0	0.39	0.98	0	0
0	0.11	0.44	0	1
2	0.46	0.47	0	1
0	0.5	0.42	1	1
1	0.36	0.96	1	1
1	1	0	0	1
2	0.39	0.95	1	0

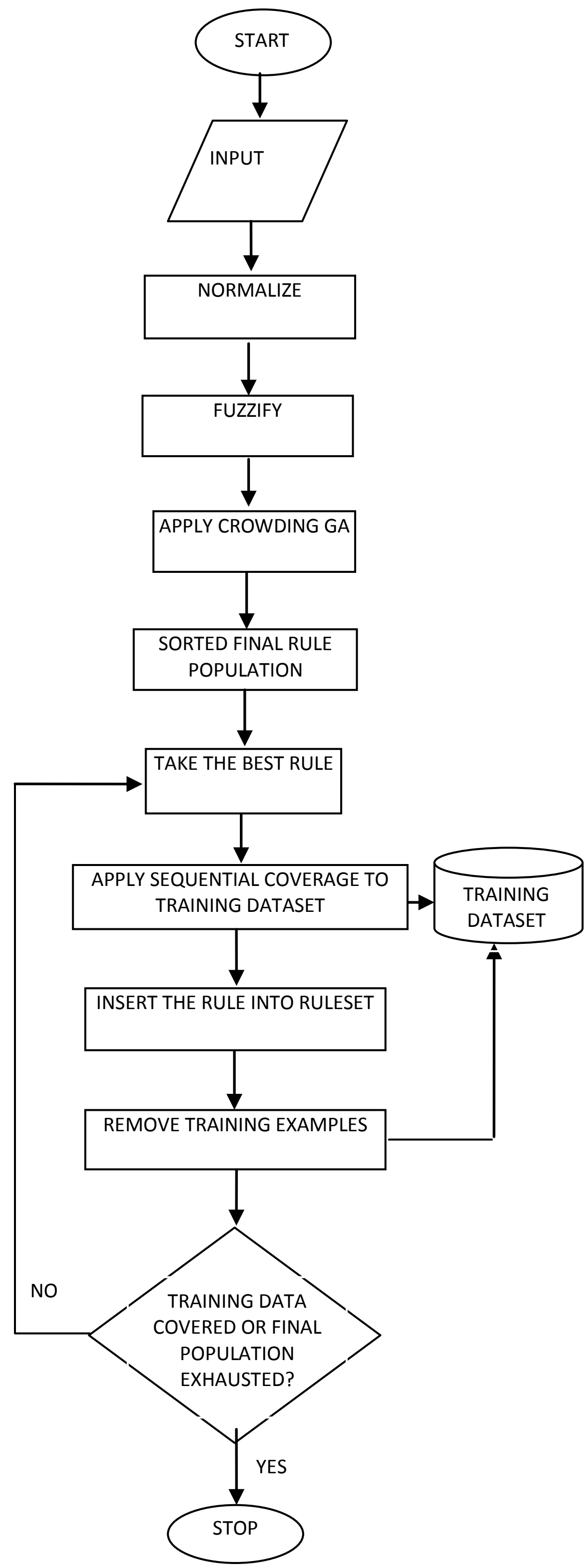

Figure1: Proposed system for automated discovery of FCRs

3.2 Fuzzification

Normalized attribute values are fuzzified by computing values corresponding to three fuzzy modifiers small, medium and large for all the attributes of the given dataset. Triangular fuzzy membership method is shown in figure 2:

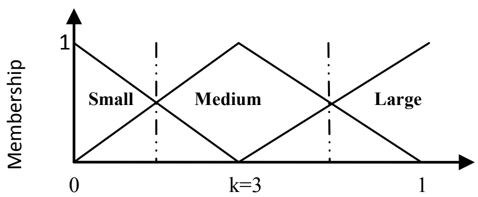

Figure 2: Triangular fuzzy membership function

In fig.2, K represents the number of fuzzy subsets in which each attribute is partitioned. The value of membership function μ_k^i for each linguistic variable A_k^i is given by:

$$\mu_K^i = \max\{(1 - \frac{X - a_K^i}{b^K}), 0\}$$

Where,

$$a_K^i = \frac{i-1}{K-1};\ i = 1,2,...,K \quad \text{and} \quad b^K = \frac{1}{K-1}$$

After employing the above fuzzification method, the fuzzified weather dataset is shown in Table 3

Table 3. Fuzzified Weather dataset

Outlook	Temp.			Humidity			Windy	Play/Notplay
	Low	Medium	High	Low	Medium	High		
0	0	0.07	0.93	0	0.18	0.82	0	0
0	0	0.14	0.86	0	0.11	0.89	1	0
1	0	0.21	0.79	0	0	1	0	1
2	0.14	0.86	0	0	0.15	0.85	0	1
2	0.93	0.07	0	0.09	0.91	0	0	1
2	0.86	0.14	0	0	0.98	0.02	1	0
1	1	0	0	0.13	0.87	0	1	1
0	0.21	0.79	0	0	0.04	0.96	0	0
0	0.79	0.21	0	0.13	0.87	0	0	1
2	0.07	0.93	0	0.05	0.95	0	0	1
0	0	1	0	0.16	0.84	0	1	1
1	0.29	0.71	0	0	0.07	0.93	1	1
1	0	0	1	1	0	0	0	1
2	0.21	0.79	0	0	0.11	0.89	1	0

3.3 Encoding

It has been shown by Ishibuchi that a Fuzzy Classifier Systems based on Michigan approach achieves higher classification accuracy than the Pittsburgh approach [9]; we have used a Michigan style approach where the whole GA population represents a rule set. Each individual rule is represented by a pure fixed length binary string genetic encoding which is very easily mapped to variable length rules. A FCR contains two parts- antecedent (IF part) and consequent (the THEN part) which can be coded as one chromosome with several segments.

Each fuzzied attribute is assigned three bits with following interpretations.

Binary String	**Interpretation**
001	A=High
010	A=Medium
011	A=Medium OR High
101	A=Low OR High
110	A=Low OR Medium

Strings 000 and 111 represent don't care and means absence of attribute. A categorical attribute is assigned bits equal to the number of values for it. For example, for outlook variable in the weather dataset that has three values 'overcast', 'rainy' and 'sunny three bits are assigned and a disjunction is allowed similarly as in the case of fuzzified continuous attributes. Corresponding to the weather dataset the chromosome 110:000:111:00:1 would be interpreted into a FCR given below.

If ((outlool=overcast)||(outlook=rainy)) Then Play
Count: 7, Coverage: 0.5, Fitness: 0.604938

3.4 Fitness Computation

The quality of a rule is known through fitness computations. Many fitness computation criteria like entropy, accuracy, sensitivity, coverage and predictive accuracy have been used in evolutionary approaches to rule mining. Most of these rule quality measures give results that are comparable. The following fitness function is used in this work:

$$\frac{Fitness(R_i) = confidence(R_i) * \mathrm{cov}\,erage(R_i)}{complexity(R_i)}$$

$$confidence(R_i) = \frac{|P \wedge D|}{|P|} \quad \mathrm{cov}\,erage(R_i) = \frac{|P \wedge D|}{|D|}$$

$$complexity(R_i) = no.of\ active\ attributes\ in\ (R_i)\ .$$

In the above fitness function, |P and D| represent number of examples in the training data that satisfy premise as well as the class in the consequent part of the rule. |P| and |D| are number of instances that satisfy the only the premise and class of the rule respectively. The suggested fitness function also works for the datasets with class imbalance. To compute the confidence and coverage of a rule on training dataset, we have designed a procedure that matches the fuzzified continuous attributed on the basis of Mumdani's min-max rule match model whereas a direct match is sought for categorical attributed. To implement the match procedure each attribute is marked continuous or categorical with the help of setting a flag array right in the beginning. The beginning of each attribute in the chromosome is also stored in a list. The complete algorithm for evaluating a rule with respect to the training examples is given in figures 3.

The fitness computation of the example rule **"If ((outlook=overcast)||(outlook=rainy)) Then Play"** is given below:

$$\mathrm{cov}\,erage(R_i) = \frac{|P \wedge D|}{|D|} = \frac{7}{9} = 0.777$$

$$confidence(R_i) = \frac{|P \wedge D|}{|P|} = \frac{7}{9} = 0.777$$

$$complexity(R_i) = 1$$

Fitness(R_i) = : 0.604

Procedure Evaluate Rule
Input: Fuzzified Training dataset, Rule R_i from GA Population, Flag List, List of Attribute beginning
Output: *Fitness* R_i
Flag list = [0,1,0,1] // 0 for continuous attribute and 1 for categorical attribute
Attribute beginning = [I_1,I_2, … I_k]
PandD=0
P=0
D=0
Fitness (R_i)=0
match_continuous = true
match_categorical = ture
For each example in the fuzzified training dataset E_j
  For each attribute *k* of R_i
    If not_active (A_{ik})
      continue;
    else If continuous (*Aik*)
      index=*attribute beginning*[I_k];
      match_continuous=
      $\min_k(\max(\min_m(R_i.index.bit_m, E_j.val_m)))$
        > alpha // *m* varies from 1 to 3 & alpha is α_cut value.
    Else // categorical attribute
        value= E_j.*val*;
        index= *attribute beginning*[I_k]+*val*
        match_categorical= (R_i.*index*==1)
    Endif
  End For each attribute *k*
  If (*Match_continuous* **AND** *Match_categorical* **AND** R_i.*class* = E_j.*class*)
    Increment *Pand D*
  Endif
  If (*Match_continuous* **AND** *Match_Categorical*)
    Increment *P*
  Endif
  If (R_i.*class*= E_j.*class*)
    Increment *D*
  Endif
End For each training example
Compute confidence (R_i)
Compute coverage (R_i)
Compute complexity (R_i)
Compute Fitness (R_i)
Return Fitness (R_i)
End Procedure

Figure 3: Complete Algorithm for evaluating Fitness of a Rule.

4. EXPERIMENTAL SIMULATION AND RESULTS

We have implemented the proposed system using GALIB247. The suggested approach has been tested on six datasets taken from UCI Machine learning laboratory. Two sets are of purely continuous attributes, and the remaining four datasets have mixed attributes, contain both attribute types continuous as well as categorical. The properties of the datasets are described and summarized in the Table 4.

Table 4. Datasets Summary information

Name of the Dataset	Total No. of Attributes	No. of Continuous Attributes	No. of Categorical Attributes	No. of Examples	No. of Classes	Class-wise Distribution of Attributes
Echocardiogram	9	7	2	74	2	50+24
Heart Disease	13	5	8	219	2	164+55
Contraceptive Method	8	2	6	1473	3	629+333+511
Weather-edited	4	2	2	14	2	9+5
Iris	4	4	0	150	3	50+50+50
Wine	13	13	0	178	3	59+71+48

The parameters for each data set were tuned before running the GA for taking final results. A crossover probability of 0.66 and a mutation rate equals to 0.1 seemed to work fine across all the datasets. However, the population size and number of generations were set differently depending on the size of candidate rules search space. Bigger is the dataset's rule search space, bigger the population size and number of generations are required by the genetic algorithms. These parameters for each dataset are given in the table 5.

Table 5. Simulation Parameters

Sr. No.	Parameters	Values
1	Population Size (N_{pop})	50
2.	Crossover Probability (P_c)	0.60
3.	Mutation Probability (P_m)	0.1
4.	Maximum Generations(max_gen)	100

Each dataset was divided into training and test dataset randomly. We used 50 % instances for the training and the remaining 50% were used for testing the predictive accuracy of the fuzzy decision rules discovered. The FCRs discovered for the Weather-edited dataset, Echocardiogram dataset and Heart Dataset along with their fitness are shown in Table 6, Table 7 and Table 8 respectively.

The predictive accuracies of the discovered rules for all the experimental datasets listed in the Table 9 prove the efficacy of proposed technique. To justify the context, the predictive accuracy of rule sets discovered by WEKA using J48, Naïve-Bayes and Neural Network Classifiers are also provided along with proposed system's predictive accuracy for comparison.

The FCRs discovered are more accurate and simple. Moreover, the rules have been discovered by employing fuzzy logic to continuous attributes making them more comprehensible.

5. CONCLUSION AND SCOPE FOR FUTURE RESEARCH

The main contribution of this paper is to propose a genetic algorithm approach for the automated discovery of FCRs from datasets of mixed attributes i.e. continuous and categorical. The proposed approach has integrated fuzzy logic with genetic algorithm for the discovery of FCRs. The discovered rules are in CNF form which is comprehensible. Moreover, the mined rules contain linguistic terms young, middle_aged or senior for a continuous variable like age and not conditions involving crisp vaues like age <= 25. The suggested approach is tested on various datasets from machine learning repository and results are encouraging. The discovered rule set is capable of handling the uncertainty imperative to the decision making processes. The devised matching procedure has been effective on various datasets containing pure categorical attributes, pure continuous attributes, categorical as well as continuous attributes and class imbalance. The fitness function employed is suitable to mine rules with more generalization power and lesser number of attributes. The results are also comprehensive in the performance of predictive accuracy as well. In current work, simple triangular fuzzy membership function is used for fuzzifying continuous attributes. In future, the performance of trapezium, Gaussian and other advanced fuzzy membership could be applied and tested for discovering the fuzzy decision rules. It would also be interesting to discover Fuzzy Censored Production Rules from datasets of mixed attributes.

Table 6. Weather-Edited Dataset Results

Sr. No	Rules: FCRs	Fitness
1.	*If* ((Outlook=Overcast)||(Outlook=Rainy)) *Then* Play	0.6049
2.	*If* (Outlook=Sunny)&& (Humidity=High) *Then* Not-Play	0.6012
3.	*If* ((Humidity=Normal)||(Humidity=High)) *Then* Play	0.6154

Table 7. Echocardiogram Dataset Results

Sr. No.	Rules: FCRs	Fitness
1.	*If* ((Months Survive=Medium)||(Months Survive=Many)) &&((WallMotionScore=Low)||(WallMotionScore=Medium))&&((WallMotionIndex=Low)||(WallMotionIndex=Medium)) *Then* Patient Survive Less Than 1 Year	0.90322
2.	*If* (StillSurvive=No)&&((EPSS=Low)||(EPSS=Medium)) &&((WallMotionIndex=Low)||(WallMotionIndex=Medium)) *Then* Patient Survive Less Than 1 Year	0.87096
3.	*If* (MonthSurvive=Few)&&(StillSurvive=Yes)&&((Age= Low)||(Age=Medium))&&((FractionalShortening=Low)||(FractionalShortening=Medium)) *Then* Patient Survive More Than 1 Year	0.76923

Table 8. Heart Disease Dataset Results

Sr. No.	Rules: FCRs	Fitness
1.	*If* ((ChestPain=TypicalAnginal)||(ChestPain=Atypical)||ChestPain=NonAngial)) &&((Bp=Low)||(Bp=Medium))&&(BloodSugar<120)&&(EnducedEngina=No) &&(Slope=Upslopping) *Then* No Angiographic Disease Less Than 50% Narrow Diameter.	0.6528
2.	*If* ((OldPeak=Low)||(OldPeak=Medium))&&((VesselColorType0)||(VesselColorType2)) *Then* No Angiographic Disease Less Than 50% Narrow Diameter.	0.6216
3.	*If* ((Age=Less)||(Age=Medium))&&(Sex=Male)&&(BP=Medium) &&(Cholestrol=Medium)&&(BloodSugar>120)&&((HeartRate=Low)&&((VesselColorType2)||(VesselColorType3)) *Then* Angiographic Disease Greater Than 50% Narrow Diameter.	0.5152

Table 9. Predictive Accuracy (%) for various approaches on Different Datasets

S. No.	Dataset Name	J48	Naïve Bayes	Neural Network	Proposed Approach
1.	Weather	68	64	78	93
2.	Echocardiogram	92	94	93	90
3.	Heart	75	77	73	74
4.	Contraceptive	52	50	51	67
5.	Wine	93	96	97	95
6.	Iris	96	94	96	95

6. REFERENCES

[1] Akbarzadeh Vahab, Sadeghian Alireza and Santos Marcus V.dos: "Determination of Relational Fuzzy Classification Rules Using Evolutionary Computation," Fuzzy Systems IEEE Int. Con. Fuzzy Syst, pp. 1689-1693, 2008.

[2] Cordon O., F. A. C. Gomide, F. Herrera, F. Hoffmann and L. Magdalena: "Ten Years of Genetic Fuzzy Systems: Current Framework and New Trends", Fuzzy Sets and Systems, pp. 5-31, 2004.

[3] Dries Anton, Raedt Luc De, Nijssen Siegfried: "Mining Predictive k-CNF Expressions,"IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 5, pp. 743-748, May 2010.

[4] Freitas Alex A.: "A Survey of Evolutionary Algorithms for Data Mining and Knowledge Discovery", Advances in Evolutionary Computation Theory and Applications, Springr-Verlag, New York, USA, pp. 819-845, 2003.

[5] Goldberg D.E.: "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley Publishing Company, Inc. MA, New York, 1989.

[6] Han J. and Kamber M.: "Data Mining: Concepts and Techniques," Morgan Kaufmann Publishers, San Francisco, 2000.

[7] Ishibuchi H., T. Nakashima, and T. Murata: "A Fuzzy Classifier System that Generates Fuzzy If–Then Rules for Pattern Classification Problems," in Proc. 2nd IEEE Int. Conf. Evolutionary Computation, Perth, Australia, pp. 759–764, Nov. 29–Dec. 1, 1995.

[8] Ishibuchi H., K. Nozaki,and H. Tanaka: "Adaptive Fuzzy Rule-Based Classification Systems," IEEE Trans. on Fuzzy Systems, vol. 4, no. 3, pp. 238-250, 1996.

[9] Ishibuchi H., T. Nakashima and T. Murata: "Performance Evaluation of Fuzzy Classifier Systems for Multi-Dimensional Pattern Classification Problems", IEEE Trans. Syst.,Man, Cybern., Part B, vol. 29, pp. 601-618, 1999.

[10] Janikow C. Z.: "A Genetic Algorithm for Optimizing Fuzzy Decision Trees," in Proc. 6th Int. Conf. Genetic Algorithms, Univ. Pittsburgh, Pittsburgh, PA, July 15–19, 1995, pp. 421–428.

[11] Lee C. C.: "Fuzzy Logic in Control Systems: Fuzzy Logic Controller," IEEE Transaction System, Man, Cybern., vol. 20, pp. 404–435, Mar./Apr. 1990.

[12] Li Ji-Dong, Xue-Jie Zhang and Yun-Shan Chen: "Applying Expert Experience to Interpretable Fuzzy Classification System using Genetic Algorithms," In Proc. 4th IEEE Int.Conf. on Fuzzy Syst & Knwldg Disc., vol. 02, pp. 129-133, Haikou, Hainan, China, Aug. 2007.

[13] Limin Jia, Ruyan Zhang, Yong Zhang, Zongyi Xing and Guoqiang Cai: "Approach of Fuzzy Classification Based on Hybrid Co-Evolution Algorithm," In Proc. 4th IEEE Int. Conf. on Ntwrk Comp. & Adv. Info. Mgt., vol. 2, pp. 266-271, Gyeongju, Sept. 2008.

[14] Mansoori Eghbal G., Zolghadri Mansoor J. and Katebi Seraj D.: "SGERD: A Steady-State Genetic Algorithm For Extracting Fuzzy Classification Rules From Data," IEEE Trans. Fuzzy Syst., vol.16, no. 4, pp. 1061-1071, Aug.2008.

[15] Mendes Roberto R. F., Voznika Fabricio de B., Freitas Alex A. and Nievola Julio C.: "Discovering Fuzzy Classification Rules with Genetic Programming and Co-Evolution," In Proc. 5th European Con. Knwldg Disc., Lecture Notes In Computer Science, Springer Verlag, vol 2168, pp. 314-325, Sep. 2001.

[16] Noda E., Freitas Alex A. and Lopes H.S.: "Discovering Interesting Prediction Rules with a Genetic Algorithm," In Proc. Congress on Evolutionary Computation (CEC-99), pp. 1322-1329. Washington D.C., USA, July 1999.

[17] Romao Wesley, Frietas Alex A. and Pacheco Roberto C.S.: "A Genetic Algorithm for Discovering Interesting Fuzzy Prediction Rules", Applications to science and technology data," pp. 1188 - 1195, 2002.

[18] Roubos J.A., Setnes M. and Abonyi J.: "Learning Fuzzy Classification Rules from Labeled Data," IEEE Trans. Fuzzy Syst., vol.8, no. 54, pp. 509-522, May 2001.

[19] Saroj and K.K. Bharadwaj: "A Parallel Genetic Algorithm Approach for Automated Discovery of Censored Production Rules", Proc. IASTED Int. Conf. on Artificial Intelligence and Application, ACTA Press, Innsbruck, Austria, pp. 435-441, 2007.

[20] Wang Dianhui, Dillon Tharam S. and Chang Elizabeth J.: "A Data Mining Approach for Fuzzy Classification Rule Generation," IEEE Trans. Fuzzy Syst., vol. 5, pp. 2960-2964, Vancouver, July 2001.

[21] Yuan Yufei and Zhuang Huijun: "A Genetic Algorithm for Generating Fuzzy Classification Rules," ELSEVIER Fuzzy Sets, Syst, vol. 84, pp. 1-19, Nov. 1996.

[22] Zadeh Lotfi A.: "Fuzzy Logic = Computing with Words," IEEE Trans. Fuzzy Syst., vol.4, no. 2, pp. 103-111, May 1996.

[23] UCI Machine Learning Repositiory Databases, http://www.ics.uci.edu/MLRepository.html

[24] Dmitri A Viattchenin: "Derivation of Fuzzy Rules from Interval Valued Data", International Journal of Computer Application, vol. 7(3), pp. 13-20, September 2010.

[25] Andy Song, Thomas Loveard and Victor Ciesielski: "Towards Genetic Learning", SEAL02, vol. 2, pp. 487-491, Orchid Country Club, Singapore, 2002.