

A genetic algorithm approach to automated discovery of hierarchical production rules with fuzzy hierarchy

Basheer Mohamad Al-Maqalet^{1,*} and K. K. Bharadwaj²

¹Faculty of Computer Sciences and Information Systems, Tamar University, Yemen.

²School of Computer & Systems Sciences, Jawaharlal Nehru University (JNU), New Delhi, India.

Abstract

Automated discovery of hierarchical structures in large datasets has been an active research area in the recent years. A concept hierarchy can facilitate mining knowledge at multiple level of abstraction. Crisp description for a concept hierarchy usually cannot represent human knowledge completely and practically. This paper focuses on the issue of mining generalized rules with fuzzy hierarchical structure using Genetic Algorithm (GA) to knowledge discovery. A fuzzy subsumption relation and suitable fitness functions are proposed. Appropriate genetic operators are proposed for the suggested encoding. Finally, Hierarchical Production Rules with Fuzzy Hierarchy (HPRFH) are generated from the discovered hierarchy. Experimental results are presented to demonstrate the performance of the proposed approach.

Keywords: Knowledge Discovery in Databases (KDD), data mining, fuzzy subsumption relation and genetic algorithm.

Received on 11 June 2018, accepted on 27 December 2018, published on 22 January 2019

Copyright © 2019 Basheer Mohamad Al-Maqalet *et al.*, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eai.13-7-2018.156381

*Corresponding author. basheer.almaqalet.dm@gmail.com

1. Introduction

Over the last decade there has been an increasing amount of research in the field of automated learning and discovery, in general, and Knowledge Discovery in Databases (KDD), in particular [8]. KDD can be defined as the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [8]. Data mining is a core stage in the entire process of KDD which applies an algorithm to extract patterns of data [11,17]. Improving the quality of the discovered knowledge is important for making correct decisions in an unpredictable environment [17, 20].

Genetic Algorithms (GAs) consist of stochastic search algorithms that are based upon Darwin's natural selection theory of evolution, where a population is progressively improved by selectively discarding the worse and breeding new offspring from the better [10].

The most predominant representation of the discovered knowledge is the standard Production Rules (PRs) in the

form **If P Then D**. The PRs, however, are unable to handle exceptions and do not exhibit variable precision. As an extension of PR, Michalski and Winston [14] have suggested Censored Production Rule (CPR), which has the form **If P Then D Unless C**, where C (censor) is the exceptional condition.

To address various problems and shortcomings with CPRs system Bharadwaj & Jain [5] introduced the concept of Hierarchical Censored Production Rules (HCPRs) by augmenting CPRs with specificity and generality information. The general form of the HCPR is given as:

Decision **If** <condition>
Unless <sensor>
Generality <general info>
Specificity <specific info>.

Automated discovery of fuzzy hierarchical structure from large database plays fundamentally importance role in data mining because it provides comprehensible results that capture real-life inheritance of objects [12,18]. As a special case (dropping the **Unless** operator) Hierarchical

Production Rules with Fuzzy Hierarchy (HPRFH) takes the form:

$$D_k \text{ If } [P] \\ \text{Generality [general class]} \\ \text{Specificity } [D_{k1}(d1), \dots, D_{ki}(di), \dots, D_{kj}(dN)]$$

where P is the set of preconditions and the **Specificity** element $D_{ki}(di)$ means that D_{ki} is a specific class of D_k with degree of subsumption d_i . The degree of inheritance (subsumption) deg_sub between the general and specific classes gives the fuzzy hierarchical structure.

In this paper, the proposed approach focuses on the automatic generation of HPRFH for nominal type data from a dataset. The rest of this paper is organized as follows:- Section 2 discusses related work. Section 3 describes the fuzzy hierarchical concepts. The design of GA model is introduced in Section 4. Section 5 reports the results of experiments and evaluating the performance of the proposed approach on datasets used. Conclusion and future work directions are given in Section 6.

2. Related Work

Several efforts have been made towards automated discovery of hierarchical structure in large databases [1, 3, 9, 12, 13, 18]. There has been increasing interest in applying evolutionary computation methods [1, 2, 4, 7, 15, 16, 19] as data mining tasks to KDD.

An Evolutionary Approach to automated discovery of Censored Production Rules with Fuzzy Hierarchy (EACPRFH) is proposed in [2]. An automated discovery of Censored Production Rule with Fuzzy Hierarchy(CPRFH) using a parallel Genetic Programming (GP) approach with two advanced genetic operators, namely; fission and fusion is proposed in [6]. They suggested fission and fusion as the special genetic operators to produce offspring from a population of HCPR-trees. An evaluation of hierarchical interestingness measures for mining pairwise generalized association rules is introduced in [7]. They proposed a framework instance and extended the set of performance measures by two novel measures: relation learning and precision which take into account hierarchical relationships.

A combination of rough set and evolutionary approach for automated discovery of CPRFH is presented in [16]. They used rough set methodology to discover characterization set and censors for all the classes in the dataset and then a GA evolves CPRFH. A Hierarchical DEcision Rules (HIDER) discovery using evolutionary algorithm in continuous and discrete domain is proposed in [1]. The evolutionary algorithm used both real and binary coding for the individuals of the population. Mining multi-level association rule at different concept hierarchy is presented in [13]. They described the implementation of multi-level association rule mining efficiently using concept hierarchies.

In the present work, the proposed approach is designed to automated discovery of HPRFH from large

datasets. It integrates the process of hierarchy generation and rule discovery using a novel approach based on GA. Appropriate fuzzy subsumption relation, encoding, suitable genetic operators and effective fitness functions are suggested for the proposed approach. A concept of Frequency Matrix (FM) [3] is used to summarize the large dataset. Also, the proposed algorithm uses the fusion as advanced genetic operator which helps the proposed algorithm to discover new classes/concepts during the evolution process.

3. Fuzzy Hierarchical Concepts

3.1 Fuzzy Subsumption Relation

A class D_i can be defined by a set of properties (values of distinct attributes), $class_prop(D_i)$. Let D_i and D_j be any two classes with the set of properties $class_prop(D_i)$ and $class_prop(D_j)$, respectively.

First, we define a subsumption (i.e., knowing if a class is an ancestor of another) measure between two attributes $P_i(x)$ and $P_j(y)$ (where x and y are frequencies of P_i and P_j with respect to classes D_i and D_j respectively) as follows [2, 3]:

- $subsume(P_i(x), P_j(y))=1$ if ($x \neq \emptyset$) and ($y \neq \emptyset$) and ($x \leq y$) and (attribute $P_i =$ attribute P_j).

- $subsume(P_i(x), P_j(y))=y$ if ($x \neq \emptyset$) and ($y \neq \emptyset$) and ($x > y$) and (attribute $P_i =$ attribute P_j).

- $subsume(P_i(x), P_j(y)) = 0$ if ($x=\emptyset$) or ($y=\emptyset$) or (attribute $P_i \neq$ attribute P_j).

Further, a degree of subsumption (deg_sub) between two classes D_i and D_j is defined as follows [3]:

Let α_i is be i -th property in P_{D_i} and β_j is be j -th property in P_{D_j} .

$$deg_sub(D_i, D_j) = \frac{\sum_{i=1}^{|P_{D_i}|} \sum_{j=1}^{|P_{D_j}|} subsume(\alpha_i, \beta_j)}{|P_{D_i}|} \quad (1)$$

Only the deg_sub with $\max(deg_sub(D_i, D_j), deg_sub(D_j, D_i)) \geq threshold$ will be considered during the construction of fuzzy hierarchy.

A general rule can be represented as:

$P_1(x_1) \wedge P_2(x_2) \wedge \dots \wedge P_i(x_i) \wedge \dots \wedge P_m(x_m) \rightarrow D$, where x_i ($1 \leq i \leq M$) denotes the frequency of the property P_i with respect to class D , where M is the total number of distinct values of attributes, and x_i is computed as follows:

$$x_i = \frac{|P_i \wedge D|}{|D|} \in [0..1] \quad (2)$$

For example, consider D_i and D_j two classes with the set of properties $\{P_1(1), P_2(0.7), P_3(1)\}$ and $\{P_1(1), P_2(1), P_3(0.8), P_4(1)\}$ respectively. The deg_sub between class D_i and D_j is computed as follows:

$$deg_sub(D_i, D_j) = \frac{1 + 1 + 0.8}{3} = 0.933, \text{ and}$$

$$deg_sub(D_j, D_i) = \frac{1 + 0.7 + 0.8}{3} = 0.833$$

If D_i subsumes D_j , then D_i is more general than D_j .

3.2 Discovering Root Properties

Root Property (P_r) is the property(ies), which covers maximum number of classes, i.e. $(FM[P_r, D_k]) \geq threshold$ in the data being mined. The degree of an element, P_i ($1 \leq i \leq M$) with respect to the class D_k ($1 \leq k \leq N$), where N is the number of distinct classes in the dataset, $deg_element(P_i, D_k)$ is defined as follows [2,3]:

$$deg_element(P_i, D_k) = \begin{cases} 1 & \text{if } (FM[P_i, D_k]) \geq threshold \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The degree of a tuple, t_i corresponding to the property P_i , $deg_element(P_i, D_k)$ is computed using the following formula [2,3]:

$$deg_tuple(t_i) = \frac{\sum_{k=1}^N deg_element(P_i, D_k)}{N} \quad (4)$$

The following cases would be considered for the discovery of root property [3]:

Case a: If there is a unique tuple with the highest degree, then the property, P_r , corresponding to that tuple is chosen as the root property.

Case b: In case more than one tuple t_i has the highest value, $deg_tuple(t_i)$, we can compute the sum of frequencies for the corresponding tuple t_i , using the following formula:

$$sum_freq_tuple(t_i) = \sum_{k=1}^N (FM[P_i, D_k]) \quad (5)$$

where $(FM[P_i, D_k]) \geq threshold$.

If there is a unique tuple t_i having the highest value of the $sum_freq_tuple(t_i)$, then the property, P_r , corresponding to that tuple is chosen as the root property.

Case c: In case more than one tuple t_i has the highest value of $sum_tuple(t_i)$, then the set of properties corresponding to the combination of tuples covering the maximum number of classes would form the root property.

3.3 Discovering Strong Properties and Classes

For any class D_k ($1 \leq k \leq N$) a property P_i ($1 \leq i \leq M$) is considered as strong property or weak property, according to the following definition [2,3]:

$$\forall i P_i = \begin{cases} \text{strong} & \text{if } (FM[P_i, D_k]) \geq threshold \\ \text{weak} & \text{otherwise} \end{cases} \quad (6)$$

Only the strong properties for the specified class D_k would participate in the evolutionary process. The weak properties for the specified class D_k , will not participate in the evolutionary process for the class under consideration (D_k) and will not take a part in the fuzzy hierarchy generation.

Strong class is the class which is covered by the root property (ies) and this class would certainly appear in the fuzzy hierarchy. For this class the proposed scheme could discover a HPRFH. Weak class is the class which is not covered by the root property (ies) and this class will not take part in the fuzzy hierarchy. In this case, the proposed approach could discover a standard PR for the weak class. For any class D_k ($1 \leq k \leq N$) is considered as strong class or weak class, according to the following definition [2,3]:

$$\forall k D_k = \begin{cases} \text{strong} & \text{if } (FM[P_i, D_k]) \geq threshold \\ \text{weak} & \text{otherwise} \end{cases} \quad (7)$$

4. Applying GA to HPRFH Discovery

In this Section a GA is presented for the automated discovery of HPRFH as the underlying knowledge representation. The proposed approach handles only categorical attributes and also cannot cope with the missing values. That is why, in each dataset the few instances that contained missing values were simply removed.

4.1 Individual Representation

Suppose there are l predicting attributes with the total number of distinct values of predicting attributes M and a single goal attribute D with the number of distinct values (classes) N in the data being mined. The linear representation of an individual (chromosome) as shown in Figure 1 is divided into three parts: **If** part consisting of a conjunction of conditions on the values (M) of the predicting attributes (l), **Generality**(G) part consisting of one general class for the class under consideration D_k , and the third part is **Specificity**(S) part consisting of a set of specific classes for the class under consideration D_k .

The set of properties $[P_1, \dots, P_M]$	General class $G[D_k]$	The set of specific classes $S[D_{k1}, \dots, D_{ks1}, \dots, D_{ksN}]$
--	---------------------------	--

Figure 1. The chromosome structure.

It is to be noted that any of the l predicting attributes (one or more) can form conditions in the If part. Further, for any PRFH the specific classes are mutually exclusive, i.e. $properties(D_{ks1}) \cap \dots \cap properties(D_{ks1}) \cap \dots \cap properties(D_{ksN}) = \emptyset$.

The decision part does not need to be encoded into the chromosome, as it would be explained later. The structure of chromosome shown in Figure 1 is equivalent to the HPRFH:

$$D_k \text{ If } [P_1, \dots, P_M] \\ G[D_{kf}] \\ S[D_{ks1}, \dots, D_{ksi}, \dots, D_{ksN}]$$

The corresponding tree structure is shown in Figure 2.

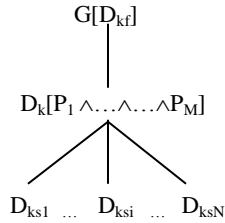


Figure 2. The tree structure of HPRFH.

It is to be noted that the **Generality** part would be empty for the root class and the **Specificity** part would be empty for the leaf classes.

4.2 Genetic Operators

A number of genetic operators suitable for the proposed approach are suggested. We used fitness proportional selection, one-point crossover, with probability 0.90, and mutation operator namely (insert, dropping) with probability 0.20. A crossover point is randomly chosen, represented in Figure 3 by the dotted line and the genes to the right of the crossover point are swapped between two individuals, yielding the new offspring individuals.

Parent 1	P ₁ P ₃ P ₄	D ₂	D ₃ D ₄ D ₅
Parent 2	P ₂ P ₆ P ₈ P ₉	D ₂	D ₉ D ₁₀
Offspring 1	P ₁ P ₆ P ₈ P ₉	D ₂	D ₃ D ₁₀
Offspring 2	P ₂ P ₃ P ₄	D ₂	D ₉ D ₄ D ₅

Figure 3. Crossover operator

Mutation is an operator that acts on a single individual at a time [12]. We developed two mutation operators tailored for our chromosome representation, namely insert, and dropping. While applying mutation operator the proposed approach chooses insert or dropping randomly. As an example in Figure 4, the mutation operator replaces the allele ‘P₃’ by the allele ‘P₄’ in the **If** part, deletes the allele ‘P₅’ in the **If** part, inserts the allele ‘D₂’ in the **Generality** part, whereas, it inserts the allele ‘D₅’ in the **Specificity** part.

Parent	P ₁ P ₂ P ₃ P ₅		D ₃ D ₄
Offspring	P ₁ P ₂ P ₄	D ₂	D ₃ D ₄ D ₅

Figure 4. Mutation operator.

Fusion as an advanced genetic operator is a combination of two HPRFH-trees to generate one or more new HPRFH-trees [5,6]. During construction of the hierarchy a new class D_{new} may be introduced with an appropriate label assigned in accordance with its properties using fusion operator. Whenever the fusion operator introduces new class to the hierarchy, the deg_{sub} for this new class and its descendant is always equal to 1 because the newly created class D_{new} totally subsumes all the classes in the FM [3].

4.3 Fitness Function

The most difficult and most important concept of GA is the fitness function. For the proposed approach, the fitness measure is defined as:

$$\text{Fitness} = F1 + F2 + F3 + F4 \tag{8}$$

$$F1 = \text{Accuracy} * \text{Coverage (for If part)} \tag{9}$$

$$\text{Accuracy} = \frac{|P \wedge D_k|}{|P|} \tag{10}$$

$$\text{Coverage} = \frac{|P \wedge D_k|}{|D_k|} \tag{11}$$

$$F2 = \text{Fitness corresponding to the root tuple (Section 3.2)} \tag{12}$$

$$F3 = \text{deg_sub}(D_{kf}, D_k) \text{ (for G part)} \tag{13}$$

$$F4 = \sum_{i=1}^N \text{deg_sub}(D_k, D_{ksi}) \text{ (for S part)} \tag{14}$$

where N is the number of classes in S part.

5. Computational Results

The proposed approach is implemented and tested on several datasets available publicly from UCI Repository of Machine Learning datasets [http://www.ics.uci.edu/~mllearn/MLRepository.html].

Each GA run consisted of a population of 100 individuals evolving during 200 generations. The proposed approach was terminated when the best fitness did not change continually throughout 20 generations. In the following experiments each decision (class) in a dataset is dealt with separately. During each run, the same decision (class) under consideration is assigned to all the individuals in the population.

The **Then** (decision) part of the rule does not need to be encoded into the individual. This approach simplifies the design of the proposed scheme and it is particularly natural when the user is not interested in a complete classification rule set (where different rules predict different classes) [10]. Each data set was randomly partitioned into two parts with 2/3 of the instances used

for training and 1/3 of the instances used for testing the quality of the discovered rules.

We have compared the predictive accuracy of the rules discovered by the proposed approach with the predictive accuracy of the rules discovered by the EACPRFH approach which is presented in [2]. The performance of the proposed approach on different datasets is demonstrated below:

5.1 Experiments

• Experiment 1

Bridge dataset has 108 examples, 13 attributes (7 specifications, 5 design descriptions and 1 identifier (not used for the classification)) and 6 classes (D_1 = Wood; D_2 =Susten; D_3 = Simple-T; D_4 = Arch; D_5 = Cantilev; and D_6 = Cont-T); P_1 (Erected= Modern); P_2 (Purpose= Highway); P_3 (Length= Medium); P_4 (Clear-G= G); P_5 (T-OR-D= Through); P_6 (Material= Wood); P_7 (Material= Steel) P_8 (Purpose = RR); P_9 (Span = Medium); and P_{10} (T-OR-D= Deck). The proposed approach discovered six rules (5 ad-hoc HPRFH and on PR) as shown in Table 1.

Table 1. Result from the Bridge dataset (*threshold*= 0.60).

No	Discovered Rules	F1	F2	F3	F4	Fitness
1	D_1 If $P_2 \wedge P_5 \wedge P_6$ G [D_2] S []	0.89	4.54	0.67	0.00	6.10
2	D_2 If $P_2 \wedge P_5$ G [] S [D_1, D_4]	0.86	4.54	0.00	1.96	7.36
3	D_3 If $P_8 \wedge P_9$	0.46	0.00	0.00	0.00	0.46
4	D_4 If $P_2 \wedge P_4 \wedge P_5 \wedge P_7$ G [D_2] S [D_5]	0.82	4.54	0.64	0.86	6.86
5	D_5 If $P_2 \wedge P_3 \wedge P_4 \wedge P_5 \wedge P_7$ G [D_4] S []	0.44	4.54	0.69	0.00	5.67
6	D_6 If $P_1 \wedge P_2 \wedge P_{10}$ G [] S []	0.67	4.54	0.00	0.00	5.21

Using the discovered rules in Table 1 and based on the fusion operator the following fuzzy hierarchy is generated as shown in Figure 5.

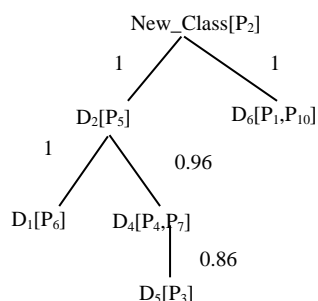


Figure 5. Fuzzy hierarchy for the Bridge dataset.

From the discovered fuzzy hierarchy shown in Figure 5 the following HPRFH are generated.

level 0	New_Class If P_2 Generality [] Specificity [$D_2(1), D_6(1)$]
level 1	D_2 If P_5 Generality [New_Class] Specificity [$D_1(1), D_4(0.96)$]
	D_6 If $P_1 \wedge P_{10}$ Generality [New_Class] Specificity []
level 2	D_1 If P_6 Generality [D_2] Specificity []
	D_4 If $P_4 \wedge P_7$ Generality [D_2] Specificity [$D_5(0.86)$]
level 3	D_5 If P_3 Generality [D_4] Specificity []

• Experiment 2

Nursery dataset has 12960 examples with no missing attributes, 8 attributes and 5 classes. From this dataset, the proposed approach discovered four rules (two ad-hoc HPRFH and two PR) which are given in Table 2.

Table 2. Result from the Nursery dataset (*threshold*=0.60).

No	Discovered Rules	F1	F2	F3	F4	Fitness
1	not_recom If health = not_recom	1.00	0.00	0.00	0.00	1.00
2	recommend If has_nurs = proper \wedge housing = convenient \wedge finance= convenient \wedge health = recommended G [very_recom] S []	1.00	2.00	1.00	0.00	4.00
3	very_recom If housing= convenient \wedge finance = convenient \wedge health = recommended G [] S [recommend]	0.48	2.00	0.00	1.00	3.48
4	spec_prior If health = priority	0.59	0.00	0.00	0.00	0.59

From the discovered fuzzy hierarchy the following HPRFH are generated:

level 0

very_recom **If** housing= convenient \wedge finance= convenient \wedge health= recommended
Generality[]
Specificity[recommend(1)]
 level 1
 recommend **If** has_nurs=proper
Generality[very_recom]
Specificity[]

Experiment 3

The Zoo dataset was used for this experiment. This dataset has 101 examples, 17 predicting attributes and a goal attribute, which can take 7 classes. The predicting attributes were nominal. The proposed approach would discover the following fuzzy hierarchy as shown in Figure 6.

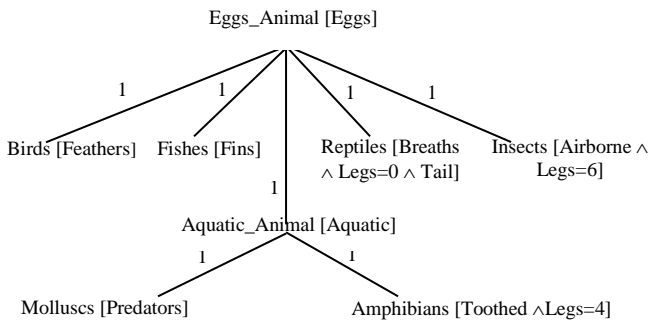


Figure 6. Fuzzy hierarchy for the Zoo dataset (*threshold=0.60*).

Finally, from the discovered fuzzy hierarchy shown in Figure 6, the HPRFH are generated.

Experiment 4

The Balance Scale dataset was used for this experiments. It has 625 instances with no missing attribute values, 4 attributes and 3 classes ($D_1=L$ (288 instances), $D_2=B$ (49 instances) and $D_3=R$ (288 instances)). Choosing a *threshold* =0.50, and based on the FM none of these classes (D_1 , D_2 , and D_3) has strong properties. So the proposed approach would not discover any rule for any class. Therefore, no HPRFH can be discovered for this dataset

5.2 Predictive Accuracy

In the context of rule discovery, it is very common practice to evaluate the quality of discovered rules with respect to their predictive accuracy. As usual in the literature, this evaluation was done by measuring the accuracy rate on the test set.

The results comparing the performance of the proposed approach with the EACPRFH approach [2] are reported in Table 3. First of all, just looking into the performance in Table 3, it is clear that the proposed approach achieves better predictive accuracy results than the EACPRFH approach in Bridge and Nursery datasets.

In Zoo dataset the EACPRFH approach achieves better predictive accuracy result than the proposed approach.

As mentioned in [2] this dataset contains some exceptions. So, the EACPRFH approach performed considerably better than the proposed approach because the EACPRFH approach discovered rules are used much better to classify test examples containing exceptions than the proposed approach.

Table 3. Summary of predictive accuracy results.

Datasets	The proposed algorithm	EACPRFH
Bridge	95.50%	88.75%
Nursery	93.60%	90.20%
Zoo	89.95%	92.85%

Figure 7 depicts the comparative performance of the two algorithms.

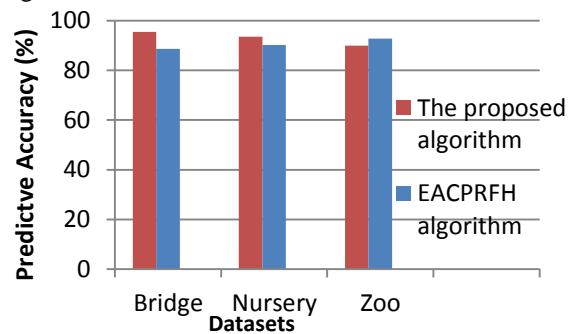


Figure 7. Predictive accuracy of mined rules by the proposed algorithm and EACPRFH algorithm.

6. Conclusion and Future Work

As an attempt towards integrating hierarchy generation processes with the data mining algorithm, a GA approach is presented that discovers Hierarchical Production Rules with Fuzzy Hierarchy (HPRFH).

The discovered hierarchy enables knowledge managers to easily handle the information through generation and specification relationships, and it helps information users to understand the overall structure of a dataset quickly. The proposed approach uses the fusion as advanced genetic operator which helps to discover new classes which are not present explicitly in the dataset being mined. Experimental results have demonstrated the effectiveness of the proposed approach.

One of the most important future research directions would be the discovery of Fuzzy Hierarchical Censored Production Rules (FH CPRs) from large datasets using evolutionary algorithm.

References

[1] Aguilar-Ruiz S., Riquelme C., and Toro M. (2003) Evolutionary learning of hierarchical decision rules. *IEEE Transactions on Systems, Man, and Cybernetics- Part B*, 33(2), pp. 324-331.
 [2] Al-Maqaleh B. M. and Bharadwaj K. K. (2007) Evolutionary approach to automated discovery of censored

- production rules with fuzzy hierarchy. In *Proceedings of the International Multi Conference of Engineers and Computer Scientists, IMECS(2007)*, Hong Kong, China, pp. 716-721.
- [3] Ba-Alwi F. (2004) Discovery of hierarchical production rules with exceptions. *Ph.D Thesis*, SC&SS, JNU, New Delhi, India.
- [4] Bharadwaj K. K. and Al-Maqaleh B. M. (2006) Evolutionary approach for automated discovery of augmented production rules. *International Journal of Computational Intelligence*, 3(4), pp. 267-275.
- [5] Bharadwaj K. and Jain K. (1992) Hierarchical censored production rules (HCPRs) system. *Data & Knowledge Engineering*, North Holland, vol. 8, pp. 19-34.
- [6] Bharadwaj K. and Saroj R. (2010) A parallel genetic programming based intelligent miner for discovery of censored production rules with fuzzy hierarchy. *Expert Systems with Applications*, vol. 37, pp. 4601-4610.
- [7] Benities F., and Sapozhnikova E. (2014) Evaluation of hierarchical interestingness measures for mining pairwise generalized association rules. *IEEE Transactions on Knowledge and Engineering*, 26(12), pp. 3012-3025.
- [8] Fayyad M., Shapiro P., and Smyth P. (1996) The KDD process for extracting useful knowledge from volumes of data. *Communication of ACM*, 39(11), pp. 27-34.
- [9] Feng Q., Miao D., and Cheng Y. (2010) Hierarchical decision rules mining. *Expert Systems with Applications*, 37(3), pp. 2081-2091.
- [10] Freitas, A. (2002) *Data Mining and Knowledge Discovery with Evolutionary Algorithms*, Springer- Verlag Berlin Heidelberg.
- [11] Han J. and Kamber M. (2004) *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, San Francisco.
- [12] Ruchika Y., Kanwal G. and Mittar V. (2015) Representation of concept hierarchy using an efficient encoding scheme. *International Journal of Computer Applications (IJCA)*, 115(11), pp. 28-32.
- [13] Dinesh J. P. and Jagruti H. P. (2015), "Mining multilevel association rule at different concept hierarchy," *Data Mining and Knowledge Engineering*, 7(10).
- [14] Michalski S. and Winston H. (1986) Variable precision logic. *Artificial Intelligence*, vol. 29, pp. 121-146.
- [15] Sarkar K., Sana S. and Chaudhuri K. (2012) A genetic algorithm-based rule extraction system. *Applied Soft Computing*, vol.12, pp.238-254.
- [16] Saroj R, and Bharadwaj K. (2010) Combining rough set and evolutionary approach for automated discovery of censored production rules with fuzzy hierarchy. *International Journal of Intelligent Systems*, 19(1), pp. 47-78.
- [17] Liu Q., Feng G., Wang N., and Tayi G. K. (2018) A multi-objective model for discovering high-quality knowledge based on data quality and prior knowledge. *Information Systems Frontiers*, Springer US, 20(2), pp. 401-416.
- [18] Wang C-H, Lee W-H, and Pang C-T. (2018) Mining association rules for concept hierarchy in large database. *Journal of Interdisciplinary Mathematics*, 21(2), pp. 457-467.
- [19] Kabir M. J., Xu S. Kang B. H., and Zaho Z. (2017) A new multiple seeds based genetic algorithm for discovering a set of interesting Boolean association rules. *Expert Systems with Applications*, vol. 74, pp. 55-69.
- [20] Al-Tapan A. A. and Al-Maqaleh B. M. (2017) An effective mining of exception class association rules from medical datasets. *International Journal of Computer Science Engineering(IJCSE)*, 6(8), pp. 191-198.