International Journal of Approximate Reasoning 67 (2015) 37–58

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

An interpretability improvement for fuzzy rule bases obtained by the iterative rule learning approach

David García [a,1], Juan Carlos Gámez [b,2], Antonio González [a,1], Raúl Pérez [a,*,1]

[a] Dpto de Ciencias de la Computación e IA, ETSIIT, Universidad de Granada, Spain
[b] Dpto de Arquitectura de Computadores, Electrónica y Tecnología Electrónica, Universidad de Córdoba, Spain

ARTICLE INFO

Article history:
Received 15 January 2015
Received in revised form 8 September 2015
Accepted 10 September 2015
Available online 16 September 2015

Keywords:
Fuzzy rule modeling
Interpretability
Complexity reduction
Genetic algorithms
Classification problems

ABSTRACT

Interpretability is one of the key concepts in many of the applications using the fuzzy rule-based approach. It is well known that there are many different criteria around this concept, the complexity being one of them. In this paper, we focus our efforts in reducing the complexity of the fuzzy rule sets. One of the most interesting approaches for learning fuzzy rules is the iterative rule learning approach. It is mainly characterized by obtaining rules covering few examples in final stages, being in most cases useless to represent the knowledge. This behavior is due to the specificity of the extracted rules, which eventually creates more complex set of rules. Thus, we propose a modified version of the iterative rule learning algorithm in order to extract simple rules relaxing this natural trend. The main idea is to change the rule extraction process to be able to obtain more general rules, using pruned searching spaces together with a knowledge simplification scheme able to replace learned rules. The experimental results prove that this purpose is achieved. The new proposal reduces the complexity at both, the rule and rule base levels, maintaining the accuracy regarding to previous versions of the algorithm.

1. Introduction

One of the most important issues when working with genetic fuzzy rule learning algorithms dealing with classification problems is to obtain a good balance between accuracy and interpretability. So, when we face a complex problem, it is interesting not only to obtain a complete rule base with high accuracy, but also to generate interpretable rules easy to understand. In [1], a categorization of the interpretability measures based on two factors is shown:

- Complexity versus semantic interpretability.
- Rule base versus fuzzy partition.

The taxonomy proposed in [1] comes from the combination of both factors. Particularly, we are interested in the combination of complexity-based interpretability at the rule base level. Basically, interpretability in this case is related to the measure of the number of conditions and the number of rules of the fuzzy rule base.

* Corresponding author.
E-mail addresses: dgarcia@decsai.ugr.es (D. García), jcgamez@uco.es (J.C. Gámez), A.Gonzalez@decsai.ugr.es (A. González), Raul_Perez@decsai.ugr.es (R. Pérez).
[1] Tel.: +34 958 244019; fax: +34 958 243317.
[2] Tel.: +34 957 212 039; fax: +34 957 218 315.

http://dx.doi.org/10.1016/j.ijar.2015.09.001
0888-613X/

In this sense, some works have been written related to this topic. In [2], the complexity of the knowledge bases is reduced applying hierarchical-interpolative fuzzy rule bases. Another example is found in [3], where a complexity management methodology for feedback fuzzy rule based systems is proposed. The main advantages of this methodology are its systematic nature and universal scope as well as its focus on reducing qualitative complexity in fuzzy systems. In [4], a two-phase adaptive schema with rule base reduction for fuzzy logic controllers (FLCs) is shown. The first phase is related to the adaptive learning method while the second one automatically generates the fuzzy rules and at the same time applies a genetic reduction technique to obtain the minimum number of fuzzy rules needed for building the fuzzy models. In [5] is proposed a novel method to reduce the number of rules that are fired, keeping the performance of that of large rules, using a membership value of the linguistic variable to calculate an equilibrium value. Thus, rules are fired only if the inputs have membership values higher than the equilibrium value. Some other contributions about this topic can be found, like [6], in which a self adaptive gesture fuzzy classifier is presented. It uses the maximum entropy principle for preserving the most promising rules and removing redundant rules from the rule set. On the other hand, in [7] is proposed a post-processing methodology to improve the interpretability of some well-known algorithms for fuzzy modeling through a complexity reduction based on an accuracy-interpretability trade-off. Anyway, the interpretability and simplicity of the rule bases is a widespread field of study and we are able to find many examples like in [8] and [9]. Some other examples of approaches with a good trade-off between accuracy and complexity of the rule bases can be found in [10–13]. Finally, in [14], a historical review of evolutionary learning methods for designing interpretable genetic fuzzy systems is presented.

As it was said before, a way to understand a rule base as interpretable is that having a reduced number of rules, enough to efficiently represent the knowledge. The Iterative Rule Learning (IRL) approach [15] is based on the use of a Sequential Covering (SC) strategy [16] together with a Genetic Algorithm (GA) as the search component. The algorithms based on this approach, in many cases, extract in final stages rules covering few examples. In fact, these rules are in most cases useless to classify test examples. This effect is produced by the decomposition of the problem followed by the SC strategy, which accelerates the process more than other genetic learning methods. The main disadvantage is that it inherits the usual problems of using greedy methods. One way to solve the problem is to modify the strategy so that when a new rule is learned, apart from adding it to the final set of rules it is possible to review the already learned knowledge, being able to modify the rule base. Thereby, avoiding too specific rules, the rule base is simplified.

So, the aim of this paper is to propose a modified version of the SC and a learning algorithm based on this strategy which tries to reduce the searching space by means of a completeness condition in order to extract rules supported by a representative number of examples. At the same time, the idea is to review the knowledge in each iteration so that rules can be replaced in the rule base. Thereby, on the one hand by replacing rules the rule base is reduced and on the other hand the pruning through the completeness condition allows extracting more general rules.

Thus, this paper is organized in eight sections. Next section is devoted to introduce the main motivation to develop our proposal. Section 3 describes the sequential covering strategy and the fuzzy rule learning algorithm base for our developments, called NSLV. In Section 4, we present a modified version of the sequential covering strategy able to replace learned rules. Next, in Section 5, an improvement of the sequential covering strategy able to search on pruned spaces is proposed. Section 6 explains the integrated version of the proposals described in Sections 4 and 5. Section 7 shows the experimental results and the study associated to these results. Finally, in Section 8, the conclusions are presented.

2. Motivation

The IRL approach is based on the use of the SC strategy together with a GA as a searching component, and it is the base of several fuzzy rule-based learning algorithms [17–19]. The SC strategy [16] uses a decomposition strategy in order to iteratively learn a single rule [20] on each step. By this way, each learned rule is added to the rule base. Commonly, those rules which are being added in each iteration, keep being part of the rule base till the end of the learning process. However, it could be also interesting for a new rule not only to increase the accuracy of the rule base, but also to consider the fact that it is able to improve any other parameter, for example the simplicity and/or the interpretability.

In [21] we explored a modified version of the iterative rule learning approach that included the capability of reviewing the knowledge that was being extracted in previous iterations. In this paper we extend and improve this idea by acting on pruned searching spaces. In this way, the purpose is to simplify the knowledge extracted without losing accuracy.

Thus, the main idea is that instead of only being able to add rules to the rule base, we also consider the possibility of replacing one or more rules by the new learned one. In this case, an important decision is to determine when it is better to add a rule improving the accuracy of the rule base or a rule replacing one or more existing ones. The fact is that it is not easy to measure and compare, from a quantitative point of view, which one among both rules is better. A possible criterion is to give priority to the reduction of the number of rules when the number of successes (examples correctly classified), is smaller than the rules that can be replaced. This makes sense as we are looking for improving the simplicity of the rule base rather than the accuracy. When the proposed algorithm finds a rule with the ability of replacing some others, this rule is usually more general than those which are going to be replaced. This is an important point, because the behavior of the IRL approach tends to iteratively search for the most general rule that correctly classifies a set of examples. Normally, at the beginning of the learning process the algorithm extracts more general rules than in final stages, but it is not always satisfied as it depends on the examples distribution. Usually, when one or more rules are replaced, it is because the new rule is more general than those. Thus, early replacements may condition the rest of the rule extraction because the