Applied Soft Computing 11 (2011) 3463–3475

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

On the extraction of decision support rules from fuzzy predictive models

Félix Castro[a], Àngela Nebot[b,*], Francisco Mugica[b]

[a] Centro de Investigación en Tecnologías de Información y Sistemas (CITIS), Universidad Autónoma del Estado de Hidalgo, Hidalgo, Mexico

[b] Departament de LSI, Universitat Politècnica de Catalunya, Campus Nord, Jordi Girona Salgado, 1-3, Barcelona 08034, Spain

A R T I C L E I N F O

Article history:
Received 15 May 2009
Received in revised form 8 November 2010
Accepted 16 January 2011
Available online 22 January 2011

Keywords:
Fuzzy inductive reasoning
Fuzzy logic
Rule extraction
Virtual campus
Global heat
Brain tumours
Iris
Pima Indian Diabetes

A B S T R A C T

In several application areas like banking, insurance, medicine, education, business, to name just a few, there are huge, sometimes unstructured data collections, and there is a need to convert data into information. On the other hand, decision makers in industry are usually not statisticians, mathematicians, systems engineers, or AI experts. So it is important, to organize the results of the data analysis process and to present it in a form that can be easily interpreted by non-experts.

One of the potential drawbacks affecting the application of computational intelligence (CI) methods in general to the analysis of data is the often limited interpretability of the results they yield. One way to overcome interpretability limitations is by explaining the operation of CI models using rule extraction methods. The interpretability of the model results should be greatly improved by their description in terms of reasonably simple and actionable rules that decision makers could rely on. In fact, rule extraction should provide whom the final responsibility for taking decisions rests, with an explanation about how a CI or related computer-based method has reached its decision.

This paper describes a novel rule-extraction algorithm based on fuzzy logic, name LR-FIR (linguistic rules in FIR), that starts from the fuzzy inductive reasoning (FIR) methodology. FIR is able to obtain good qualitative relationships between the variables that compose the system and to predict the future behaviour of that system. The proposed algorithm (LR-FIR) is able to derive linguistic rules from a FIR model. The LR-FIR functioning is similar to those used in Boolean algebra. However the premises and consequences of rules are not necessarily binary in nature, hence the algorithm must be able to deal with multi-valued logic, and accept partial do-not-care conditions. Due to the fact that LR-FIR was developed within the FIR methodology, the obtained rules could be considered as predictive rules and deal naturally with the uncertainty captured in the FIR models. LR-FIR, in this paper, was evaluated using five data-sets from different domains: e-learning, global change temperature, brain tumour diagnosis, and two of the most used classical UCI data-sets: IRIS and Pima Indian Diabetes. The rules extracted by LR-FIR capture the main behaviour of each application, from the domain experts' point of view, demonstrating in this sense, the efficiency of the proposed algorithm.

1. Introduction

Many methods of data analysis formulated in pattern recognition, neural networks, evolutionary computation, and related fields are aimed mainly at building predictive data models, adapting internal parameters of the data models to account for the known (training) data samples and allowing for predictions to be made on the unknown (test) data samples. Usually these kinds of methods are considered as black box prediction or classification systems. However they accomplish a very high prediction or classification accuracy, nevertheless, frequently they are not very interpretative from the domain experts' point of view. In this sense, reasoning with logical rules is more acceptable than the recommendations given by black box systems, because such reasoning is comprehensible, provides explanations, and may be validated by human inspection. It also increases confidence in the system, and may help to discover important relationships and combination of features, if the expressive power of rules is sufficient for that. Interpretability means that human beings are able to understand the systems behaviour by inspecting the rule base. It is crucial in the data mining field where the knowledge should be extracted from data bases and represented in a comprehensible fashion. It is also essential for decision support systems where the reasoning process should be transparent to the user.

Rule extraction algorithms have been addressed as a topic in areas as diverse as medical, business, climate and e-learning, to name just a few. Due to the fact that there are a lot of research

* Corresponding author.
E-mail addresses: fcastro@uaeh.reduaeh.edu.mx (F. Castro), angela@lsi.upc.edu (À. Nebot), fmugica@lsi.upc.edu (F. Mugica).

1568-4946/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2011.01.018

efforts in rules extraction methods, we consider reasonable organizing the provided information according to the type of modelling techniques used:

- *Neural networks*. A wide study of all aspects of rule generation, optimization, and application can be found in [1], including a discussion about accuracy and simplicity at the rule-extraction stage. It also describes the main approaches to logical rules extraction, i.e. neural networks, decision trees, machine learning, and statistical methods. In [2,3] the author presents two similar approaches (KT and CFNet) to derive a set of rules from a trained neural network. Both KT and CFNet algorithms perform a tree search for the most general valid rules. CFNet uses a feedforward multilayer neural network where the activation function is based on the certainty factor (CF) model. A connectionist (NN-like) mechanism for a rule inference problem alternative to the usual chaining method has been presented in [4]. The proposed approach translates symbolic representation into network representation. A modified relaxation method which improves the convergence speed has been also proposed. Etchells develops an algorithm (OSRE) that extracts comprehensible rules from smooth models, such as those created by neural networks. OSRE extracts rules by taking each data item, which the model predicts to be in a particular class, and searches in the direction of each variable to find the limits of the space regions for which the model prediction is in that class. These regions form hyper-boxes that capture in-class data and they are converted into conjunctive rules in terms of the variables and their values [5].
- *Evolutionary algorithms*. Carvalho and Freitas [6] present a hybrid system to detect and address rules that covers a small number of training examples by means of a decision-tree/genetic-algorithm approach. The authors emphasize that if the rule extraction algorithm disregards the sets of small number of training examples, classification accuracy is considerably degraded. In [7] a genetic algorithm is introduced as an alternative to the classical Pittsburgh and Michigan approaches that obtains a useful rule in each iteration.
- *Fuzzy models*. A high research effort was performed involving the building of fuzzy models to extract rules [7–9]. In [7] a fuzzy theory refinement algorithm (SLAVE + R) including an heuristic process of generalization, specification, addition and elimination of rules is proposed. Lawry [10] presents a methodology for converting probabilistic inference rules into linguistic inference rules, improving their interpretability.
- *Neuro-fuzzy models*. In [11] a modified fuzzy min–max (FMM) supervised neural network to improve the classification performance when a small number of large hyperboxes are formed in the network is proposed. A confidence factor is calculated for each FMM hyperbox, and a user-defined threshold is used to prune the hyperboxes with low confidence factors. From the pruned network fuzzy IF–THEN rules are then extracted. In [12] a learning method for fuzzy classification rules is presented. The proposed approach is based on NEFCLASS. NEFCLASS is used to derive fuzzy classification rules and to learn the shape of the membership functions from a set of data that can be separated in different crisp classes. Chakraborty and Pal [13] present a neuro-fuzzy approach for both, design a classifier and perform a feature selection process in an integrated way. The proposed scheme includes a four-layered feed-forward network for performing a fuzzy rule-based classifier. In the first step, the network learns the relevant features and the classification rules. In the other phases the network is pruned to obtain the best architecture that represents an optimal set of rules, by removing redundant nodes, incompatible and not used rules. The last step is tuning the membership functions of the final rules in order to improve the performance.
- *Tree search based algorithms*. The research works using tree search based models and/or associative algorithms are very usual in the rule extraction area, due to the fact that they are especially easy to combine with other techniques such as neural networks or fuzzy models. In [14] a new rule based method for the automatic selection of kernels based on statistical measures of the data sets is presented. The study evaluates the kernels' performance in terms of accuracy measures with 112 different classification problems, using the popular kernel based statistical learning algorithm SVM. Thabtah and Cowling [15] present a new associative classification technique, i.e. Ranked Multilabel Rule (RMR) algorithm, which generates rules with multiple labels. RMR avoids the overlapping between rules by generating rules that do not share training objects during the training phase, resulting in a more accurate classifier.

Although the wide range of research in the rule extraction area, few of them consider multiple quality factors simultaneously. Some approaches apply multiple quality criteria selection by means of genetic algorithms. For instance, Lopes et al. [16] and Radcliffe and Surry [17] use a GA which fitness function combines multiple factors in a weighted sum or a product. In [18] a GA to find the best subset of features that minimize both the error rate and the size of the tree discovered by a tree induction algorithm is presented.

The LR-FIR algorithm proposed in this paper is developed with the goal to be a useful tool for decision makers. With this purpose in mind the rules extracted by LR-FIR describe in a very intuitive and actionable way the system behaviour.

LR-FIR extracts predictive rules of the type IF–THEN and allows to represent multi-valued logic functions, i.e. neither inputs nor outputs are restricted to binary logic. Instead to avoid overlapping rules, these rules are treated in the compaction and unification steps where rules sharing contiguous input spaces in a feature and the same values in the remaining features are unified in a unique rule. In this way LR-FIR represents as much accurately as possible the system behaviour while preserving the main goal, i.e. the simplicity of the resulting rule base.

The rest of the paper is structured as follows: Section 2 presents the basic concepts of FIR methodology. The LR-FIR algorithm is described in Section 3. In Section 4 three real applications, i.e. e-learning, global climate temperature and brain tumour diagnosis, are introduced and the results are shown and compared with those obtained using other methodologies. Section 5 presents the results obtained by LR-FIR when it is applied to Iris and Pima Indian Diabetes benchmark problems. A discussion about the results obtained is presented in Section 6. Finally, Section 7 wraps up the paper with some conclusions.

2. The fuzzy inductive reasoning approach

The conceptualization of the fuzzy inductive reasoning (FIR) methodology arises from the General Systems Theory proposed by Klir [19]. This modelling and qualitative simulation methodology is based on systems behaviour rather that on structural knowledge. It is able to obtain good qualitative relationships between the variables that compose the system and to infer the future behaviour of that system. It also has the ability to describe systems that cannot easily be described by classical mathematics (e.g. differential equations), i.e. systems for which the underlying physical laws are not well understood. A FIR model is a qualitative, non-parametric, shallow model based on fuzzy logic. Currently FIR methodology runs under the user friendly platform named Visual-FIR, which is available as a Matlab toolkit.

FIR consists of four main processes, namely: *fuzzification, qualitative model identification, fuzzy forecast and defuzzification*. Fig. 1 describes the structure of the FIR methodology.