Neuronal Communication Genetic Algorithm-Based Inductive Learning

Abdiya Alaoui, Department of Computer Science, Djillali Liabes University Sidi Belabbes, Algeria
https://orcid.org/0000-0002-4879-3850

Zakaria Elberrichi, Department of Computer Science, Djillali Liabes University Sidi Belabbes, Algeria
https://orcid.org/0000-0002-3391-6280

ABSTRACT

The development of powerful learning strategies in the medical domain constitutes a real challenge. Machine learning algorithms are used to extract high-level knowledge from medical datasets. Rule-based machine learning algorithms are easily interpreted by humans. To build a robust rule-based algorithm, a new hybrid metaheuristic was proposed for the classification of medical datasets. The hybrid approach uses neural communication and genetic algorithm-based inductive learning to build a robust model for disease prediction. The resulting classification models are characterized by good predictive accuracy and relatively small size. The results on 16 well-known medical datasets from the UCI machine learning repository shows the efficiency of the proposed approach compared to other states-of-the-art approaches.

KEYWORDS
GABIL, Inductive Learning, Medical Datasets, NCA, NCGABIL, Rule Classification

INTRODUCTION

The medical diagnosis identifies the disease or the condition that explains the symptoms and signs of a malady. The information necessary for the diagnosis is usually obtained from the history and physical examination of the patient concerned with medical care.

Medical data has a number of characteristics that make their classification a complex task, the supervised machine learning algorithms are used to treat the healthcare labeled data. Learn from data can be achieved by the design and development of algorithms using Machine learning with understanding intelligence, to enhance their learning behavior over time and obtain knowledge from experience. The pertinent knowledge is hidden in arbitrarily high dimensional spaces, which is not accessible to a human, so the defy is to find them. Machine learning is implementing in various areas, like smart health (Holzinger, 2017).

The classification task of medical datasets in the form of a rule set is an NP-Hard (Cotta & Moscato, 2003) problem. Metaheuristics are solutions to solve these types of problems like Neuronal Communication Algorithm (NCA) and Genetic algorithms (GA). NCA based inductive learning can find the global maximum accuracy for classification of medical data using Genetic Algorithm based inductive learning (GABIL). The GA can be used to many optimization problems as a key element in the design of robust learning strategies like feature selection (Jourdan et al., 2001), Instance selection (Ishibuchi et al., 2001; Min, 2016) and inductive learning (DeJong et al., 1993).
The main advantage of Genetic algorithms is their ability to exploit the data space in a much-detailed level than traditional methods based on gradient or search heuristics and thus to find better solutions. It can be applied in Neuronal Communication Algorithm based inductive learning approach to avoid trapping in local maximum.

The objective of Rule induction algorithms is to discover an explicit rule set formulated in terms of tests for certain values of the attributes, those rules are used as knowledge representation, they can correctly know the instances of the target concept and distinguish them from objects that do not appertain to it. The rules can be read easily by human experts. (Alberto et al., 2010; Furnkranz, 1999), they were very competitive in terms of interpretability.

In this paper, a new hybrid Algorithm for classification model learning is presented. The result classification model is an explicit rule set.

The hybrid approach consists of two Algorithms. Many models are constructed by Neuronal Communication Algorithm (Asil-Gharebaghi & Ardalan-Asl, 2017) based on different instances of the datasets. Genetic Algorithm based inductive learning (DeJong et al., 1993) algorithm is used to generate new solutions, this algorithm replaces the solutions with the worst fitness function with the new solutions, the use of new solutions avoids trapping in local maximum and enhances the result models in terms of classification accuracy.

The paper is structured as follows. In the next section, related works are presented. Section3 describes the involved Algorithms NCA and GABIL. Section 4 explains the proposed hybrid approach. Empirical results are shown and discussed in Section5. Concluding remarks are given in the last section.

RELATED WORK

This section explores the research works which are related to the proposed hybrid approach.

The authors (Alberto et al., 2010) give an exhaustive survey of the genetics-based machine learning (GBML) Algorithms for rule induction, an experimental analysis for classification is shown in this article. A comparative study of the GBML algorithms with other non-evolutionary algorithms is presented, the evolutionary algorithms (EA) give a competitive results compared to a classical machine learning algorithms (CART analysis (Breiman et al., 1984), AQ (Michalksi et al., 1986), CN2 (Clark & Niblett, 1989), C4.5 (Quinlan, 1993), C4.5-Rules (Quinlan, 1995), Ripper (Cohen, 1995) in terms of classification accuracy, those algorithms with their references are shown in Table 1

In the Michigan approach, the rule set is updating by a consecutive observation of training examples with their classification (Booker et al., 1989; Holland, 1986). When the rules are learned by EA Repeatedly and it eliminates the examples wrapped by each new rule from the training set, the approach is iterative rule learning (IRL) (Venturini, 1999).

If the rules/chromosomes are improved in the Evolutionary Algorithms by competition between them to be the best rule and to rest in there the rule base, the approach is The genetic cooperative competitive learning (GCCL) (Greene & Smith, 1993). The rules set in the Pittsburgh Approaches are encoded into the chromosomes and evolved until convergence. In Hybrid Evolutionary Decision Trees (HEDT) the authors find the use of decision trees with GA.

Self-adjusting Associative Rules Generator (SARG) is proposed by the author (Lavangnananda, 2006) to produce associative rules for classification. SARG is an enhancement of Genetic Programming for Inductive learning (GPIL) (Lavangnananda, 2004). In GAPIL, the crossover method generates new chromosomes randomly with fitness values that can be higher by chance, the rank selection is not appropriate when the better rules cannot be produced after a number of generations. But in SARG, the authors used two methods MaxToMin Crossover and Self-adjusting selection.

MaxToMin crossover can generate better chromosomes while self-adjusting selection can give a new track to generate new chromosomes when there is no development after a consecutive number of generations.
High Sierra Industries: IS Gridlock
www.igi-global.com/article/high-sierra-industries/3203?camid=4v1a