AIRS-GA: A Hybrid Deterministic Classifier Based on Artificial Immune Recognition System and Genetic Algorithm

Ilyes Jenhani  
College of Computer Engineering and Science  
Prince Mohammad Bin Fahd University  
Kharob, KSA  
Email: ijenhani@pmu.edu.sa

Zied Elouedi  
LARODEC  
ISG Tunis, Université de Tunis  
Tunis, Tunisia  
Email: zied.elouedi@gmx.fr

Abstract—For many engineering and decision-making problems, a deterministic model is more suitable than a stochastic one. In fact, in domains like medicine, security, military, etc., decision makers require to utilize models that, for the same input parameters (i.e. factors), provide the same output(s) (i.e. risk-free outcome(s)). The Artificial Immune Recognition System (AIRS); an immune system-inspired classifier, is one such stochastic method. In this paper, we propose AIRS-GA: a hybrid approach based on AIRS and a deterministic version of the Genetic Algorithm (GA). This hybridization lets AIRS develop its memory cells in a deterministic way. Experiments carried out on real datasets obtained from the U.C.I. machine learning repository showed that AIRS-GA often outperforms the original AIRS2 in terms of classification accuracy and time.

Keywords—Artificial Immune Recognition System, Genetic Algorithm, Machine Learning, Classification.

I. INTRODUCTION

Artificial Immune Systems (AIS) [1] are a class of bio-inspired systems that mimic natural immune system processes. Among the AIS-based methods, we are interested in the Artificial Immune Recognition System (AIRS): an immune-inspired supervised learning technique based on clonal selection theory of natural immune systems [2]. AIRS is considered as one of the most widely applied AIS-based classifiers which has shown to be a good competitor to well-known classifiers such as decision trees, naive Bayes networks, artificial neural networks, etc, [3].

Roughly speaking, the AIRS classifier involves a training-reduction stage followed by a classification stage which uses the k-Nearest Neighbor (kNN) approach. More specifically, the training-reduction stage will result on a pool of memory cells which represents a reduced training set containing the most representative training instances. These latter will be used by a kNN classifier to classify unseen instances.

The AIRS algorithm is a non-deterministic classifier since it uses some random mechanisms which may fluctuate its classification performance, thus producing some uncertainty in decision making. Furthermore, AIRS requires an important number of predefined parameters, values of which can profoundly sway the classification performance.

Generally, each time a given set of inputs is presented, a deterministic algorithm makes the same computations and choices to end up with the same output. However, a non-deterministic algorithm makes random choices producing thus different outputs. Hence, non-deterministic algorithms exhibit a variety of fluctuations resulting from random variation that can have an important effect on outputs. In the classification context, a non-deterministic classifier, also called set-valued classifier [4], is a classification algorithm that results in different classifier models each time a training is performed, thus possibly providing different predictions for a same set of instances. Such behaviour is not desirable for a wide range of problems such as scheduling and task allocation, prediction and forecasting, diagnosis, etc.

In this paper, we propose a deterministic version of the AIRS2 classifier (i.e. version 2 of AIRS) by using a deterministic version of the Genetic Algorithm (GA) [5] as a basis of its memory cell search and development procedure. Genetic Algorithms have been successfully applied to a several real-world optimization and search problems [6]. As such, the genetic algorithm has been adapted to our problem and hybridized with the AIRS2 algorithm so that mutation will be made in a deterministic manner. The proposed AIRS-GA hybrid approach shows interesting classification performance with a minimum number of required parameters.

The paper is organized as follows: Section II provides an overview of the AIRS2 and Genetic algorithms. Section III presents the new features of the AIRS-GA approach and shows how the Genetic Algorithm is used within AIRS2. Section IV analyzes results obtained from applying AIRS2 and AIRS-GA on selected datasets from the U.C.I. machine learning repository. Finally, conclusions, challenges and future work are presented in section V.

II. BACKGROUND

In this section, we provide a brief overview of the two bio-inspired computational methods used in our approach, namely,