Proposing a Features Preprocessing Method Based on Artificial Immune and Minimum Classification Errors Methods

M. Miralvand, S. Rasoolzadeh* and M. Majidi

Department of Computer Engineering
Malayer Branch, Islamic Azad university
Malayer, Iran
* Siam.rasoolzade@gmail.com

ABSTRACT
Artificial immune systems that have been inspired from organic immune systems, have drawn many attentions in recent years (and have been considered) as an evolutionary algorithm and have been applied in different papers. This algorithm can be used in two different areas of optimization and classification. In this paper, an artificial immune algorithm has been applied in optimization problem. In particular artificial immune systems have been used for computing the mapping matrices and improving features. Comparison of results of proposed method with other preprocessing methods shows the superiority of the proposed method so that in 90% of cases it has the best performance based on different measures. Evaluation measures are including classification rate, variance and compression measure.

Keywords: Artificial Immune Systems, Evolutionary Algorithm, Optimization Problem.

1. Introduction

Bio-Inspired Computing lies within the realm of Natural Computing, a field of research that is concerned with both the use of biology as an inspiration for solving computational problems and the use of the natural world experiences to solve real world problems. The increasing interest in this field lies in the fact that nowadays the world is facing more and more complex, large, distributed and ill-structured systems, while on the other hand, people notice that the apparently simple structures and organizations in nature are capable of dealing with most complex systems and tasks with ease. Artificial Immune Systems (AIS) is one among such computing paradigms, which has been receiving more attention recently and have widely been used in different application areas such as clustering/classification, anomaly detection, computer security, learning, web mining, numeric function optimization and combinatory optimization [1].

Artificial immune systems (AIS) have been inspired from mammal's immune systems against various diseases. These methods have been used in two areas of optimization and classification. In [2,3,4] the artificial immune algorithms have been grouped into two categories: 1.population based methods, 2. network dependent methods. Each one of these groups has two branches. This categorization is illustrated in Figure. 1. Generally clonal model is applied in optimization problems and negative selection in classification and clustering [5, 6]. Since the present work mainly focuses on population based methods, we introduce this model in the next sections.

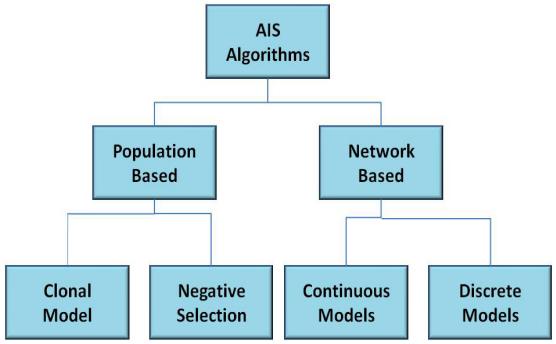

Figure 1. AIS categories.

In this paper artificial immune systems based on clonal selection model have been used and the

optimization and features space transformation problems have been solved based on these algorithms.

Feature improvement and feature space transformation are one of the preprocessing methods that cause the amount of overlapping between dataset classes decrease and then classification of packages can be done very carefully at new feature space. One of the features transformation methods is based on Minimum Classification Error (MCE) algorithms [7].

Minimum Classification Error algorithms are efficient and effective methods of feature transformation and have many applications in the fields of data mining and machine learning [8]. This method as a discriminate method can be combined with each classification method and can be proposed as a preprocessing operation. The goal of this method is that with the aim of mapping on the features they could be transmitted into new space so the overlapping of classes in new space becomes less than previous state. This operation causes the classification error reduction. In this method the mapping is done with the multiplying of mapping matrix into data sets samples [9,10].

Mapping matrix of minimum classification error method is computed based on classification error and for computing this matrix in each step, mapping matrix changes so that classifying error reduces. The most important part of minimum classification error method is mapping matrix. Various methods have been presented for computing the matrix [8]. In this paper we use Artificial Immune System to calculate the mapping matrix [11].

The rest of the paper is organized as follow: in next section we have a brief explanation on AIS cycle and its operators. In section 3, the proposed method for calculating the transformation matrix is described. In section 4, a number of datasets are given and the experimental results from these datasets are presented and compared with the results of other methods. Finally in section 5 we discuss the benefits and limitations of our approach and came to a conclusion.

2. Artificial immune systems in solving optimization problems

All living beings are endowed with an immune system whose complexity varies according to their characteristics. Animals containing bones developed a highly effective and complex immune system. It's composed of a vast array of cells molecules and organs that work together to maintain life. The focus here will be on the immune system of vertebrates, more specifically of humans. When human's body faces with an external and unknown factor that has been entered into its body, bone marrow begins to generate and to proliferate cells and antibodies that can detect the external factor. For this purpose the bone marrow begins to generate different cells and each cell secretes a special type of antibody. Cells that secreted antibody by them better identifies and nurtures the external entered factor, is recognized as winner cell and from this time force bone marrow generates more of this cell, finally the winner cell will fill the most body's immune systems cells against external factors and secreted antibody from this cell is saved on bone marrows memory.

The mentioned procedure in this subsection is called clonal model. In this model two functions play the main role, these are the maturity function and affinity function. Fig. 2 shows different parts of the model.

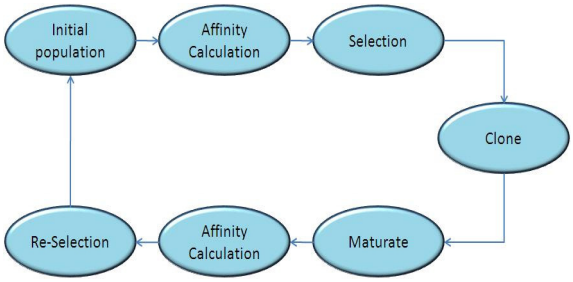

Figure 1. The AIS cycle.

Initial population. the first step is the initial population valuation. Each member of this set is called a chromosome. In this step a set of valid solutions for the problem is produced randomly, additionally, there is a memory that is called M and it exists beside this set. In each execution