Available online at www.sciencedirect.com
ScienceDirect
Engineering Applications of Artificial Intelligence 20 (2007) 571–585
Engineering Applications of ARTIFICIAL INTELLIGENCE
www.elsevier.com/locate/engappai

Comparing binary and real-valued coding in hybrid immune algorithm for feature selection and classification of ECG signals

Michał Bereta[a], Tadeusz Burczyński[a,b,*]

[a]Institute of Computer Modelling, Artificial Intelligence Division, Cracow University of Technology, Warszawska 24, 31-155 Cracow, Poland
[b]Department for Strength of Materials and Computational Mechanics, Silesian University of Technology, Konarskiego 18a, 44-100 Gliwice, Poland

Received 10 November 2006; accepted 20 November 2006
Available online 4 January 2007

Abstract

The paper presents a new algorithm for feature selection and classification. The algorithm is based on an immune metaphor, and combines both negative and clonal selection mechanisms characteristic for B- and T-lymphocytes. The main goal of the algorithm is to select the best subset of features for classification. Two level evolution is used in the proposed system for detectors creation and feature selection. Subpopulations of evolving detectors (T-lymphocytes) are able to discover subsets of features well suited for classification. The subpopulations cooperate during evolution by means of a novel suppression mechanism which is compared to the traditional suppression mechanism. The proposed suppression method proved to be superior to the traditional suppression in both recognition performance and its ability to select the proper number of subpopulations dynamically. Some results in the task of ECG signals classification are presented. The results for binary and real coded T-lymphocytes are compared and discussed.

Keywords: Artificial immune system; Feature selection; ECG signals classification; Negative selection; Clonal selection; Evolutionary feature selection; Evolutionary algorithms; Immune metaphors; Hybrid immune algorithm

1. Introduction

Selecting a proper subset of features is an important task in classification problems, especially when the dimensionality of the measured space is high. The goal of feature selection is to reduce the dimensionality of the problem and to retain the characteristics necessary for recognition, classification and/or data mining process. Feature selection is a subject of great interest in the research areas such as pattern recognition (Kudo and Sklansky, 2000), statistics (Chen, 2003; Boyce et al., 1974; Miller, 1990), data mining (Chen et al., 1996), machine learning (Hall, 1998), neural networks and many others. There are many methods for this task and the methods based on biological metaphors are among them.

Let us denote F as a space of all available features. The cardinality of F is n. The task of feature selection is to find a subset of features F' with cardinality of n' ($n' \leqslant n$) and with $\mathrm{Cr}(F')$ optimised, where $\mathrm{Cr}(F')$ is a given criterion of the selection of a given feature space. If n' is known, the number of possible combinations is equal to $\binom{n}{n'}$. The number of combinations grows exponentially and the problem is even more difficult if there is no assumption made on n' and there are 2^n possible combinations. For that reason a feature selection task very seldom can be solved by means of exhausted search and can be considered as belonging to the group of NP-hard problems. The possible solutions for finding suboptimal feature subsets can be categorised by means of different criteria. From the point of view of heuristic search methods examples of these criteria can be the starting point in the feature space, the organisation of search procedure, the evaluation strategy for selected feature subsets or stopping criteria. All feature selection methods can be categorised in two general groups: classifier-specific, if the usefulness of selected features is evaluated by a given classifier, and classifier-independent,

*Corresponding author. Department for Strength of Materials and Computational Mechanics, Silesian University of Technology, Konarskiego 18a, 44-100 Gliwice, Poland. Tel.: +48322371204; fax: +48322371282.
E-mail addresses: beretam@torus.uck.pk.edu.pl (M. Bereta), Tadeusz.Burczynski@polsl.pl, tb@polsl.pl (T. Burczyński).

0952-1976/$ - see front matter
doi:10.1016/j.engappai.2006.11.004

where feature subset selection is performed considering no specific classification method which will be used after feature subset selection procedure. Thus, classifier independent feature subset selection has to use its own criterion to evaluate the selected feature subset. This leads to three typical distinctions among existing approaches to feature selection methods: filter selection models, wrappers selection models and embedded selection models. Selection methods that utilise an independent criterion to select a proper subset of features before the learning process of a given classifier is performed were termed as the filter selection methods (John et al., 1994). The irrelevant features are filtered out independently of the classification method. Some examples of filter methods are the FOCUS algorithm (Almuallim and Dietterich, 1991) or the RELIEF algorithm (Kira and Rendell, 1992). The wrappers methods use the performance of a given classifier as the evaluation method. For that reason the selected features can be tuned for the specific classification methods and better results can be achieved. However, if the procedure of classifier creation is computationally demanding, wrappers methods become intractable. Genetic algorithms (GAs) applied for feature selection described in the next section are examples of wrappers methods. The last group of selection methods is the group of embedded selection methods. In this group the selection process is not considered as a wrapper around the classifier learning process but is rather an implicit part of the classifier construction and training. Examples of embedded models are the decision tree algorithms ID3 and C4.5 (Quinlan, 1986, 1992) or CART (Breiman et al., 1984).

It is worth mentioning that feature selection can be considered an optimisation problem with many competing criteria, which can be, for example, minimising the number of selected features, maximising the percentage of correctly classified training samples, time complexity etc.

The rest of the paper is organised as follows. In Section 2 some work on evolutionary methods applied to feature selection are mentioned. Section 3 gives a brief overview of the main concepts of the artificial immune systems (AIS) and in Section 4 the proposed method for feature selection and classification based on AIS is described. Section 5 presents some results in the problem of recognising pathological ECG signals and a comparison of binary and real coded negative selection methods in AIS is given. Conclusion is drawn in Section 6.

2. Previous work on evolutionary feature selection

As mentioned in the previous section, evolutionary methods (like GA) as feature selectors belongs in general to the group of wrappers methods. GA can be hybridised with a huge range of possible classification models as the GA can be considered as a general search engine in optimisation. Siedlecki and Sklansky (1993) designed the GA for feature selection. Genetic feature selection was extended by Raymer et al. (2000) where not only a feature subset was selected but also features' weights were optimised. Other works on applying the GA to a feature selection task are the works of Lee (2004) or Cantú-Paz (2004). What all methods based on the GA have in common is that they optimise not a single solution but the population of solutions. In this population each chromosome of length n (the number of features) consists of zeros and ones indicating selected features. Both the classification error and the number of selected features can be minimised constituting a multi-objective optimisation problem. The results depend on the properly defined fitness function. Weighted sum approach is one of the widely used approaches where each objective is considered in the fitness function with a given weight giving scalar fitness value. Penalty functions are used as well, however, it is a known fact that it makes the algorithms very sensitive for changing the parameters of the penalty function. For that reason some researchers aim to find the optimal feature subset by means of Pareto based selection approach. An example is the work of Emmanouilidis et al. (1999) where the multi-objective niching GA tries to find the Pareto optimal feature subset.

AIS, described in detail in Section 3, have been seldom applied to feature selection tasks. In the work of Ando and Iba (2003) AIS was applied to evolve weights for features and the criteria were maximising the classification rate and the margin of the linear classifier. Polat et al. (2005) developed a hybrid immune algorithm FS-AIRS (feature selection artificial immune recognition system), however, the feature selection task was performed by a sub-procedure of C4.5 algorithm for decision tree induction. The approach presented in this paper is different.

This paper presents a new algorithm based on the immune metaphor, which is suitable for feature selection task. While the existing AIS are often based on only one immune metaphor, negative or clonal selection, the presented method involves both of them, resulting in a new, hybrid method when compared to the existing ones. In opposite to the methods described earlier, in the proposed system the feature selection, classification and evaluation of selected subsets of features are performed by artificial immune methods. Thus the proposed method can be classified as belonging to the group of embedded feature selection methods. Usefulness of this method is presented in ECG signals classification task.

3. Artificial immune systems

The AIS (Wierzchoń, 2001; de Castro and Timmis, 2002) have become popular over the last years. Many specialised AIS have been developed for different tasks such as intrusion and anomaly detection, optimisation and clustering (Dasgupta and Forrest, 1996, 1999; D'haeseleer et al., 1996). They all share the same principle of imitating real immune system, its B- and T-lymphocytes and mechanisms like somatic hyper-mutation, clonal and negative selection.