OPT-*i*
An International Conference on
Engineering and Applied Sciences Optimization
M. Papadrakakis, M.G. Karlaftis, N.D. Lagaros (eds.)
Kos Island, Greece, 4-6 June 2014

SELF-ADAPTIVE MULTI-OBJECTIVE GENETIC ALGORITHMS FOR FEATURE SELECTION[1]

C. Brester[2], E. Semenkin[2], M. Sidorov[3], W. Minker[3]

[2] Siberian State Aerospace University,
Krasnoyarsky rabochy avenue, 31, 660014, Krasnoyarsk, Russia,
[2]abahachy@mail.ru, [2]eugenesemenkin@yandex.ru

[3] University of Ulm,
Albert-Einstein-Allee, 43, 89081 Ulm, Germany,
[3]{ maxim.sidorov, wolfgang.minker}@uni-ulm.de

Keywords: Multi-Objective Optimization, Genetic Algorithm, Self-Adaptation, Feature Selection, Emotion Recognition.

Abstract. *Incorporating multi-objective optimization procedures in the classification process allows to achieve a trade-off between the accuracy of the final solution and the number of features involved in supervised learning. In our research we investigate a number of self-adaptive multi-objective genetic algorithms as a tool to select the most essential attributes from the database. A probabilistic neural network is implemented to evaluate the relevancy of reduced feature sets. The high performance of the developed approach is demonstrated on the speech-based emotion recognition problem. For the engaged data set it became possible not only to improve the classification accuracy by up to 28.83% but also to decrease the number of features from 384 to 85.5 on average.*

[1] Research is fulfilled with the support of the Russian Science Foundation.

1 INTRODUCTION

Classification procedures based on the supervised learning approach imply the presence of labeled sampling data. In most cases we have to deal with raw information which includes highly-correlated features, measures with errors or attributes with the low variation level. As a result, using of irrelevant data by the learning algorithm is likely to deteriorate its performance.

Apparently, improvement of the classifier's predictive ability might be achieved through eliminating non-informative features from the database. In this paper we propose to accomplish the heuristic search of essential attributes by means of multi-objective genetic algorithms (MOGAs). The possibility to take into account several criteria at once allows to minimize the relative classification error and the number of selected features simultaneously.

To incorporate MOGAs in the developed scheme it was necessary to elaborate some significant aspects. First, the self-adaptation concept [1] was borrowed to provide the realized approach with automatically adjusted items. Actually, it was a good alternative to occasional choice of genetic operator variants or multiple runs of the algorithm to reveal the best settings. Secondly, applied MOGAs should be oriented to the nature of optimized criteria because algorithms had to operate with both continuous and discrete objective functions in diverse value areas.

The proposed approach was applied to the speech-based emotion recognition problem that was one of the crucial opportunities to improve the quality of spoken dialogue systems. Two data sets (37- and 384-dimensional feature vectors) representing the acted German language corpus were engaged in the series of experiments for comparison the effectiveness of different MOGAs. It was found that due to implementation of the heuristic search there was an opportunity to improve the emotion recognition accuracy by up to 28.83% and reduce the number of features from 384 to 85.5 on average.

This paper is comprised of the following parts: Section 2 provides the brief overview of important studies related to application of heuristic methods in the feature selection process. Section 3 includes the description of used MOGAs, their basic stages and modifications. The problem definition, conducted experiments and results are introduced in Section 4. Conclusion and future plans are presented in Section 5.

2 RELATED WORKS

Generally, the feature selection procedure can be organized as the *wrapper* approach or the *filter* one [2]. The first technique involves classification models to evaluate the relevancy of each feature subset. Although it requires high computational resources, this approach demonstrates the effective work due to adjustment to an applied classifier. The second technique is referred to the preprocessing stage because it extracts information from the data set and reduces the number of attributes taking into consideration such measures as consistency, dependency, and distance. On the one hand, this approach needs significantly fewer calculations therefore it is rather effective in the computational effort sense. However, it does not co-operate with a learning algorithm during feature selection and so ignores its performance entirely.

Yang and Hanovar (1998) used one-criterion genetic algorithm (GA) to determine relevant attributes in order to improve quality of classification realized with neural networks [3]. Li Zhuo *et al*. (2008) accomplished classification of hyperspectral images with support vector machine, they also engaged one-criterion GA to remove non-informative features [4]. In both cases the feature selection procedure was combined with supervised learning algorithms based on the wrapper approach scheme.