



Multiobjective optimization of classifiers by means of 3D convex-hull-based evolutionary algorithms

Jiaqi Zhao^{a,*}, Vitor Basto Fernandes^b, Licheng Jiao^a, Iryna Yevseyeva^c,
Asep Maulana^d, Rui Li^e, Thomas Bäck^d, Ke Tang^f, Michael T.M. Emmerich^d

^a Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, International Research Center for Intelligent Perception and Computation, Joint International Research Laboratory of Intelligent Perception and Computation, Xidian University, Xi'an, Shaanxi Province 710071, China

^b School of Technology and Management, Computer Science and Communications Research Centre, Polytechnic Institute of Leiria, Leiria 2411-901, Portugal

^c Faculty of Technology, De Montfort University, Gateway House 5.33, The Gateway, LE1 9BH Leicester, UK

^d Multicriteria Optimization, Design, and Analytics Group, LIACS, Leiden University, Niels Bohrweg 1, 2333-CA Leiden, The Netherlands

^e Microsoft Research Asia, Beijing 100190, China

^f Nature Inspired Computation and Applications Laboratory (NICAL), the USTC-Birmingham Joint Research Institute in Intelligent Computation and Its Applications (UBRI), School of Computer Science and Technology, University of Science and Technology of China (USTC), Hefei, Anhui 230027, China

ARTICLE INFO

Article history:

Received 27 July 2015

Revised 22 April 2016

Accepted 22 May 2016

Available online 27 May 2016

Keywords:

Convex hull

Classification

Evolutionary multiobjective optimization

Parsimony

ROC analysis

Anti-spam filters

ABSTRACT

The receiver operating characteristic (ROC) and detection error tradeoff (DET) curves are frequently used in the machine learning community to analyze the performance of binary classifiers. Recently, the convex-hull-based multiobjective genetic programming algorithm was proposed and successfully applied to maximize the convex hull area for binary classification problems by minimizing false positive rate and maximizing true positive rate at the same time using indicator-based evolutionary algorithms. The area under the ROC curve was used for the performance assessment and to guide the search. Here we extend this research and propose two major advancements: Firstly we formulate the algorithm in detection error tradeoff space, minimizing false positives and false negatives, with the advantage that misclassification cost tradeoff can be assessed directly. Secondly, we add complexity as an objective function, which gives rise to a 3D objective space (as opposed to a 2D previous ROC space). A domain specific performance indicator for 3D Pareto front approximations, the volume above DET surface, is introduced, and used to guide the indicator-based evolutionary algorithm to find optimal approximation sets. We assess the performance of the new algorithm on designed theoretical problems with different geometries of Pareto fronts and DET surfaces, and two application-oriented benchmarks: (1) Designing spam filters with low numbers of false rejects, false accepts, and low computational cost using rule ensembles, and (2) finding sparse neural networks for binary classification of test data from the UCI machine learning benchmark. The results show a high performance of the new algorithm as compared to conventional methods for multicriteria optimization.

© 2016 Elsevier Inc. All rights reserved.

* Corresponding author. Tel.: +86 18629588969.

E-mail address: jiaqizhao88@126.com (J. Zhao).

1. Introduction

Classification is one of the most common problems in machine learning. The task of classification is to assign instances in a dataset to target classes based on previously trained classifiers. The ROC (Receiver Operating Characteristic) curve is a technique for visualizing, organizing and selecting binary classifiers based on their performance [24]. ROC curves are typically used to evaluate and compare the performance of classifiers and they also have properties that make them especially useful for domains with skewed class distributions and different classes of problems that assign costs to misclassification. Originating from the field of object classification in radar images, ROC analysis has become increasingly important in many other areas with cost sensitive classification [15] and/or unbalanced data distribution [49], such as medical decision making [50], signal detection [20] and diagnostic systems [52]. As opposed to ROC curves, which show the tradeoff between true positive rate and false positive rate, DET (Detection Error Tradeoff) curves [41] show tradeoffs between false positive and false negative error rates. With DET it is easier to visualize the tradeoff between misclassification cost for binary classifiers than with ROC curves.

More recently, research has drawn attention to ROC convex hull (ROCCH) analysis that covers potentially optimal points for a given set of classifiers [24]. ROCCH makes use of the finding that two hard classifiers can be combined into a classifier that has characteristics in ROC space that correspond to linear combinations of the characteristics of single classifiers and thus, when searching for an approximation to the Pareto front, these linear combinations do not have to be explicitly represented in ROC space. A performance indicator for sets of hard binary classifiers that is compliant with the improvement of ROCCH is the area under the convex hull (AUC). And likewise the area above the DET convex hull can serve as an indicator of how well a Pareto front has been approximated. It measures the area attained by the current Pareto front approximation in DET space.

Some evolutionary multiobjective optimization algorithms (EMOAs) [31,32,54,57,58] have been applied to machine learning [2,33] and image processing areas [36,39]. One of the first algorithms where EMOAs were used for ROC optimization was proposed in [35]. Here a niched Pareto multiobjective genetic algorithm was used for classifier optimization by optimizing biobjective ROC curve. The generalization improvement in multiobjective learning was discussed in [27], where the generation of binary neural network classifiers based on the ROC analysis using an evolutionary multiobjective optimization algorithm was presented. It showed that the generalization ability can be more efficiently improved with a multiobjective framework than within a single objective one. ROC for multiclass classification was analyzed in [22], where a multiobjective optimization algorithm was used for classifiers training based on multiclass ROC analysis. The ROC front concept was introduced as an alternative to the ROC curve representation in [13], and the strategy was applied to the selection of classifiers in a pool using a multiobjective optimization algorithm. Moreover, the maximization of the performance of ROC representations with respect to this indicator has been subject to a recent study by Wang et al. [55], who showed that the proposed algorithm, convex-hull-based multiobjective genetic programming algorithm (CH-MOGP), is capable of showing a strong performance for improving ROCCH with respect to AUC as compared to using state-of-the-art EMOAs for the same task, such as NSGA-II (Nondominated Sorting Genetic Algorithm II) [16], GDE3 (the third evolution step of Generalized Differential Evolution) [34], SPEA2 (Strength Pareto Evolutionary Algorithm 2) [63], MOEA/D (Multiobjective Evolutionary Algorithm based on Decomposition) [61], and SMS-EMOA (multiobjective selection based on dominated hypervolume) [7].

So far algorithms that seek to maximize ROCCH performance have only focused on the problem of optimizing binary classifiers with respect to two criteria, i.e., minimization of false positive rate (fpr) and maximization of true positive rate (tpr). There is however an increasing interest in extending ROCCH performance analysis to more than two criteria. In this research we consider the complexity as an additional objective. The objective here is to find models with maximum simplicity (parsimony) or minimum computational costs. For rule-based systems, it can be described as the number of rules defining a classifier in proportion to the number of all possible rules. As it is easier to see the tradeoff between misclassification costs (i.e., fpr and fnr) when using DET space than when using ROC space, we use DET curve to describe the performance of binary classifiers.

In the past, convex-hull-based selection operators were employed in EMOA to maintain a well-distributed set or make the non-dominated sorting more effective (cf. [30,43]). In [14] a multiobjective evolutionary algorithm based on the properties of the convex hulls defined in the ROC space was proposed. It was applied to determine a set of fuzzy rule-based binary classifiers with different tradeoffs between false positive rate (fpr) and true positive rate (tpr). NSGA-II was used to generate an approximation of a Pareto front composed of genetic fuzzy classifiers with different tradeoffs among sensitivity, specificity, and interpretability in [17]. After projecting the overall Pareto front onto the ROC space, ROC convex hull method was used to determine the potentially optimal classifiers on the ROC plane.

In this paper, we add the complexity minimization for parsimony maximization as a third objective function and formulate the problem from the misclassification error optimization point of view by minimizing false positive and false negative error rates objectives. For this we model the problem as a triobjective optimization in augmented DET space, and we propose a 3D convex-hull-based evolutionary multiobjective algorithm (3DCH-EMOA) that takes into account domain specific properties of the 3D augmented DET space. Moreover, we analyze and assess the performance of the algorithm in different studies on, partly new, academic problems and practical applications. To analyze the capability of different algorithms to maximize convex hull volume, in a more fundamental study, a set of test problems named ZEJD (Zhao, Emmerich, Jiao, Deutz) [21] are designed and the capability of 3DCH-EMOA to capture only the convex part of a Pareto front is assessed. Besides, we include a study on spam filter design, in which the number of rules determines the complexity objective in

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات