

EA Multi-Model Selection for SVM

Gilles Lebrun

University of Caen Basse-Normandie, France

Olivier Lezoray

University of Caen Basse-Normandie, France

Christophe Charrier

University of Caen Basse-Normandie, France

Hubert Cardot

University François-Rabelais of Tours, France

INTRODUCTION

Evolutionary algorithms (EA) (Rechenberg, 1965) belong to a family of stochastic search algorithms inspired by natural evolution. In the last years, EA were used successfully to produce efficient solutions for a great number of hard optimization problems (Beasley, 1997). These algorithms operate on a population of potential solutions and apply a survival principle according to a **fitness** measure associated to each solution to produce better approximations of the optimal solution. At each iteration, a new set of solutions is created by selecting individuals according to their level of fitness and by applying to them several operators. These operators model natural processes, such as selection, recombination, mutation, migration, locality and neighborhood. Although the basic idea of EA is straightforward, solutions coding, size of population, fitness function and operators must be defined in compliance with the kind of problem to optimize.

Multi-class problems with binary SVM (Support Vector Machine) classifiers are commonly treated as a decomposition in several binary sub-problems. An open question is how to properly choose all models for these sub-problems in order to have the lowest error rate for a specific SVM multi-class scheme. In this paper, we propose a new approach to optimize the **generalization capacity** of such SVM multi-class schemes. This approach consists in a global selection of models for sub-problems altogether and is denoted as **multi-model selection**. A multi-model selection can outperform the classical individual model selection used until now in the literature, but this type of selection defines a hard optimisation problem, because it corresponds to a search

a efficient solution into a huge space. Therefore, we propose an adapted EA to achieve **that multi-model selection** by defining specific **fitness** function and **recombination operator**.

BACKGROUND

The multi-class classification problem refers to assigning a class to a feature vector in a set of possible ones. Among all the possible inducers, Support Vector Machine (SVM) have particular high generalization abilities (Vapnik, 1998) and have become very popular in the last few years. However, SVM are binary classifiers and several **combination schemes** were developed to extend SVM for problems with more two classes (Rifkin & Klautau, 2005). These schemes are based on different principles: probabilities (Price, Knerr, Personnaz & Dreyfus, 1994), error correcting codes (Dietterich, & Bakiri, 1995), correcting classifiers (Moreira, & Mayoraz, 1998) and evidence theory (Quost, Denoeux & Masson, 2006). All these **combination schemes** involve the following three steps: 1) decomposition of a multi-class problem into several binary sub-problems, 2) SVM training on all sub-problems to produce the corresponding binary decision functions and 3) **decoding strategy** to take a final decision from all binary decisions. Difficulties rely on the choice of the **combination scheme** (Duan & Keerthi, 2005) and how to optimize it (Lebrun, Charrier, Lezoray & Cardot, 2005).

In this paper, we focus on step 2) when steps 1) and 3) are fixed. For that step, each binary problem needs to properly tune the SVM hyper-parameters (model) in

order to have a global low multi-class error rate with the combination of all binary decision functions involved in. The search for efficient values of hyper-parameters is commonly designed by the term of *model selection*. The classical way to achieve optimization of multi-class schemes is an individual model selection for each related binary sub-problem. This methodology overtones that a multi-class scheme based on SVM combination is optimal when each binary classifier involved in that **combination scheme** is optimal on the dedicated binary problem. But, if it is supposed that a **decoding strategy** can more or less easily correct binary classifiers errors, then individual binary model selection on each binary sub-problem cannot take into account error correcting possibilities. For this main reason, we are thinking that another way to achieve optimization of multi-class schemes is a global **multi-model selection** for binary problems altogether. In fact, the goal is to have a minimum of errors on a multi-class problem. The selection of all sub-problem models (multi-model selection) has to be globally performed to achieve that goal, even if that means that error rates are not optimal on all binary sub-problems when they are observed individually. EA is an efficient meta-heuristic approach to realize that **multi-model selection**.

EA MULTI-MODEL SELECTION

This section is decomposed in 3 subsections. In the first section, the multi-model optimization problem for multi-class **combination schemes** is exposed. More details than in previous section and useful notations for next subsections are introduced. In the second section, our EA multi-model selection is exposed. Details on fitness estimation of multi-model and crossover operator over them are described. In the third section, experimental protocol and results with our EA multi-model selection are provided.

Multi-Model Optimization Problem

A multi-class combination scheme induces several binary sub-problems. The number k and the nature of binary sub-problems depend on the decomposition involved in the combination scheme. For each binary sub-problem, a **SVM** must be trained to produce an appropriate binary decision function h_i ($1 < i < k$). The quality of h_i is greatly dependent on the selected model

θ_i and is characterized by the expected error rate e_i for new datasets with the same binary decomposition. Each model θ_i contains all hyper-parameters values for training a SVM on dedicated binary sub-problem. Expected error rate e_i associated to a model θ_i is commonly determined by cross-validation techniques. All the θ_i models constitute the multi-model $\theta = (\theta_1, \dots, \theta_k)$. The expected error rate e of a SVM multi-class combination scheme is directly dependent on the selected multi-model θ . Let Θ denote the multi-model space for a multi-class problem (*i.e.* $\forall \theta : \theta \in \Theta$) and Θ_i the model space for the i^{th} binary sub-problem. The best θ^* multi-model is the one for which expected error e is minimum and corresponds to the following optimization problem:

$$\theta^* = \arg \min_{\theta \in \Theta} e(\theta) \quad (1)$$

where $e(\theta)$ denotes the expected error e of a multi-class combination scheme with the multi-model θ . The huge size of the multi-model space ($\Theta = \times_{i \in [1, k]} \Theta_i$) makes the optimization problem (0.1) very hard. To reduce the optimization problem complexity, it is classic to use the following approximation:

$$\tilde{\theta} = \{\arg \min_{\theta \in \Theta} e(\theta_i) \mid i \in [1, k]\} \quad (2)$$

Hypothesis is made that

$$e(\tilde{\theta}) \approx e(\theta^*).$$

This hypothesis also supposes that

$$e(\tilde{\theta}_i) \approx e(\theta_i^*).$$

If it is evident that each individual model θ_i in the best multi-model θ^* must correspond to efficient **SVM** (*i.e.* low value of e_i) on the corresponding i^{th} binary sub-problem, all best individual models ($\theta_1^*, \dots, \theta_k^*$) do not necessarily define the best multi-model θ^* . The first reason is that all error rates e_i are estimated with some tolerance and combination of all these deviations can have a great impact on the final multi-class error rate e . The second reason is that even if all the binary classifiers of a **combination scheme** have identical e_i error rates for different multi-models, these binary classifiers can have different binary class predictions for a same

4 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/multi-model-selection-svm/10296

Related Content

Wireless Sensor Node Placement Using Hybrid Genetic Programming and Genetic Algorithms

Arpit Tripathi, Pulkit Gupta, Aditya Trivedi and Rahul Kala (2011). *International Journal of Intelligent Information Technologies* (pp. 63-83).

www.irma-international.org/article/wireless-sensor-node-placement-using/54067

Modelling of Cloud Computing Enablers Using MICMAC Analysis and TISM

Nitin Chawla and Deepak Kumar (2018). *International Journal of Ambient Computing and Intelligence* (pp. 31-43).

www.irma-international.org/article/modelling-of-cloud-computing-enablers-using-micmac-analysis-and-tism/204347

An Introduction to the Basic Concepts in QSAR-Aided Drug Design

Maryam Hamzeh-Mivehroud, Babak Sokouti and Siavoush Dastmalchi (2017). *Artificial Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 32-78).

www.irma-international.org/chapter/an-introduction-to-the-basic-concepts-in-qsar-aided-drug-design/173330

How Artificial Intelligence Can Help Accounting in Information Management

Albérico Travassos Rosário (2024). *Artificial Intelligence Approaches to Sustainable Accounting* (pp. 65-92).

www.irma-international.org/chapter/how-artificial-intelligence-can-help-accounting-in-information-management/343354

A Brownian Agent Model for Analyzing Changes in a Nation's Product Space Structure

Bin Jiang, Chao Yang, Takashi Yamada and Takao Terano (2015). *International Journal of Intelligent Information Technologies* (pp. 52-71).

www.irma-international.org/article/a-brownian-agent-model-for-analyzing-changes-in-a-nations-product-space-structure/128839