An efficient and speeded-up tree for multi-class classification

Ranganathan, P.[a] , Ramanan, A.[a] and Niranjan, M.[b]

[a] Department of Computer Science, University of Jaffna, Sri Lanka
[b] School of Electronics and Computer Science, University of Southampton, United Kingdom

Abstract

Support vector machine is a state-of-the-art learning machine that is used in areas, such as pattern recognition, computer vision, data mining and bioinformatics. SVMs were originally developed for solving binary classification problems, but binary SVMs have also been extended to solve the problem of multi-class pattern classification. There are different techniques employed by SVMs to tackle multi-class problems, namely oneversus-one (OVO), one-versus-all (OVA), and directed acyclic graph (DAG). When dealing with multi-class classification, one needs an appropriate technique to effectively extend these binary classification methods for multi-class classification. We address this issue by extending a novel architecture that we refer to as unbalanced decision tree (UDT). UDT is a binary decision tree arranged in a top-down manner, using the optimal margin classifier at each split to relieve the excessive time in classifying the test data when compared with the DAG-SVMs. The initial version of the UDT required a longer training time in finding the optimal model for each decision node of the tree. In this work, we have drastically reduced the excessive training time by finding the order of classifiers based on their performances during the selection of the root node and fix this order to form the hierarchy of the decision tree. UDT involves fewer classifiers than OVO, OVA and DAG-SVMs, while maintaining accuracy comparable to those standard techniques.

Author keywords

Decision Trees; Multi-class Classification; Support Vector Machine

Indexed keywords

Appropriate techniques; Binary classification methods; Binary classification problems; Binary decision trees; Directed acyclic graph (DAG); Learning machines; Margin classifiers; Multi-class classification; Multi-class pattern classifications; Multi-class problems; Novel architecture; Optimal model; Root nodes; Test data; Top-down manner; Training time

Engineering controlled terms: Bioinformatics; Classifiers; Data mining; Decision trees; Pattern recognition; Problem solving; Support vector machines; Sustainable development

Engineering main heading: Binary trees