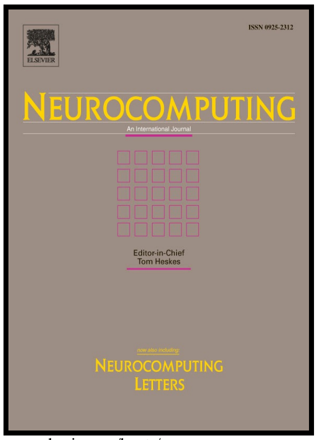
Author's Accepted Manuscript



www.elsevier.com/locate/neucom

Monotonic classification extreme learning machine

Hong Zhu, Eric C.C. Tsang, Xi-Zhao Wang, Rana Aamir Raza Ashfaq

PII: S0925-2312(16)31402-3
DOI: http://dx.doi.org/10.1016/j.neucom.2016.11.021
Reference: NEUCOM17768

To appear in: *Neurocomputing*

Received date: 21 July 2016
Revised date: 21 October 2016
Accepted date: 4 November 2016

Cite this article as: Hong Zhu, Eric C.C. Tsang, Xi-Zhao Wang and Rana Aami Raza Ashfaq, Monotonic classification extreme learning machine *Neurocomputing,* http://dx.doi.org/10.1016/j.neucom.2016.11.021

Monotonic Classification Extreme Learning Machine

Hong Zhu[a], Eric C.C. Tsang[a,*], Xi-Zhao Wang[b], Rana Aamir Raza Ashfaq[b]

[a] *Faculty of Information Technology, Macau University of Science and Technology, Macau*
[b] *College of Computer Science & Software Engineering, Shenzhen University, Shenzhen 518060, China*

Abstract

Monotonic classification problems mean that both feature values and class labels are ordered and monotonicity relationships exist between some features and the decision label. Extreme Learning Machine (ELM) is a single-hidden layer feedforward neural network with fast training rate and good generalization capability, but due to the existence of training error, ELM cannot be directly used to handle monotonic classification problems. This work proposes a generalization of ELM for processing the monotonic classification, named as Monotonic Classification Extreme Learning Machine (MCELM) in which the monotonicity constraints are imposed to the original ELM model. Mathematically, MCELM is a quadratic programming problem in which the monotonicity relationships are considered as constraints and the training error is the objective to be minimized. The mathematical model of MCELM not only can make the generated classifier monotonic but also can minimize the classification error. MCELM does not need to tune parameters iteratively, and therefore, keeps the advantage of extremely fast training which is the essential characteristic of ELM. MCELM does not require that the monotonic relationships existing between features and the output are consistent, which essentially relaxes the assumption of consistent monotonicity used in most existing approaches to handling monotonic classification problems. In comparison with exiting approaches to handling monotonic classification, MCELM can indeed generate a monotonicity-reserving classifier which experimentally shows a much better generalization capability on both artificial and real world datasets.

Keywords: Monotonic classification, Extreme learning machine, Constrained extreme learning machine, Monotonicity, Quadratic programming

1. Introduction

As a fundamental task of supervised learning, classification is to get a classifier by training a number of labeled samples, and then to predict the class label of an unseen sample based on the trained classifier. From references we can find many different algorithms proposed to solve classification problems, such as decision tree induction [1], Bayesian classifier [2], kernel methods [3], support vector machine [4], artificial neuron networks [5], etc.

Ordinal classification is a generalization of the traditional classification. In traditional classification problems the class labels are nominal while in the ordinal classification the class labels are ordered, in other words, there is an order relationship among the class labels in ordinal classifications. For example, the damage degree after a typhoon hit can be classified into three levels: slight, moderate and serious. It is clear that there is an order relationship among these class labels. In comparison with the traditional classification, the ordinal classification sufficiently employs the information of order among the labels.

The monotonic classification is essentially an ordinal classification with monotonicity constraints, in which both feature values and class labels are ordered and monotonic relationships exist between them. The monotonicity can be either increasing or decreasing. If the decision value increases (decreases) with the increasing (decreasing) of a particular feature value, then the monotonicity between the decision attribute and the feature is increasing; otherwise it is decreasing. For example, the grade of scholarship is monotonically increasing with respect to the student's academic performance and the satisfaction degree of a car is monotonically decreasing with respect to the maintenance cost. Monotonic classification problems widely exist in many real application fields such as disease diagnoses, bankruptcy risk assessments and employee selections.

In recent decades, the monotonic classification has attracted lots of attention from researchers. The following is an incomplete survey on monotonic classifications.

In 1989, Ben David et al. introduced the first algorithm OLM [6] for ordinal classification with monotonic constraints in the machine learning community. This method consists of two components. It first chooses a subset of training objects and then classifies the samples of the subset by using a function. Unfortunately this approach may produce non-monotonic results. Greco et al. proposed a dominance-based rough set approach [7] to deal with the multi-criteria sorting, which is an extension of the classical rough set approach [8]. The approach handles inconsistencies coming from violation of the monotonicity constraints by substituting the indiscernibility relation with the dominance relation. It is the first approach using the comprehensive theory for monotonic classification tasks in the domain of knowledge discovery. Further in [9] and [10], the model of dominance rough set was used to extract rules for monotonic classification. Similar pieces of work can be found from many other researchers' studies, such as [11]-[16]. Among the existing approaches to learning with monotonicity constraints, the nonparametric approach is the most general one. In [17], the authors introduced a probabilistic model for ordinal classifications with monotonicity constraints based on the concept of stochastic dominance and

*Corresponding authors.
Email addresses: xszhuhong@163.com (Hong Zhu), cctsang@must.edu.mo (Eric C.C. Tsang), xizhaowang@ieee.org (Xi-Zhao Wang), aamir@szu.edu.cn (Rana Aamir Raza Ashfaq)