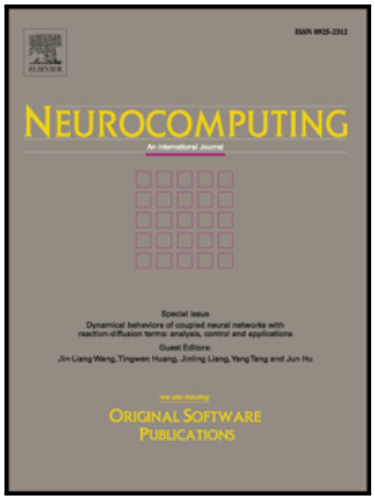
Communicated by Dr Yiming Ying

Accepted Manuscript



Universal Consistency and Robustness of Localized Support Vector Machines

Florian Dumpert, Andreas Christmann

PII: S0925-2312(18)30799-9
DOI: 10.1016/j.neucom.2018.06.061
Reference: NEUCOM 19739

To appear in: *Neurocomputing*

Received date: 13 October 2017
Revised date: 9 May 2018
Accepted date: 28 June 2018

Please cite this article as: Florian Dumpert, Andreas Christmann, Universal Consistency and Robustness of Localized Support Vector Machines, *Neurocomputing* (2018), doi: 10.1016/j.neucom.2018.06.061

Universal Consistency and Robustness of Localized Support Vector Machines

Florian Dumpert[*], **Andreas Christmann**
Department of Mathematics, University of Bayreuth, Germany

Abstract

The massive amount of available data potentially used to discover patterns in machine learning is a challenge for kernel based algorithms with respect to runtime and storage capacities. Local approaches might help to relieve these issues. From a statistical point of view local approaches allow additionally to deal with different structures in the data in different ways. This paper analyses properties of localized kernel based, non-parametric statistical machine learning methods, in particular of support vector machines (SVMs) and methods close to them. We will show there that locally learnt kernel methods are universally consistent. Furthermore, we give an upper bound for the maxbias in order to show statistical robustness of the proposed method.

Key words and phrases. Machine learning; universal consistency; robustness; localized learning; reproducing kernel Hilbert space.

1 Introduction

This paper analyzes properties of localized kernel based, non-parametric statistical machine learning methods, in particular of support vector machines (SVMs) and methods close to them. Caused by the enormous research activities there is abundance of general introductions to this field of computer science and statistics. Beside many publications in international journals there are summarizing textbooks like for example Cristianini & Shawe-Taylor (2000), Schölkopf & Smola (2001), Steinwart & Christmann (2008) or Cucker & Zhou (2007) from a mathematical or statistical point of view. Nevertheless, we want to give a short overview over the analyzed topic.

Support vector machines were initially introduced by Boser, Guyon & Vapnik (1992) and Cortes & Vapnik (1995), based on earlier work like the Russian original of Vapnik, Chervonenkis & Červonenkis (1979). The basic ideas are presented in a comprehensive way in Vapnik (1995, 2000) and Vapnik (1998). An early discussion is provided by Bennett & Campbell (2000). Although SVMs and related kernel based methods are much more recent then other very well-established statistical techniques like for example ordinary least squares regression or their related generalized linear models for regression and classification, they became pretty popular in many fields of science, see for example Ma & Guo (2014). The analysis provided by this paper usually refers to classification or regression problems and therefore to so-called supervised learning. Beyond this, support vector machines are a suitable method for unsupervised learning (e. g. novelty detection), too.

The paper is organized as follows: Section 2.1 gives an overview on support vector machines, Section 2.2 introduces the idea of local approaches. The consistency and robustness results are provided in Section 3, the proofs can be found in the appendix. Section 4 summarizes the paper.

[*]corresponding author