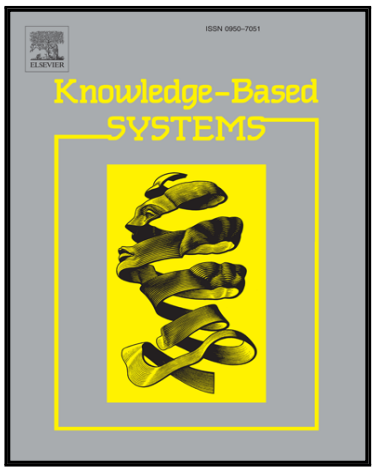

Sparse Online Feature Maps

Nima Salehi-M., Reza Monsefi, Hadi Sadoghi Yazdi

PII: S0950-7051(18)30155-2
DOI: 10.1016/j.knosys.2018.03.023
Reference: KNOSYS 4272

To appear in: *Knowledge-Based Systems*

Received date: 12 September 2017
Revised date: 13 March 2018
Accepted date: 15 March 2018

Please cite this article as: Nima Salehi-M., Reza Monsefi, Hadi Sadoghi Yazdi, Sparse Online Feature Maps, *Knowledge-Based Systems* (2018), doi: 10.1016/j.knosys.2018.03.023

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

00 (2018) 1–32

Knowledge-based system

Sparse Online Feature Maps

Nima Salehi-M.[1], Reza Monsefi[1,*], Hadi Sadoghi Yazdi[1]

Computer Department, Engineering Faculty, Ferdowsi University of Mashhad (FUM), Mashhad, IRAN

Abstract

Online kernel methods suffer from computational and memory complexity in large-scale problems. Due to these drawbacks, budget online kernel learning and kernel approximation (low-dimensional feature map approximation) methods are widely used to speed up time and to reduce memory usage of kernel approaches. In this paper, orthogonal Gram-Schmidt explicit feature maps are applied to online kernel methods. The main advantage of these feature maps come from their orthogonality property. Utilization of these feature maps leads to mutually linearly independent dimensions of feature space, hence, reduce the redundancy in this space. These feature maps can be applied to single-pass online learning methods with l_2- and l_0-norm regularization to reduce the computational and memory complexity. In this paper, the proposed methods are named: 1) Online Feature Maps (OFEMs) and 2) Sparse Online Feature Maps (SOFEMs). These methods are examined for binary and multiclass single-label classification problems. Extensive experiments are compared with the results of other state-of-the-art methods on standard and real-world datasets. The experimental results show that OFEMs and SOFEMs outperform other methods in the literature.

Keywords: Explicit Feature Map, Kernel Methods, Single-pass Online Learning, Gram-Schmidt Orthogonalization Process

1. Introduction

Batch kernel methods such as Support Vector Machines (SVMs) are widely used in machine learning. In the literature, before SVM and its variants (Scholkopf and Smola, 2001; Shawe-Taylor and

*Corresponding author

Email addresses: ni.salehi@mail.um.ac.ir (Nima Salehi-M.), monsefi@um.ac.ir (Reza Monsefi), h-sadoghi@um.ac.ir (Hadi Sadoghi Yazdi)