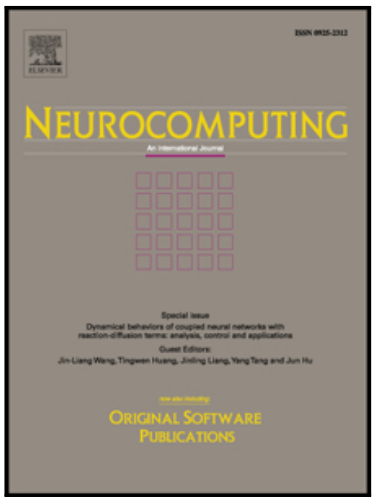

Graph classfication based on sparse graph feature selection and extreme learning machine

Yajun Yu , Zhisong Pan , Guyu Hu , Huifeng Ren

PII: S0925-2312(17)30200-X
DOI: 10.1016/j.neucom.2016.03.110
Reference: NEUCOM 17999

To appear in: *Neurocomputing*

Received date: 30 September 2015
Revised date: 12 March 2016
Accepted date: 15 March 2016

Please cite this article as: Yajun Yu , Zhisong Pan , Guyu Hu , Huifeng Ren , Graph classfication based on sparse graph feature selection and extreme learning machine, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2016.03.110

Graph classfication based on sparse graph feature selection and extreme learning machine

Yajun Yu, Zhisong Pan, Guyu Hu, Huifeng Ren

College of Command Information System, PLA University of Science and Technology, Nanjing, China 210007

Abstract

Identification and classification of graph data is a hot research issue in pattern recognition. The conventional methods of graph classification usually convert the graph data to the vector representation and then using SVM to be a classifier. These methods ignore the sparsity of graph data, and with the increase of the input sample, the storage and computation of the kernel matrix will cost a lot of memory and time. In this paper, we propose a new graph classification algorithm called graph classification based on sparse graph feature selection and extreme learning machine. The key of our method is using the lasso to select features because of the sparsity of graph data, and extreme learning machine(ELM) is introduced to the following classification task due to its good performance. Extensive experimental results on a series of benchmark graph data sets validate the effectiveness of the proposed methods.

Keywords: graph kernel, graph classification, extreme learning machine, lasso

1. Introduction

Most of the existing machine learning algorithms such as support vector machine(SVM) are only applied to deal with vector type data. But in many practical applications such as bioinformatics, drug discovery, web data mining and social networks involves the study of relationships between structured objects[1], which can not be represented in vector forms. In recent years, the research about structured data has become a hot research issue in machine learning and data mining. Graphs are usually employed to represent the structured objects, and the nodes of the graph represent objects while the edges indicate the relationships between objects. To analyse graph data, the most important thing is the similarity measurement of two graphs. In order to address this issue, several methods are proposed such as graph edit distance (GED)[2] and graph kernel. Kernel method is a good way to study graph data, and kernel function can be used to measure the similarity between two graphs.

The general method of graph classification is using the graph kernel to map the graph data into higher dimensional vector described feature space, and computing the kernel matrix K consisting of each similarity of two graphs, then the original graph data set can be classified by SVM[3]. However, on one hand, this method neglects the