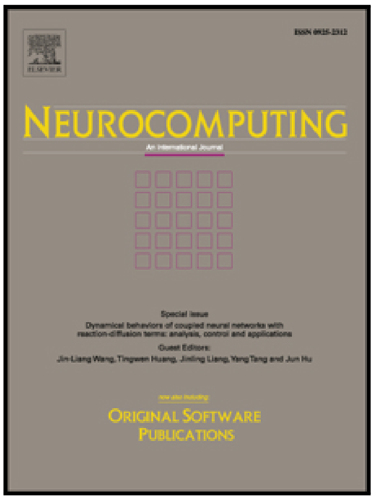
Accepted Manuscript



Empirical Kernel Map-Based Multilayer Extreme Learning Machines for Representation Learning

Chi-Man Vong , Chuangquan Chen , Pak-Kin Wong

PII: S0925-2312(18)30584-8
DOI: 10.1016/j.neucom.2018.05.032
Reference: NEUCOM 19585

To appear in: *Neurocomputing*

Received date: 22 January 2018
Revised date: 8 May 2018
Accepted date: 9 May 2018

Empirical Kernel Map-Based Multilayer Extreme Learning Machines for Representation Learning

Chi-Man Vong[1], Chuangquan Chen[1], Pak-Kin Wong[2]

[1]Department of Computer and Information Science, University of Macau, Macau
[2]Department of Electromechanical Engineering, University of Macau, Macau

Abstract

Recently, multilayer extreme learning machine (ML-ELM) and hierarchical extreme learning machine (H-ELM) were developed for representation learning whose training time can be reduced from hours to seconds compared to traditional stacked autoencoder (SAE). However, there are three practical issues in ML-ELM and H-ELM: 1) the random projection in every layer leads to unstable and suboptimal performance; 2) the manual tuning of the number of hidden nodes in every layer is time-consuming; and 3) under large hidden layer, the training time becomes relatively slow and a large storage is necessary. More recently, issues (1) and (2) have been resolved by kernel method, namely, multilayer kernel ELM (ML-KELM), which encodes the hidden layer in form of a kernel matrix (computed by using kernel function on the input data), but the storage and computation issues for kernel matrix pose a big challenge in large-scale application. In this paper, we empirically show that these issues can be alleviated by encoding the hidden layer in the form of an approximate *empirical kernel map* (EKM) computed from low-rank approximation of the kernel matrix. This proposed method is called ML-EKM-ELM, whose contributions are: 1) stable and better performance is achieved under no random projection mechanism; 2) the exhaustive manual tuning on the number of hidden nodes in every layer is eliminated; 3) EKM is scalable and produces a much smaller hidden layer for fast training and low memory storage, thereby suitable for large-scale problems. Experimental results on benchmark datasets demonstrated the effectiveness of the proposed ML-EKM-ELM. As an illustrative example, on the *NORB* dataset, ML-EKM-ELM can be respectively up to 16 times and 37 times faster than ML-KELM for training and testing with a little loss of accuracy of 0.35%, while the memory storage can be reduced up to 1/9.

Index Terms: Kernel learning, Multilayer extreme learning machine (ML-ELM), Empirical kernel map (EKM), Representation learning, stacked autoencoder (SAE).

1. Introduction

Autoencoder (AE) is an unsupervised neural network whose input layer is equal to

✉ Corresponding author: Chi-Man Vong (cmvong@umac.mo)