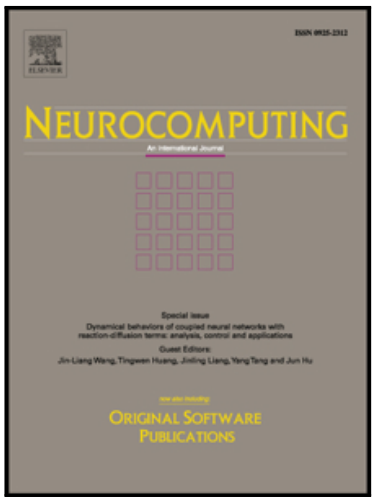

Extreme Learning Machines with Heterogeneous Data Types

Julio J. Valdés

PII: S0925-2312(17)31409-1
DOI: 10.1016/j.neucom.2017.02.103
Reference: NEUCOM 18799

To appear in: *Neurocomputing*

Received date: 30 May 2016
Revised date: 14 January 2017
Accepted date: 6 February 2017

Please cite this article as: Julio J. Valdés, Extreme Learning Machines with Heterogeneous Data Types, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2017.02.103

Extreme Learning Machines with Heterogeneous Data Types

Julio J. Valdés[a]

[a]*National Research Council Canada.*
Information and Communications Technologies.
Data Science for Complex Systems Group.
M50, 1200 Montreal Rd. Ottawa, Ontario K1A 0R6, Canada.
Email: julio.valdes@nrc-cnrc.gc.ca

Abstract

Current advances in communication, sensor and computing technologies are generating information in never before seen amounts and at constantly increasing rates (i.e. the information explosion, the Internet of Things). From the point of view of data analytics, the information is composed of a diversity of data types and it contains uncertainties and incompleteness of different degrees, which add an extra component to the original heterogeneity. Many data mining and machine learning methods do not handle heterogeneity well.

Extreme learning machines (ELM) are interesting computational algorithms because of their simplicity, their good performance and their speed. They can be extended for processing information composed of heterogeneous data types (HT-ELM), capable of addressing classification and regression problems with complex data. Two approaches are discussed: one works directly with the heterogeneous data and the other one transforms the information into simpler homogeneous spaces that preserve structural properties. In them, standard learning methods can be applied, including classical ELMs among others. Both approaches are illustrated using real world examples involving heterogeneous predictor vari-

[☆]This work was supported by the National Research Council Canada (Information and Communications Technologies, project A1-009199).