

Accepted Manuscript

Learning from Small Datasets Containing Nominal Attributes

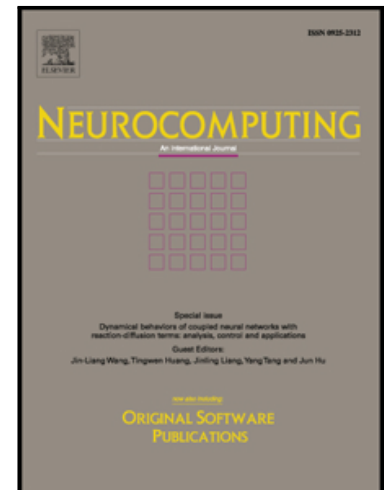
Der-Chiang Li , Hung-Yu Chen , Qi-Shi Shi

PII: S0925-2312(18)30235-2
DOI: [10.1016/j.neucom.2018.02.069](https://doi.org/10.1016/j.neucom.2018.02.069)
Reference: NEUCOM 19373

To appear in: *Neurocomputing*

Received date: 30 June 2017
Revised date: 28 December 2017
Accepted date: 15 February 2018

Please cite this article as: Der-Chiang Li , Hung-Yu Chen , Qi-Shi Shi , Learning from Small Datasets Containing Nominal Attributes, *Neurocomputing* (2018), doi: [10.1016/j.neucom.2018.02.069](https://doi.org/10.1016/j.neucom.2018.02.069)



This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Learning from Small Datasets Containing Nominal Attributes

Der-Chiang Li ^{*}, Hung-Yu Chen, Qi-Shi Shi

Department of Industrial and Information Management, National Cheng Kung

University, University Road, Tainan 70101, Taiwan, R.O.C

E-mail address: *lidc@mail.ncku.edu.tw

*Corresponding author. Tel: (886) 2757575 x53134; Fax: (886) 2374252

Abstract

In many small-data-learning problems, owing to the incomplete data structure, explicit information for decision makers is limited. Although machine learning algorithms are extensively applied to extract knowledge, most of them are developed without considering whether the training sets can fully represent the population properties. Focusing on small data which contains nominal inputs and continuous outputs, this paper develops an effective sample generating procedure based on fuzzy theories to tackle the learning issue by data preprocessing. According to the derived fuzzy relations between categories and continuous outputs, the possibilities of the combinations of categories (virtual samples) can be aggregated when continuous outputs are given. Proper virtual samples are further selected by using fuzzy alpha-cut on the possibility distributions, and these are added to the training sets to form new

Download English Version:

<https://daneshyari.com/en/article/6864218>

Download Persian Version:

<https://daneshyari.com/article/6864218>

[Daneshyari.com](https://daneshyari.com)