Chapter V
Preprocessing Perceptrons and Multivariate Decision Limits

Patrik Eklund
Umeå University, Sweden

Lena Kallin Westin
Umeå University, Sweden

ABSTRACT

Classification networks, consisting of preprocessing layers combined with well-known classification networks, are well suited for medical data analysis. Additionally, by adjusting network complexity to corresponding complexity of data, the parameters in the preprocessing network can, in comparison with networks of higher complexity, be more precisely understood and also effectively utilised as decision limits. Further, a multivariate approach to preprocessing is shown in many cases to increase correctness rates in classification tasks. Handling network complexity in this way thus leads to efficient parameter estimations as well as useful parameter interpretations.

INTRODUCTION

Decision limits, in the sense of discrimination values used for interpretative purposes (Solberg, 1987), are arrived at through traditional bio-statistical approaches to medical data analysis (see e.g. (Armitage & Berry, 1994)). In discrimination (or classification), such as typically used in differential diagnosis, involving several attributes, this analysis should not be performed on univariate level. Attributes cannot be analysed one by one, since synthesization of these decision limits must then be considered separately. Such a discriminator is then like a hybrid, with isolated univariate analysis composed with a classification network, where network parameters are trained independently from the level of univariate decision limits. This would provide a separation of univariate decision limits and discrimination, which

Copyright © 2009, IGI Global, distributing in print or electronic forms without written permission of IGI Global is prohibited.

is unfavourable for correctness rates. Furthermore, efforts to increase correctness rates are typically based on increase of network complexities, leading to other complications such as overtraining and difficulties to interpret network parameters. An analytical presentation of the usefulness to correlate data complexity with network complexity is outside the scope of this chapter. We refer e.g. to (Duda et. al., 2001) for further reading.

In this chapter, we propose to use preprocessing perceptrons which include a preprocessing layer allowing establishing optimal multivariate decision limits. Classification networks coupled with preprocessing can be kept small, and, indeed, with a network complexity in correlation with data complexity. The classification network coupled with preprocessing can be ordinary logistic regression, i.e., a perceptron using neural network terminology. Several other classification networks can be used but it has been shown (Eklund, 1994) difficult to establish significantly higher correctness rates with classifiers of higher complexities.

Hybrid networks, used either in classification or closed-loop problems, are often used in order to increase network complexity to cope with higher data complexity. Typical examples include moving from character to voice recognition or applications of intelligent control. The hybrid described in this chapter is intended to enhance decision limit related preprocessing within the classifier itself. Careful preprocessing with parameter estimations, also for the preprocessing layer, enables the use of classifier networks with low complexities. We support our claims by using some typical case studies in binary classification. Examples are drawn from chromosomal anomalies, endocrinology and virology.

This chapter is organised as follows; First, a description of the proposed preprocessing perceptron is given, including the mathematical formulas behind it. In the next section, three case studies, where the preprocessing perceptron has been used, are described together with the methods used and the results. The results are compared with the results of a multilayer perceptron run on the same data. Next, univariate versus multivariate decision limits are discussed and the final section in this chapter contains a conclusion of the chapter.

THE PREPROCESSING PERCEPTRON

A linear regression is given by the weighted sum:

$$y = \gamma + \sum_{i=1}^{n} w_i x_i$$

where x_i are inputs and w_i and γ are parameters of the linear function. A logistic regression performs a sigmoidal activation of the weighted sum, i.e.:

$$y = \frac{1}{1 + e^{\left(-\gamma + \sum_{i=1}^{n} w_i x_i\right)}}$$

Note that a logistic regression function is precisely a (single-layer) perceptron in the terminology of neural networks (Duda et. al., 2001). The preprocessing perceptron consists similarly of a weighted sum but includes preprocessing functions for each input variable. Suitable preprocessing functions are sigmoids (sigmoidal functions):

$$g[\alpha, \beta](x) = \frac{1}{1 + e^{-\beta(x-\alpha)}}, \ \beta \neq 0 \qquad (1)$$

11 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/preprocessing-perceptrons-multivariate-decision-limits/7529

Related Content

Data Mining Techniques on Earthquake Data: Recent Data Mining Approaches
Negar Sadat Soleimani Zakeriand Saeid Pashazadeh (2015). *Improving Knowledge Discovery through the Integration of Data Mining Techniques (pp. 183-199).*
www.irma-international.org/chapter/data-mining-techniques-on-earthquake-data/134539

Keyword Extraction
(2021). *Developing a Keyword Extractor and Document Classifier: Emerging Research and Opportunities (pp. 119-131).*
www.irma-international.org/chapter/keyword-extraction/268465

Acquiring Semantic Sibling Associations from Web Documents
Marko Brunzeland Myra Spiliopoulou (2007). *International Journal of Data Warehousing and Mining (pp. 83-98).*
www.irma-international.org/article/acquiring-semantic-sibling-associations-web/1795

Efficient Metaheuristic Approaches for Exploration of Online Social Networks
Zorica Stanimiroviand Stefan Miškovi (2016). *Big Data: Concepts, Methodologies, Tools, and Applications (pp. 2098-2148).*
www.irma-international.org/chapter/efficient-metaheuristic-approaches-for-exploration-of-online-social-networks/150258

A Survey of Parallel and Distributed Data Warehouses
Pedro Furtado (2009). *International Journal of Data Warehousing and Mining (pp. 57-77).*
www.irma-international.org/article/survey-parallel-distributed-data-warehouses/1826