



ON-LINE MIXTURE-BASED ALTERNATIVE TO LOGISTIC REGRESSION

I. Nagy*, E. Suzdaleva[†]

Abstract: The paper deals with a problem of modeling discrete variables depending on continuous variables. This problem is known as the logistic regression estimated by numerical methods. The paper approaches the problem via the recursive Bayesian estimation of mixture models with the purpose of exploring a possibility of constructing the continuous data dependent switching model that should be estimated on-line. Here the model of the discrete variable dependent on continuous data is represented as the model of the mixture pointer dependent on data from mixture components via their parameters, which switch according to the activity of the components. On-line estimation of the data dependent pointer model has a great potential for tasks of clustering and classification. The specific subproblems include (i) the model parameter estimation both of the pointer and of the components obtained during the learning phase, and (ii) prediction of the pointer value during the testing phase. These two phases can be joined together in the case of necessity. A real-data experimental comparison with theoretical counterparts shows a competitiveness of the approach in the discussed field.

Key words: *on-line modeling, on-line logistic regression, recursive mixture estimation, data dependent pointer*

Received: January 15, 2016

DOI: 10.14311/NNW.2016.26.024

Revised and accepted: October 3, 2016

1. Introduction

The presented paper deals with a problem of modeling discrete variables depending on continuous variables. This problem is generally known as the logistic regression [16]. Classification based on the logistic regression is widely applied in various fields. Only to enumerate, in medicine applications data observed on a patient (e.g., weight, blood pressure, cholesterol level, sex, age, results of various blood tests, etc.) can be analyzed by the logistic regression to obtain a probability of the certain disease and subsequently to classify the patient's state [2, 8, 9, 20]. In

*Ivan Nagy, Department of Signal Processing, The Institute of Information Theory and Automation of the Czech Academy of Sciences, Pod vodárenskou věží 4, 18208 Prague, Czech Republic, and Faculty of Transportation Sciences, Czech Technical University, Na Florenci 25, 11000 Prague, Czech Republic, E-mail: nagy@utia.cas.cz

[†]Evgenia Suzdaleva – Corresponding author, Department of Signal Processing, The Institute of Information Theory and Automation of the Czech Academy of Sciences, Pod vodárenskou věží 4, 18208 Prague, Czech Republic, E-mail: suzdalev@utia.cas.cz