



COMBINING MISSING DATA IMPUTATION AND PATTERN CLASSIFICATION IN A MULTI-LAYER PERCEPTRON

JOSÉ-LUIS SANCHO-GÓMEZ⁽¹⁾, PEDRO J. GARCÍA-LAENCINA⁽¹⁾ AND ANÍBAL R. FIGUEIRAS-VIDAL⁽²⁾

⁽¹⁾ *Dpto. Tecnologías de la Información y las Comunicaciones. Universidad Politécnica de Cartagena. Plaza del Hospital, 1, 30202. Cartagena (Murcia). SPAIN.*

⁽²⁾ *Dpto. Teoría de Señal y Comunicaciones. Universidad Carlos III de Madrid. Avda. de la Universidad, 30, 28911. Leganés (Madrid). SPAIN.*

ABSTRACT—Multi-Layer Perceptrons (MLPs) have been successfully applied in many pattern classification tasks. However, a drawback of these learning machines is that they cannot handle input vectors that present missing data on its features. A recommended way for dealing with missing values is imputation, i.e., to fill in missing data with plausible values. This paper presents a brief review of handling missing data, including the new Multi-Task Learning (MTL) systems. Moreover, an MLP approach for incomplete pattern classification based on MTL is proposed. This network learns in parallel the classification task (main task) and the different tasks associated to each incomplete feature (secondary tasks). During training, unknown values are imputed, being this missing data imputation process oriented by the learning of the classification task. Experimental results on five classification problems are given to show the effectiveness of the proposed approach.

Key Words: Multi-Layer Perceptron, Statistical Pattern Classification, Missing Data, Imputation, Multi-Task Learning, Artificial Neural Networks.