A Rejection Option for the Multilayer Perceptron Using Hyperplanes

Eduardo Gasca A.¹, Sergio Saldaña T.¹, José S. Sánchez G.², Valentín Velásquez G.¹, Eréndira Rendón L.¹, Itzel M. Abundez B.¹, Rosa M. Valdovinos R.², and Rafael Cruz R.¹

¹Technological Institute of Toluca, State of Mexico, Mexico
²University Jaume I, Castello, Spain
egasca@ittoluca.edu.mx

Abstract. Currently, a growing quantity of the Artificial Intelligence tasks demand a high efficiency of the classification systems (classifiers); making an error in the classification of an object or event can cause serious problems. This is worrying when the classifiers confront tasks where the classes are not linearly separable, the classifiers efficiency diminishes considerably. One solution for decreasing this complication is the Rejection Option. In several circumstances it is advantageous to not have a decision be taken and wait to obtain additional information instead of making an error.

This work contains the description of a novel reject procedure whose purpose is to identify elements with a high risk of being misclassified; like those in an overlap zone. For this, the location of the object in evaluation is calculated with regard to two hyperplanes that emulate the classifiers decision boundary. The area between these hyperplanes is named an overlap region. If the element is localized in this area, it is rejected.

Experiments conducted with the artificial neural network Multilayer Perceptron, trained with the Backpropagation algorithm, show between 12.0%- 91.4% of the objects in question would have been misclassified if they had not been rejected.

Keywords: Reject option, Multilayer Perceptron, Backpropagation, hyperplane, overlap.

1 Introduction

In pattern recognition systems a common task is to categorize an unknown object (pattern) as an element of a class; the class is included in a finite set defined previously. In these systems, the procedure responsible for assigning the class label (classifier) always makes a decision. The classifiers efficiency is characterized mainly by its accuracy in the classification.

One of these classifiers is the Artificial Neural Networks which generates decision boundaries using hyperplanes to separate classes in the observation space. Under this outline, the patterns with a high probability of being incorrectly labeled are usually one of two types: ambiguous data, which creates confusion
Fig. 1. Topology of the Multilayer Perceptron of three layers

among different classes because of their position in an overlap zone; or outliers, these patterns don’t belong to any class included in the initial group.

Currently a growing quantity of real applications require a higher reliabil-
ity level of the classifiers, mainly in tasks where making an error can be very expensive. These applications need systems with a percentage of classification error as low as it is possible which can be impeded by the presence of ambiguous and/or outliers patterns, among others issues. The number of objects in such statuses influences the classifiers efficiency strongly, even when its design is ap-
propriate. One way of decreasing the negative aspects of the ambiguous and/or outliers patterns is implementing the Rejection Option (RO) procedure. The reject concept admits the classifiers inability to formulate a correct decision in the circumstances given. It is preferable to postpone the pattern classification and wait to obtain more information than to take the risk of making a mistake. Here is shown a novel procedure for the RO implantation in an artificial neural network, Multilayer Perceptron, trained with the Backpropagation algorithm.

The content of the following sections is: In section 2, the description of the Multilayer Perceptron; research on the RO is explained in section 3; a novel reject procedure is shown in section 4; section 5 explains the experimental development; and the conclusions are in section 6.

2 Multilayer Perceptron

Figure 1 shows the Multilayer Perceptron (MLP) structure with a hidden layer. The insertion of this layer (it can be one or more) gives it the capacity of con-
fronting tasks where the classes are not linearly separable. Different papers (3, 4 and 5) have mentioned the advantages of using Perceptrons of three layers (TLP) -an input layer, a hidden layer, and an output layer- instead of Perceptron with a larger quantity of layers. This is due to their smaller computational