Dimensionality Reduction Using Rough Set Approach for Two Neural Networks-Based Applications

M. Sammany[1] and T. Medhat [2]

[1] National Water Research Center Ministry of Water Resources and Irrigation, Kornish El-Nile, Imbaba-Giza 12666- Cairo, Egypt
sammanyddr1@gmail.com
[2] Department of Physics and Engineering Mathematics, Faculty of Engineering,Tanta University, 31521, Tanta, Egypt
tmedhatm@yahoo.com

Abstract. In this paper, Rough Sets approach has been used to reduce the number of inputs for two neural networks-based applications that are, diagnosing plant diseases and intrusion detection. After the reduction process, and as a result of decreasing the complexity of the classifiers, the results obtained using Multi-Layer Perceptron (MLP) revealed a great deal of classification accuracy without affecting the classification decisions.

Keywords: Rough Sets, Neural networks, Multi-layered Perceptron (MLP).

1 Introduction

Recently, Artificial Neural Networks (ANN) has been applied successfully to create accurate and efficient models for classification problems [1]. The initial phase of MLP modeling requires selection of input parameter vectors and the corresponding output vectors, which adequately characterize the component to be modeled. Two rules are thus important regarding the input parameter selection. One is that the input parameter vector should be chosen in a manner so that they will weave a domain to cover the model parameter of interest. The other is to select the parameters without redundancy [2].

Dimensionality reduction methods in general try to find a reduced number of new dimensions to account for the original data. Several techniques are available, which can be seen as variants of factors analysis to find a smaller set of representative dimensions. Principal Component Analysis (PCA) [3] is the best known of these techniques: the new dimensions, linear combinations of the original features, are given by the eigenvectors (ordered by decreasing eigenvalue) of the covariance matrix of input data. The new features, called principal components, are uncorrelated and of maximum variance so that the new representation is now minimal. Successive components are of decreasing importance, and the

M. Kryszkiewicz et al. (Eds.): RSEISP 2007, LNAI 4585, pp. 639–647, 2007.