

Towards a Balancing Medical Database by Using the Least Mean Square Algorithm

Authors : Kamel Belammi, Houria Fatrim

Abstract : imbalanced data set, a problem often found in real world application, can cause seriously negative effect on classification performance of machine learning algorithms. There have been many attempts at dealing with classification of imbalanced data sets. In medical diagnosis classification, we often face the imbalanced number of data samples between the classes in which there are not enough samples in rare classes. In this paper, we proposed a learning method based on a cost sensitive extension of Least Mean Square (LMS) algorithm that penalizes errors of different samples with different weight and some rules of thumb to determine those weights. After the balancing phase, we apply the different classifiers (support vector machine (SVM), k- nearest neighbor (KNN) and multilayer neuronal networks (MNN)) for balanced data set. We have also compared the obtained results before and after balancing method.

Keywords : multilayer neural networks, k- nearest neighbor, support vector machine, imbalanced medical data, least mean square algorithm, diabetes

Conference Title : ICCSSP 2015 : International Conference on Circuits, Systems, and Signal Processing

Conference Location : Venice, Italy

Conference Dates : August 13-14, 2015