Murphey, Yi L.; **Guo, Hong**; **Feldkamp, Lee A.**
Neural learning from unbalanced data. (English) Zbl 1075.68075
Appl. Intell. 21, No. 2, 117-128 (2004).

Summary: This paper describes the result of our study on neural learning to solve the classification problems in which data is unbalanced and noisy. We conducted the study on three different neural network architectures, multi-layered Back Propagation, Radial Basis Function, and Fuzzy ARTMAP using three different training methods, duplicating minority class examples, Snowball technique and multidimensional Gaussian modeling of data noise. Three major issues are addressed: neural learning from unbalanced data examples, neural learning from noisy data, and making intentional biased decisions. We argue that by properly generated extra training data examples around the noise densities, we can train a neural network that has a stronger capability of generalization and better control of the classification error of the trained neural network. In particular, we focus on problems that require a neural network to make favorable classification to a particular class such as classifying normal(pass)/abnormal(fail) vehicles in an assembly plant. In addition, we present three methods that quantitatively measure the noise level of a given data set. All experiments were conducted using data examples downloaded directly from test sites of an automobile assembly plant. The experimental results showed that the proposed multidimensional Gaussian noise modeling algorithm was very effective in generating extra data examples that can be used to train a neural network to make favorable decisions for the minority class and to have increased generalization capability.

MSC:

68T05 Learning and adaptive systems in artificial intelligence

Keywords:

machine learning; neural networks; unbalanced data; data noise

Full Text: DOI