A Hybrid Approach of Neural Network and Memory-Based Learning to Data Mining

by: Chung-Kwan Shin, Ui Tak Yun, Huy Kang Kim, and Sang Chan Park
From: IEEE Transactions on Neural Networks, Vol.11, No. 3, MAY 2000.
Review by: Thimaporn Phetkaew

The paper proposed a hybrid prediction system of neural network and memory-based learning. Neural network (NN) and memory-based reasoning (MBR) are frequently applied to data mining with various objectives. They have common advantages over other learning strategies. NN and MBR can be directly applied to classification and regression without additional transformation mechanisms. They also have strength in learning the dynamic behavior of the system over a period of time. Unfortunately, they have shortcomings when applied to data mining tasks. Though the neural network is considered as one of the most powerful and universal predictors, the knowledge representation of NN is unreadable to human, and this “black box” property restricts the application of NN to data mining problems, which require proper explanations for the prediction. On the other hand, MBR suffers from the feature-weighting problem. When MBR measures the distance between cases, some input features should be treated as more important than other features. Feature weighting should be executed prior to prediction in order to provide the information on the feature importance. The k-nearest neighbor (k-NN) method is widely used as the retrieval mechanism. However, the most important assumption of k-NN is that all of the features presented are equally important, which is not true in most data mining problems.

In this hybrid system of NN and MBR, the feature weight set, which is calculated from the trained neural network, plays the core role in connecting both learning strategies, and the explanation for prediction can be given by obtaining and presenting the most similar examples from the case base. Moreover, the proposed system has advantages in the typical data mining problems such as scalability to large datasets, high dimensions, and adaptability to dynamic situations. The neural network provides on-line learning property and it can also be accomplished for MBR by updating the case base with new data and discarding old data. They set up a framework for extracting

feature weights from trained neural networks and defined four feature weighting mechanisms based on the frameworks: sensitivity, activity, saliency, and relevance. A measure of the sensitivity of an input feature is the difference in the prediction value between when the feature is removed and when it is left in place. The activity of a node is measured by the variance of activation level in the training data. The saliency of a weight is measured by estimating the second derivative of the error with respect to the weight. The variance of weights into a node is a good predictor of the node's relevance.

Experimental results show that the hybrid system has a high potential in solving data mining problems by comparing to other experiments in real world datasets: odd parity problem, sinusoidal, Wisconsin Diagnostic Breast Cancer (WDBC), credit, sonar and auto-mpg. They can observe that the error rate of this method is maintained low while the rejection ratio (RR) rises moderately as the number of irrelevant features increases. This method shows robust behavior in the noisy datasets, which has irrelevant features in the input. This implies that our method is particularly useful to large dimensional problems.