Chapter 6

A Confidence-Based RBF Neural Network Ensemble Learning Paradigm with Application to Delinquent Prediction for Credit Risk Management

Lean Yu
Chinese Academy of Sciences, China

Shouyang Wang
Chinese Academy of Sciences, China

ABSTRACT

In this study, a multistage confidence-based radial basis function (RBF) neural network ensemble learning model is proposed to design a reliable delinquent prediction system for credit risk management. In the first stage, a bagging sampling approach is used to generate different training datasets. In the second stage, the RBF neural network models are trained using various training datasets from the previous stage. In the third stage, the trained RBF neural network models are applied to the testing dataset and some prediction results and confidence values can be obtained. In the fourth stage, the confidence values are scaled into a unit interval by logistic transformation. In the final stage, the multiple different RBF neural network models are fused to obtain the final prediction results by means of confidence measure. For illustration purpose, two publicly available credit datasets are used to verify the effectiveness of the proposed confidence-based RBF neural network ensemble learning paradigm.

1. INTRODUCTION

Neural network ensemble learning has been turned out to be an efficient strategy for achieving high performance, especially in fields where the development of a powerful single learning system requires considerable efforts (Lai et al., 2006, Yu et al., 2007, 2008b). Usually, neural network ensemble learning model outperforms the individual neural network models, whose performance is limited by the imperfection of feature extraction, learn-
A Confidence-Based RBF Neural Network Ensemble Learning Paradigm

ing algorithms, and the inadequacy of training data. Due to these reasons, there is an increasing research stream on neural network ensemble learning (Hansen and Salamon, 1990; Perrone and Cooper, 1993; Krogh and Vedelsby, 1995; Rosen, 1996; Tumer and Ghosh, 1996; Yang and Browne, 2004; Lai et al., 2006; Yu et al., 2005, 2007, 2008a, 2008b). For instance, performance improvement can result from training the individual networks to be decorrelated with each other with respect to their errors (Krogh and Vedelsby, 1995). To achieve high prediction performance, there are some essential requirements to the ensemble members and the ensemble strategy. First of all, a basic condition is that the individual neural network models must have enough training data. Secondly, the ensemble members are diverse or complementary, i.e., learners could capture different patterns. Thirdly, to obtain high performance, a wise ensemble strategy is also required based on a set of complementary learners.

For a limited data set, some sampling approaches, such as bagging (Breiman, 1996) have been used for creating different samples by varying the data subsets selected or perturbing training sets. Similarly, diverse ensemble members can be obtained by varying the initial conditions or using different training data. In the ensemble model, the most important point is to select an appropriate ensemble strategy. Generally, the variety of ensemble methods can be grouped into three categories according to the level of classifier outputs: abstract level (crisp class), rank level (rank order) and measurement level (class score) (Xu et al., 1992; Suen and Lam, 2000). In the existing studies, many ensemble systems still use empirical heuristics and ad hoc ensemble schemes at the abstract level (Perrone and Cooper, 1993; Krogh and Vedelsby, 1995; Rosen, 1996; Tumer and Ghosh, 1996; Yang and Browne, 2004; Yu et al., 2005; Lai et al., 2006). Typically, majority voting (Yang and Browne, 2004) uses the abstract level of output of ensemble members. An important drawback of this ensemble strategy is that it does not take confidence degree of neural network output into account. Actually, ensemble at the measurement level is advantageous in that the output measurements contain richer information of class measures. In a sense, an appropriate ensemble strategy is more crucial, especially for integrating the learners that output diverse measurements. Furthermore, the intensive investigation of neural network ensemble has not formulated a convincing theoretical foundation and an overall process model yet (Yu et al., 2008b).

In such situations, this chapter proposes a multistage confidence-based radial basis function (RBF) neural network ensemble approach that differs in that the final ensemble strategy is determined the confidence of neural network output at the measurement level. Particularly the RBF neural network (RBFNN) models (Yu et al., 2008a) are used. In this chapter, the proposed RBF neural network ensemble model consists of five stages. In the first stage, a bagging sampling approach is used to generate different training datasets. In the second stage, the RBF neural network model is trained by various training datasets from the previous stage. In the third stage, the trained RBF neural network models are applied to testing datasets and some prediction results and confidence values can be obtained. In the fourth stage, the confidence values are scaled into a unit interval by logistic transformation. In the final stage, the multiple neural network models are integrated to obtain final output by means of confidence measurement. For testing purpose, two publicly available credit dataset are used to verify the effectiveness of the proposed neural network ensemble model.

The motivation of this chapter is to formulate a multistage confidence-based RBF neural network ensemble model for delinquent prediction and compare its performance with other existing approaches. The rest of the chapter is organized as follows. The second section presents a building process of the multistage neural network ensemble model in detail. To illustration and verification
11 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:

www.igi-global.com/chapter/confidence-based-rbf-neural-network/44251?camid=4v1


www.igi-global.com/e-resources/library-recommendation/?id=1

Related Content

Enterprise Information System and Data Mining

www.igi-global.com/article/enterprise-information-system-data-mining/45725?camid=4v1a

Pursuing Supply Chain Integration: Roles of Resources, Competences, Experience, and Industry-type

www.igi-global.com/article/pursuing-supply-chain-integration/169221?camid=4v1a

Evaluation of Pattern Based Customized Approach for Stock Market Trend Prediction With Big Data and Machine Learning Techniques


Data Warehousing Requirements Collection and Definition: Analysis of a Failure

www.igi-global.com/article/data-warehousing-requirements-collection-definition/45727?camid=4v1a