Chapter 19
Support Vector Machine based Hybrid Classifiers and Rule Extraction thereof: Application to Bankruptcy Prediction in Banks

M. A. H. Farquad
Institute for Development and Research in Banking Technology (IDRBT), INDIA University of Hyderabad, INDIA

V. Ravi
Institute for Development and Research in Banking Technology (IDRBT), INDIA

Raju S. Bapi
University of Hyderabad, INDIA

ABSTRACT
Support vector machines (SVMs) have proved to be a good alternative compared to other machine learning techniques specifically for classification problems. However just like artificial neural networks (ANN), SVMs are also black box in nature because of its inability to explain the knowledge learnt in the process of training, which is very crucial in some applications like medical diagnosis, security and bankruptcy prediction etc. In this chapter a novel hybrid approach for fuzzy rule extraction based on SVM is proposed. This approach handles rule-extraction as a learning task, which proceeds in two major steps. In the first step the authors use labeled training patterns to build an SVM model, which in turn yields the support vectors. In the second step extracted support vectors are used as input patterns to fuzzy rule based systems (FRBS) to generate fuzzy “if-then” rules. To study the effectiveness and validity of the extracted fuzzy rules, the hybrid SVM+FRBS is compared with other classification techniques like decision tree (DT), radial basis function network (RBF) and adaptive network based fuzzy inference system. To illustrate the effectiveness of the hybrid developed, the authors applied it to solve a bank bankruptcy prediction problem. The dataset used pertain to Spanish, Turkish and US banks. The quality of the extracted fuzzy rules is evaluated in terms of fidelity, coverage and comprehensibility.

DOI: 10.4018/978-1-60566-766-9.ch019
INTRODUCTION

Support Vector Machines (SVMs) (Vapnik 1995, 1998) and other linear classifiers are popular methods for building hyperplane-based classifiers from data sets and have been shown to have excellent generalization performance in a variety of applications. SVM is based on the statistical learning theory developed by Vapnik (1995) and his team at AT&T Bell Labs, which is a new learning algorithm and can be seen as an alternative training technique for Polynomial, Radial Basis Function and Multi-Layer Perceptron classifiers (Cortes & Vapnik 1995, Edgar et al. 1997). The scheme used by the SVM is that some linear weighted sum of the explanatory variables is lower (or higher) than a pre-specified threshold indicating that the given sample is classified into one class or the other. Even though such a schemes works well, it is completely non-intuitive to human experts in that it does not let us know the knowledge learnt by it during training in simple, comprehensible and transparent way. Therefore, SVM are also treated as “Black Box” models just like ANN.

There are many techniques existing for extracting knowledge embedded in trained neural networks in the form of if-then-else rules (Tickle et al. 1998). The process of converting opaque models into transparent models is often called Rule Extraction. These models are useful for understanding the nature of the problem and interpreting its solution. Using the rules extracted one can certainly understand in a better way, how a prediction is made. Gallant (1988) initiated the work of rule extraction from a neural network that defines the knowledge learnt in the form of if-then rules.

The advantages of rule extraction algorithm:

- Provision of user explanation capability (Gallant 1988). Davis et al. (1977) argues that even limited explanation can positively influence the system acceptance by the user.
- Data exploration and the induction of scientific theories. A learning system might discover salient features in the input data whose importance was not previously recognized (Craven & Shavlik 1994).
- Improves Generalization.

A taxonomy describing the techniques that are used to extract symbolic rules from the neural networks is proposed by Andrew et al. (1995). In general, rule extraction techniques are divided into two major groups i.e. decompositional and pedagogical. Decompositional techniques view the model at its minimum (or finest) level of granularity (at the level of hidden and output units in case of ANN). Rules are first extracted at individual unit level, these subset of rules are then aggregated to form global relationship. Pedagogical techniques extract global relationship between the input and the output directly without analyzing the detailed characteristics of the underlying solution. The third group for rule extraction techniques is eclectic, which combines the advantages of the decompositional and pedagogical approaches.

Earlier work (Andrews et al. 1995), (Towell & Shavlik 1993) and recent work (Kurfess 2000, Darbari 2001) includes rule extraction from neural network. Hayashi (1990) incorporated fuzzy sets with expert systems and proposed a rule extraction algorithm FNES (Fuzzy Neural Expert System). FNES relies on the involvement of an expert at input phase for transforming the input data into required format. This transformation process is then automated in Fuzzy-MLP (Mitra 1994). Later Mitra & Hayashi (2000) presented a survey of extracting fuzzy rules from neural networks. However, not much work has been done to extract rules from SVM.