2006

A semiparametric model approach for financial bankruptcy prediction

Maria H. Kim
University of Wollongong, mhykim@uow.edu.au

Paul Yoo
University of Sydney

Publication Details

Kim, M. H. & Yoo, P. (2006). A semiparametric model approach for financial bankruptcy prediction. Proceedings: International Conference on Engineering of Intelligent Systems (ICEIS' 2006) (pp. 1-6). Piscataway, NJ: IEEE.

A semiparametric model approach for financial bankruptcy prediction

Abstract

In this study, we propose a model that achieves both accurate modeling and sustainable model stability for corporate bankruptcy prediction. This model is to model the given samples accurately as well as to respond adequately to the unknown inputs by employing semiparametric approach where parametric model and nonparametric Neural Networks (NNs) are combined. By exploring the structural relationships within the available sample data, the proposed model is assumed to retain the advantages of both parametric and nonparametric models. The proposed model is compared to pure parametric models such as Multivariate Discriminant Analysis (MDA) and Logistic Regression (LR), and pure nonparametric model such as NNs. Each model predicts the default probability of a company and classifies the company into an appropriate group as either bankrupt or healthy. Experimental results demonstrated that the proposed semiparametric model showed superior performance in terms of model stability and prediction accuracy in bankruptcy prediction.

Keywords

model, semiparametric, bankruptcy, approach, prediction, financial

Disciplines

Business | Social and Behavioral Sciences

Publication Details

Kim, M. H. & Yoo, P. (2006). A semiparametric model approach for financial bankruptcy prediction. Proceedings: International Conference on Engineering of Intelligent Systems (ICEIS' 2006) (pp. 1-6). Piscataway, NJ: IEEE.

A Semiparametric Model Approach to Financial Bankruptcy Prediction

Maria H. Kim and Paul D. Yoo
School of Economics and Business
The University of Sydney
Sydney, NSW, 2006, Australia
hkim2321@usyd.edu.au

***Abstract*-In this study, we propose a model that achieves both accurate modeling and sustainable model stability for corporate bankruptcy prediction. This model is to model the given samples accurately as well as to respond adequately to the unknown inputs by employing semiparametric approach where parametric model and nonparametric Neural Networks (NNs) are combined. By exploring the structural relationships within the available sample data, the proposed model is assumed to retain the advantages of both parametric and nonparametric models. The proposed model is compared to pure parametric models such as Multivariate Discriminant Analysis (MDA) and Logistic Regression (LR), and pure nonparametric model such as NNs. Each model predicts the default probability of a company and classifies the company into an appropriate group as either bankrupt or healthy. Experimental results demonstrated that the proposed semiparametric model showed superior performance in terms of model stability and prediction accuracy in bankruptcy prediction.**

I. INTRODUCTION

A large number of companies declare bankruptcy every year. According to the American Bankruptcy Institute, 1,660,245 businesses filed for bankruptcy in 2003 [3]. Bankruptcy prediction has always been an important issue in the field of finance since it can have significant impact on many financial institutes such as banks, insurance companies and government agencies. A number of techniques and models have been proposed and widely developed to predict bankruptcy to help decision makers such as investors, credit managers, and financial analysts. Accurate prediction can help bank to make better lending decisions. Accurate prediction can also help in estimating a fair value of the interest rate of a loan and in accurately assessing the credit risk of bank loan portfolios [5]. A prediction model can also be used to aid decision makers of financial institutions in evaluating and selecting companies to collaborate with or to invest in. Such decisions have to take account of the opportunity cost and the risk of failure. In addition, reliable prediction can serve as an early warning system for managers and creditors to take corrective action to prevent corporate default [13].

The traditional approach for banks in doing bankruptcy prediction is to produce an internal scoring system or rating, which takes into account various financial ratios [36]. The problem with these ratings is that they tend to be reactive rather than predictive. Thus, there is a need to develop reliable and accurate financial prediction models that can serve as an early warning system. However, there are some generic problems in financial prediction modeling. Firstly, there are many varying factors that make it challenging to model the underlying financial system. Secondly, limited samples of the bankrupt companies make it difficult to extract the common financial characteristics of the bankrupt companies.

Moreover, while many researchers have proposed various models that achieve high accuracy in bankruptcy prediction, they faced the problem that the models with high accuracy tend to show poor generalization. Generalization is the capacity of the model to respond to unknown/unseen inputs that differ from training samples. In order to have a reliable and accurate financial prediction model, we need a model that has high prediction accuracy as well as stable generalization.

Neural Networks (NNs) have been shown to outperform other traditional prediction models, but NNs are still deemed unreliable for real life applications because of their inherent limitations such as overfitting that can cause poor generalization. Overfitting is an inherent problem of the nonparametric model approach such as NNs, which does not specify any assumption about the underlying probabilistic distribution; rather it relies heavily on available sample data only. Lack of generalization is a serious problem in prediction.

This study aims to develop a reliable and accurate bankruptcy prediction model using state-of-the-art machine learning techniques. In this paper, we present a general framework for prediction methods in bankruptcy prediction. We provide a review of historical developments in bankruptcy prediction research in terms of parametric and nonparametric model paradigms. Ultimately, we develop a semiparametric method where parametric model and nonparametric model are integrated, retaining the advantages of both approaches.

The rest of the paper is structured as follows: The next section reviews bankruptcy prediction literature. In Section III, the problem of existing models is discussed and a new model is proposed. Section IV presents research data and variable selection, followed by evaluation methods. Section V and VI describe experimental results and conclusion, respectively.

II. PREDICTION MODELS FOR BANKRUPTCY PREDICTION

A. *Multivariate Discriminant Analysis (MDA)*

Multivariate Discriminant Analysis (MDA) constructs a discriminant function by maximizing the ratio of between-

1-4244-0457-6/06/$20.00 ©2006 IEEE

groups and within-groups variances. MDA obtains a linear discriminant function relating a set of independent variables to a dependent variable. In bankruptcy prediction, financial ratios are used as independent variables while the state of being bankrupt or healthy is used as dependent variable.

Altman [1] first used MDA in his Z-score model, which has developed benchmark status in academic literature, for the purpose of detecting failure potential. The discriminant function proposed by Altman is as follows [1]:

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 0.999X_5 \qquad (1)$$

where X_1 = working capital/total assets, X_2 = retained earnings/total assets, X_3 = earnings before interest and taxes/total assets, X_4 = market value equity/book value of total liabilities, X_5 = sales/total assets, and Z = overall index.

Following conditions should be satisfied for inferences of MDA to be valid:

1. mean vectors, covariance matrices, and prior probabilities of misclassification are known;
2. independent variables in each group follow a multivariate normal distribution;
3. the covariance matrices of each group are identical.

B. *Logistic Regression (LR)*

Logistic Regression (LR) approach was first introduced by Ohlson to bankruptcy prediction problem in 1980. Ohlson [29] used nine accounting-based variables (ratios) of company's size, leverage, liquidity and performance and proposed the Y-score model. While examining the relationship between binary and ordinal response probability and independent variables, LR assigns a Y-score in a form of bankruptcy probability to each company.

The function of the probability of a company to go bankrupt can be calculated as follows [5]:

$$f(x) = 1/(1+e^{-x}) \qquad (2)$$

The LR essentially uses a sigmoid function at the output. It takes the value of one if the sample company falls into bankruptcy and zero otherwise.

C. *Neural Networks (NNs)*

In 1990, the Neural Networks (NNs) technique has jumped into the field of corporate bankruptcy prediction and it has been a very popular technique ever since. Odom and Sharda [28] were the first to apply NNs to bankruptcy prediction problem. The NNs technique dominates the literature on business failure in 1990s, and it is still most frequently used in corporate bankruptcy prediction.

A neural network is typically composed of several layers of many computing elements called nodes. Each node receives input information from external inputs or from the output signal of other nodes. While processing the signals locally through a transfer function, the node outputs a transformed signal to other nodes or final result [39].

Most neural network approaches to bankruptcy prediction use MultiLayer Perceptron (MLP). In MLP, all nodes and layers are arranged in a feedforward manner [39]. The feedforward layered network contains three kinds of layer. The first layer is called the input layer where external information is received. The last layer is called the output layer where the network produces the final solution. In between, there are one or more internal or hidden layers. As the number of hidden layers increases, the network becomes more complex [12]. The three-layer MLP is the most commonly used NNs structure for two-group classification problems like bankruptcy prediction.

Many researchers put emphasis on the superiority of NNs technique to the classical techniques. First of all, NNs can recognize complex patterns with better accuracy, and they are able to learn from training samples, without any prior knowledge about the underlying problems [5]. Secondly, Coats and Fant [12] finds that non-numeric data can be easily included in a NN because the input data do not need to conform to some linearity assumptions. The third advantage is that a NN is perfectly suited for pattern recognition and classification in unstructured environments with noisy data, which may be incomplete or inconsistent [21].

III. PROPOSED METHODS

A. *Bias and Variance Tradeoff*

There have been many efforts to find a model that will provide the best possible output for a set of new/unknown input data. As a flexible, model-free approach, NNs can fit the training data very well and thus provide a low learning bias. However, their good data fitting capacity also makes them more susceptible to the overfitting which can cause large generalization variance.

The bias and variance terms provide useful information on how the estimation differs from the desired function. The bias measures the extent to which the average of the estimation function over all possible data sets differs from the desired function. The variance, on the other hand, measures the sensitivity of the estimation function to the data sample [17]. A simple parametric model that is less dependent on the data may misrepresent the true functional relationship and have a large bias, but a low variance because of simple predictable characteristics of linear models. In contrast, a complex nonparametric model that fits the data well tends to have low bias but high variance when applied to different data sets [39].

It is desirable to have both low bias and low variance in generalization; however, we cannot reduce both at the same time for a given data set because bias and variance are often incompatible. With a fixed data set, if we try to reduce one, it will inevitably cause the other to increase [17]. Hence, clearly what we want is to find a good tradeoff between model bias and model variance, that is, a model which is powerful enough to represent the underlying structure of the data, but not so powerful that the model output becomes unstable for prediction.

There have been a number of approaches to improve generalization by trade-off between the bias and variance.

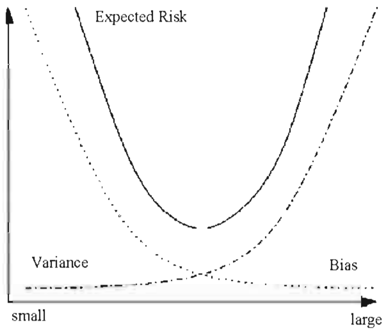

Figure 1. Complexity of Function Set.

Some well-known such methods include cross-validation [26] [38], training with penalty terms [37], and weight decay and node pruning [31] [32]. Trade-off is often made between the states of overfitting and underfitting. Somewhere between the states of overfitting and underfitting, an adequate model is assumed to exist. In many cases, tradeoff is made by the changes in model complexity as shown in Fig. 1. Constructive model approach begins with a small model and increments the model size until overall empirical risk is minimized. Pruning model approach begins with a large model and reduces the model complexity until the overall empirical risk is minimized.

However, tradeoff methods have some limitations for financial prediction model. The main problem is that they tend to reduce the variance at the cost of increased bias. Financial prediction requires accurate learning because the observed data is the main source of information. In such applications, underfitting is not very feasible option as the most information is carried by the observed data. It is also very difficult to identify which information can be discarded. The information provided by the observed data must be faithfully preserved while stable generalization is maintained. In short, financial prediction model requires retaining both low bias and variance simultaneously.

B. *Proposed Model*

So what we need is a model which is powerful enough to accurately fit the data as well as become stable for prediction. The flexible models such as ANNs can fit the observed data well; however, they can lose too much structural information about the underlying function, and neglect some existing linear dependencies among them. It may be possible to reduce the overfitting by utilizing the linear dependencies or structural relationship between the observed data.

Many real life problems are nonlinear in nature; however, they can be approximated by linear models in a small range of data. Combining these linear models can approximate an underlying nonlinear system. There are many different ways in which multiple local linear models may be combined. The most popular approach is via simple average of outputs from individual local models. Combining can also be done with weighted averaging that treats the contribution or accuracy of the local model differently [20].

In this study, we propose a semiparametric model for use. Semiparametric modeling is, as its name suggests, a hybrid form of parametric and nonparametric approaches to building and validating prediction models. In this study, we pursue to combine the advantages of classical parametric models with flexible NNs and show that the proposed model can provide both accurate model capacity and model stability in bankruptcy prediction.

C. *Model Description*

The idea of the semiparametric model is to combine a parametric form for structural relation between the dependent variable and input variables with weak nonparametric restrictions on the remainder of the model [30]. It means that big changes are modeled by parametric model and small changes are modeled by nonparametric component. If nonparametric approach is used to model the big changes, then it becomes a problem since variation (or changes) is too large for nonparametric model to learn well.

In the classical parametric approach, it is generally assumed that the dependent variable is functionally dependent on the input variables and unobservable errors according to a fixed structural relation of a function. The drawback to parametric modeling is the requirement that both the structural model and the error distribution are correctly specified. However, it is particularly difficult to correctly specify the error distribution, which represents the unpredictable component of the relation of the dependent variable to input variables [30].

On the other hand, nonparametric modeling typically imposes few restrictions on the joint distribution of the data; hence there are few chances for misspecification. Consistency of an estimator of a function is established under much more general conditions than for parametric modeling [30].

As a hybrid of the parametric and nonparametric approaches, semiparametric modeling is assumed to retain the advantages and disadvantages of each model. The proposed model consists of the parametric model embedded in NNs. In general, there can be a number of different architectures in which the parametric models may be embedded in the NNs model. We tried to embed LR in NNs in this experiment. The idea is derived from one of semiparametric models [11] as follows:

$$Y_t = \phi(Z_{1,t}, b_0) + \eta_0(Z_{2,t}) + e_t \quad (3)$$

The equation shows that it uses the linear weighted sum of parametric function and nonparametric function. Z_1 is for the parameter function, Z_2 for a nonparametric function and e_t is the error sequence with mean zero and finite variance.

We choose the proposed model to combine the structural information of LR (Z_1) with accurate data-fitting capacity of NNs (Z_2). The final output (Y) is the linear combination of Z_i with corresponding weight values. While varying the weight value for each model, we could find the best final output.

IV. Research Design and Methodology

Models are constructed to represent the relationship between the probability of bankruptcy and the respective financial features in corporations. In the following subsections, we describe the various steps of the data analysis. We present the reasoning behind the selection of appropriate variables for the experiment. Cross validation techniques are employed to enhance the generalization capacity of the models in this study.

A. Data Analysis

The data set of bankrupt and healthy companies used for this study was obtained from Baek & Cho [6]. The data was originally collected from a total of 4231 yearly financial statements from both 606 healthy and 56 bankrupt Korean companies. The financial statement data that are collected three or fewer years before the actual bankruptcy were used as bankrupt-representative data and the rest as healthy [6]. The data set consisted of 3893 cases of healthy companies and 248 of bankrupt companies during the period from 1994 to 2000. Table I shows the number of data. We can see that the number of bankrupt data is far smaller than that of healthy data.

The data set has a major problem that is 'the imbalance of data' [6]. As we can see from Table I, the bankrupt data accounts for only 5.8 percent of the whole data. The imbalance of data can be problematic because it may affect the predictive performance of the prediction models. Boritz, et al [10] examined the effect of various proportions of bankrupt companies in the training and testing data on the prediction ability of statistical techniques and neural networks. They found that having a smaller proportion of bankrupt companies to healthy companies would increase the Type I error rates. They concluded that the techniques had better capability to discriminate between bankrupt and healthy groups as the bankruptcy proportion increased in the training sample.

In this study, we construct a data set with an equal proportion of bankrupt companies to healthy companies to avoid the data imbalance problem. We increase the number of bankrupt cases by cross-multiplying the bankrupt data (248) to be of an equal number to the healthy data (3983). In summary, a data set consists of a total of 7966 cases containing fifty percent bankrupt companies.

Simple multiplication of training data can introduce unnecessary bias to the model. In order to avoid such bias, a random noise has been added to the multiplied training samples. This process is called 'jittering'. Jittering improves a model's ability to generalize. Generalization means the ability of a model to predict on the unknown inputs outside the training data. When a model yields a very high predictive accuracy, the result sometimes shows sign of overfitting, which can cause instability in generalization. It means that the model is too restricted to the training data, so when an unusual input which differs from the training data is presented to the model as an input, the model output may be unreliable. That is, it shows large variance to the testing data input, which is out of the range of its usual training data. Adding noise or jitter to the inputs during training, which is one of the most widely used means for preventing overfitting, has been found empirically to improve model generalization [4] [23] [25]. This is attributed to the idea that the noise added to the training patterns will smooth out each data point, thus prevent the model from learning the individual data points too precisely [8]. In practice, training with added noise has been shown to reduce overfitting and thus produce better-regularized model.

TABLE I
NUMBER OF DATA

The period of data	Healthy	Bankrupt	Total
1994-1996	1707	156	1863
1997-2000	2276	92	2368
Total	3983	248	4231

B. Variable Selection

Financial ratios have long been used as a tool for predicting potential bankruptcy of a company since the early bankruptcy prediction studies. The financial ratios are used as the input variables for the models. Based on the pioneering work by Altman [1], many researchers have used the Altman's five ratios as input variables for the models of their studies [6] [9] [12] [28] [34] [39]. Often a few other predictor variables are also used for prediction. Tam and Kiang [35] select 19 financial variables in their study. Boritz and Kennedy [9] examine the effectiveness of NNs in bankruptcy prediction with Ohlson's nine and 11 variables as well as Altman's five ratios. It is interesting to note one study uses as many as 57 predictor variables [27] while Fletcher & Goss [15] and Fanning & Cogger [14] use only three variables.

This paper uses Altman's five financial ratios as input variables. These variables are listed in Table II. Table III presents the generally accepted interpretations of the selected ratios. This list is selected from ratios proven to be popular and useful in earlier research on bankruptcy prediction [5] [13] [18].

C. Evaluation and Analysis

In order to accurately assess the predictive performance of each model, we adopt a cross-validation scheme for our model evaluation. As with jittering, cross-validation is regarded as one of the effective approaches to improve generalization of a model [22] [33] [39].

TABLE II
FINANCIAL RATIOS USED

variables	
x1	working capital/total assets
x2	retained earnings/total assets
x3	earnings before interests and taxes/total assets
x4	market value of equity/book value of total liabilities
x5	sales/total assets

TABLE III
INTERPRETATION OF RATIOS

variables	
x1	Short term liquidity of a company to pay its current debts
x2	Financial leverage to use liabilities to purchase assets
x3	Measures for true profitability of a company
x4	Solvency to issue and sell new shares to repay its debts
x5	Activity of a company to generate sales

First, we apply the holdout method to our data set. The idea is that it randomly splits a data set into two sets, called the training set and the testing set. The training set is used for model building or parameter tuning and the test set is used only to assess the predictive performance of the fully-trained model. However, the holdout method has a key drawback in that the single random division of a sample into training and testing sets may introduce bias in model selection and evaluation [39]. Since the estimated prediction rate can be very different depending on the characteristic of data, the holdout estimate can be misleading if we happen to get an unfortunate split.

So we employ the random subsampling method to overcome the limitation of the holdout method. It is a resampling technique using multiple random training and testing subsamples. Random subsampling performs a number of separate data splits of the data set. Each split randomly selects a fixed number of samples without replacement. For each data split, we retain the model with the training samples and estimate the generalization error with the test samples. In our experiment, this random data set selection and testing process will be conducted five times. When multiple random train-and-test experiments are performed, a model is learned from each training sample. The estimated error rate is the average of the error rates for the models derived from the independently and randomly generated test divisions. Random subsampling is expected to produce better error estimates than a single train-and-test division.

We evaluate the quality of each model by looking at the average error rates and the variance of the error rates. The bias and variance from semiparametric models will be compared to those of MDA, LR and NNs. And with the results compared, we will validate the usefulness of our proposed model in terms of accuracy and stability. We believe that the cross validation analysis will provide us valuable insights into the reliability and robustness of the models in regard of sample variation.

V. Experimental Results and Analysis

A. *Experimental Results*

The predictive performance of the semiparametric model has been compared with that of the MDA, LR and NN models in terms of bias and variance. The aim of experiment is to prove that the semiparametric models can achieve a good tradeoff between the model bias and variance. Table IV presents the average error rates and the variance of error rates of each prediction model on the five independent testing samples.

TABLE IV
The Bias and Variance of Error Rates of Prediction Models

Models	MDA	LR	NNs	NNs with Embedded LR
Sample 1	33.04%	32.45%	20.30%	20.30%
Sample 2	33.04%	30.60%	23.51%	23.44%
Sample 3	32.51%	32.93%	25.14%	24.36%
Sample 4	32.78%	32.71%	23.77%	23.66%
Sample 5	34.18%	31.23%	26.47%	26.47%
Mean	33.11	31.98	23.84	23.65
St.Dev	0.0064	0.0102	0.0231	0.0222

Within the prediction test in Table IV, the result shows that NNs Embedded LR produced the smaller error rates in average (23.65) and the smaller variance (0.0222) than those of NNs (23.84; 0.0231), which are the encouraging results of which we have expected.

Now, we analyze the results from the Table IV in terms of bias and variance. The average prediction error rates (Mean) represents bias, and variance can be estimated by measuring the standard deviation of the error rates (St.Dev) over the five independently different testing samples.

As shown in Table IV, the parametric models such as MDA and LR produced larger error rates, which mean that they have high bias. The nonparametric NNs model showed fairly high accuracy (low error rate); hence it has low bias. As we expected, our proposed model showed highest accuracy, thus it retains low bias.

The variance indicates the generalization capacity of a model to various unknown inputs. Theoretically, the parametric methods are likely to have low variance and nonparametric models tend to have high variance. Our experiments just showed the same results as expected. Table IV shows that the NNs with embedded LR model is successful to have produced lower variance than NNs model. Even though it seems only a fraction of improvement, it turns out a huge impact when we deal with the trillion dollars that are lost due to the incorrect prediction.

In short, parametric models such as MDA and LR, which do not depend on the sample data much, tend to have low variance but high bias. On the other hand, the nonparametric models like NNs can fit the training data very well; thus they provide a low learning bias. But, at the same time, they are more susceptible to the overfitting which can cause high variance.

As it is seen from Fig. 1, the parametric models are to reside at the left side of the graph; that is, it shows high bias and low variance. The nonparametric models are likely to occupy the right side with low bias and high variance. Our proposed semiparametric model is theoretically assumed to fit in the middle providing both low bias and low variance [11].

The experimental results confirm our hypothesis that the proposed semiparametric model (NNs with embedded LR) produces improved generalization while it is preserving the accurate data modeling.

B. *Assumptions*

The financial ratios chosen for consideration in this study were based on the previous research in bankruptcy prediction applying statistical techniques such as MDA. The use of financial ratio variables precludes the possible influence of external factors such as national and international economic conditions, litigation, taxes or other unforeseen environments. Some researchers put emphasis on the macroeconomic variables for bankruptcy prediction as they assessed that the different economic environments lead to different models for the prediction of bankruptcy [13] [16] [40]. In this study, we did not consider the influence of macroeconomic factors on the financial condition of a company. Further, we believe that the analysis and results are accurate to the extent that the reported data reflect the actual financial condition of a company.

VI. Conclusion

In this study, we presented recent development in bankruptcy prediction in terms of different types of modeling approaches. Overall, there are some advantages and disadvantages associated with each model. Parametric models are less accurate; however, more reliable, whereas the nonparametric models are more accurate yet unstable.

A general approach is to find an optimal model that is between the state of underfitting and overfitting. In this study, we proposed to use semiparametric model approach that is assumed to retain the advantages of both parametric and nonparametric models. By combining the advantages of classical parametric models and emerging NNs, the proposed model promises to provide both inherent stability (of parametric models) as well as accurate model capacity (of NNs).

In general, bias and variance are often incompatible, so reducing the variance often leads to the increased bias. However, we expect that the semiparametric approach will provide a good tradeoff between the bias and variance. That is, the semiparametric model promises to maintain sustainable model stability in generalization while preserving the accurate data-fitting capacity of nonparametric models. From our preliminary experimental results, the semiparametric model seems to give promising results that the model could achieve a good tradeoff between the bias and variance. We expect that future research will further confirm the feasibility of our proposed model.

References

[1] E.I. Altman, "Financial Ratios discriminant analysis and the prediction of corporate bankruptcy", *Journal of Finance*, vol.23, pp. 589-609, 1968.

[2] E.I. Altman, R.G. Haldeman and P. Narayanan, "Zeta Analysis: A new model to identify bankruptcy risk of corporations", *Journal of Banking and Finance*, vol. 1, pp. 29-51, 1977.

[3] ABI, "Bankruptcy filings set new calendar year record in 2003", *American Bankruptcy Institute*, 2004.

[4] G.An, "The effects of adding noise during backpropagation training on a generalization performance", *Neural Comp*, vol. 8, pp. 643-674, 1996.

[5] A.F. Atiya, "Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results", *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 929-935, 2001.

[6] J. Baek, and S. Cho, "Bankruptcy prediction for credit risk using an auto-associative neural network in Korean firms", *IEEE Int. Conf. on Computational Intelligence forFinancial Engineering*. pp. 25-29, 2003.

[7] R.E. Bellman, *Adaptive Control Processes*, Princeton University Press, New Jersey, 1961.

[8] C.M. Bishop, *Neural networks for pattern recognition*, Oxford University Press, New York, 1995.

[9] J.E. Boritz and D.B. Kennedy, "Effectiveness of neural network types for prediction of business failure", *Expert Systems With Application*, vol. 9, no. 4, pp. 503-512, 1995.

[10] J.E. Boritz, D.B. Kennedy and Albuquerque, A., "Predicting corporate failure using a neural network approach", *Int. Journal of Intelligent Systems in Acc, Fin and Management*, vol. 4, no. 2, pp. 95-111, 1995.

[11] X. Chen, J. Racine, and N.R. Swanson, "Semiparametric ARX Neural-network models with an application to forecasting inflation", *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 674-683, 2001.

[12] P.K Coats and L.F. Fant, "A neural network approach to forecasting financial distress", *The Journal of Business Forecasting*, pp. 9-12, 1992.

[13] A.I. Dimitas, S.H. Zanakis and C. Zopouniis, "A survey of business failures with an emphasis on prediction methods", *European journal of Operational Research*, vol. 90, pp. 487-513, 1996.

[14] K.M. Fanning and K.O. Cogger, "A comparative analysis of artificial neural networks using financial distress prediction", *Intelligent Systems in Accounting, Finance and Management*, vol. 3, pp. 241-252, 1994.

[15] D. Fletcher and E. Goss, "Forecasting with NNs: An application using bankruptcy data", *Info and Management*, vol. 24, pp. 159-167, 1993.

[16] G. Foster, *Financial Statement Analysis*, PrenticeHall, NewJersey, 1986.

[17] S. Geman, E. Bienenstock and T. Dousat, "Neural networks and the bias/variance dilemma", *Neural Computation*, vol. 5, pp. 1-58, 1992.

[18] J.S. Grice and R.W. Ingram, "Tests of the generalizability of Altman's prediction model", *Journal of Bus Research*, vol. 54, pp. 53-61, 2001.

[19] M.M. Hamer, "Failure prediction: Sensitivity of classification accuracy to alternative statistical methods and variable sets", *Journal of Accounting and Public Policy*, vol. 2, pp. 289-307, 1983.

[20] S. Hashem, and B. Schmeiser, "Improving model accuracy using optimal linear combinations of trained neural networks", *IEEE Transactions on Neural Networks*, vol. 6, pp. 792-794, 1995.

[21] D.D. Hawley, J.D. Johnson and D. Raina, "Artificial neural systems: A new tool for financial decision-making", *Financial Analysts Journal*, vol. 46, pp. 63-72, 1990.

[22] J.S.U. Hjorth, *Computer intensive statistical methods validation, model selection, and bootstrap*, Chapman & Hall, London, 1994.

[23] L. Holmstr and P. Koistinen, "Using additive noise in backpropagation training", *IEEE Transaction on NNs*, vol. 3, pp. 24-38, 1992.

[24] H. Jo and I. Han, "Integration of case-based forecasting, neural network and discriminant analysis for bankruptcy prediction", *Expert Systems with Applications*, vol. 11, no. 4, pp. 415-422, 1996.

[25] P. McCullagh and J.A. Nelder, *Generalized linear models*, 2nd ed., Chapman & Hall, London, 1989.

[26] N. Morgan and H. Bourlard, "Generalization and parameter estimation in feedforward nets: Some experiments", *Advanced Neural Information Process Systems*, vol. 2, pp. 630-637, 1990.

[27] K. Lee, I. Han and Y. Kwon, "Hybrid NN models for bankruptcy predictions", *Decision Support Systems*, vol. 18, pp. 63-72, 1996.

[28] M.D. Odom and R. Sharda, "A neural network model for bankruptcy prediction", *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN)*, San Diego, CA, pp. 163-168, 1990.

[29] J. Ohlson, "Financial ratios and probabilistic prediction of bankruptcy", *Journal of Accounting Research*, vol. 19, pp. 190-131, 1980.

[30] J.L. Powell, "Estimation of semiparametric models" in Engle, R.F. & McFadden, D.L. (ed.), *Handbook of Econometrics*, Elsevier B.V., North-Holland, vol. 4, ch. 41, pp.2443-2521, 1994.

[31] R. Reed, "Pruning algorithms – A survey", *IEEE Transactions on Neural Networks*, vol. 4, pp. 740-747, 1993.

[32] C. Schittenkopf, F. Deco, and W. Brauer, 'Two strategies to avoid overfitting in FFNNs", *Neural Networks*, vol. 10, pp. 505-516, 1997.

[33] J. Shao, "Linear model selection by cross-validation", *Journal of the American Statistical Association*, vol. 88, pp. 486-494, 1993.

[34] R. Sharda, and R.L. Wilson, "Neural network experiments in business-failure forecasting: Predictive performance measurement issues", *International Journal of Computational Intelligence and Organizations*, vol. 1, pp. 107-117, 1996.

[35] K.Y. Tam and M.Y. Kiang, "Managerial applications of neural networks: The case of bank failure predictions", *Management Science*, vol. 38, pp. 926-947, 1992.

[36] W. Treacy and M. Carey, "Credit risk rating at large US banks", *Journal of Banking & Finance*, vol. 24, pp. 167–201, 2000.

[37] A. Weigend, D. Rumelhart and B. Huberman, "Predicting the future: A connectionist approach", *International Journal of Neural Systems*, vol. 3, pp. 193-209, 1990.

[38] S.M. Weiss and C.A. Kulikowski, *Computer systems that learn*, Morgan Kaufmann Publishers, San Francisco, 1991.

[39] G. Zhang, M.Y. Hu, B.E. Patuwo, and D.C Indro, "Artificial neural networks in bankruptcy prediction", *European Journal of Operational Research*, vol. 116, pp. 16-32, 1999.

[40] C. Zopounidis and A.I. Dimitras, *Multicriteria Decision Aid Methods for the Prediction of Bus. Failure*, Kluwer publishers, the Netherlands, 1998.