Eötvös Loránd University
Faculty of Infomatics

APPLICATION OF HEURISTIC METHODS IN DEFAULT FORECASTING

THESES OF PHD DISSERTATION

József Bozsik

Supervisor:
Róbert Fullér, professor, D.Sc.

ELTE DOCTORAL SCHOOL OF INFORMATICS
Erzsébet Csuhaj-Varjú, professor, D.Sc.

Basics and methodology of informatics Doctoral Program
János Demetrovics, professor, D.Sc.

Budapest, 2016

Introduction

In the state of today's global financial market - let us think about the US financial crisis - the reliable and accurate decision support systems are attributed to have a very important role by banks, may they be either in the public or the private sphere. The effectiveness of these systems lays on their diversity and pluralism. In these heterogeneous systems they try to produce the best possible results that are satisfactory in all aspects, utilizing a variety of technologies and theories with information from various origin.

The default forecast model is a key component of a complex financial decision support system. This of course comes from the simple fact that most financial market participants prefer secure and predictable credits that promise a high profit and mean a lower risk at the same time.

Bankruptcy is important not only economically, but also may have significant moral effect socially. The bankruptcy of a larger company could affect a large segment of society, therefore the importance and legitimacy of default forecasting is supported by both financial and moral reasons, and not only for economic organizations.

In light of these findings the default forecasting systems are all the more important. Of course, these systems realize a number of different theoretical solutions, and they present the cumulative result compared to certain partial results. Like many parts of the fundamental economics, this field also heavily relies on mathematics, especially on the toolbox of mathematical statistics. The main default forecasting models use statistics based logistic regression and timeline analysis.

These methods are very well useable, but the science of economics has been extended with a lot of new methods and tools. Each method has its own advantages and disadvantages, so they do not compete with each other and we are not looking for the best method in every aspect. These methods give a well useable and - in a term - best result for the financial experts when they are applied side by side.

The objective of the dissertation is the creation and study of new and novel artificial intelligence models in default forecasting. Since default forecasting is considered to be a relatively young discipline in Hungary, few people have dealt with the applicability of artificial intelligence methods in this area in domestic relations. As a result, this area offers many opportunities for the application of both classical and new hybrid artificial intelligence methods.

The classical economic methods in many cases are 'one-armed giants', because they define lot of restrictions, and here we should primarily think about the linearity and the dispersion. These restrictions can be lifted by using artificial intelligence methods. This is one of the reasons why heuristic methods can be applied to solve optimization problems successfully.

In addition, the ever-changing legal and economic environment also warrant the development of new, flexibly adaptable models. The flexibility and adaptability are important virtues of the hybrid artificial intelligence methods, but the successful application abroad also supports the legitimacy of these methods.

Default forecasting can be viewed as a segmentation, clustering or classification problem, and the classical artificial intelligence methods - especially the heuristic methods - are very well suited for solving such tasks. Examples of such methods include neural networks, fuzzy systems, decision trees and genetic algorithms. We can consider some of the procedures and their parts as optimization problems, and if we can define a measure for the compliance of the procedure then we can compare them to the results of other methods. This can usually be done in the classification tasks, and based on this we can compare the results of the new methods with the results of other methods.

The theses of the dissertation

As the central results of my research I have developed and summarized those artificial intelligence methods which provide a well useable solution to the core part of this dissertation: default forecasting. I introduced the efficiency indicators of the methods, and I provided a tool for their efficient usage. Besides the precise description of the presented methods I also thrived for placing the researches they were based on into an international context as well. In the dissertation I also placed an emphasis on the detailed presentation of the methodology underlying the models.

In my research I examined the traditional economic methods first, and studying them deeper I learned about their beneficial properties and explored their drawbacks, which supported the further direction of my research, as it is detailed in Chapter 2. The published results and the international experience confirmed that it is possible to develop efficient artificial intelligence based models and methods for default forecasting based on the data of domestic companies as well.

A further direction of the research was first determined by the procedures belonging to the family of traditional artificial intelligence methods, which I showed in Chapter 3 in detail, and discussed the related results there [1], [2], [3], [4], [5] [6]. I made progressive steps towards the hybrid methods that were explained in detail in Chapter 4, and I described my own results [7], [8], [9], [10].

Thesis 1: The specialization of the ID3 algorithm for credit assesment

I developed the model of ID3 algorithm used for credit assessment in banks. I extended the original ID3 method, and specialized it to the problem in my implemented model.

In this thesis I summarized the results of my research carried out with decision trees, which are described in Chapter 3.1. I presented in detail the basis of my model - the Hunt algorithm and the ID3 algorithm, which can be derived from it. Behind the implementation that provides my own results I map the variables with constant values with the help of dynamically created attributes in the decision tree. The obtained results of my model are presented in detail in Chapter 3.1.2., where an outstanding feature in addition to the rapid and effective running time is a classification accuracy of 96%. I published these results on the LINDI conference held in 2011 and referred by IEEE [1].

Thesis 2: The genetic model in default forecasting

As a result of my research with genetic algorithms I developed two models: a so called basic model and a refined model. In the final model I used refined crossing with additional extra genes, which proved to have an outstanding 84.5% classification accuracy compared to the 79% of discriminant analysis.

I presented the results of my researches with promising genetic algorithms based on international professional literature in Chapter 3.2, first with the help of a basic model in Chapter 3.2.2, and then with a refined model in Chapter 3.2.3. I created the final model by using the results and experience I gained with the two models, in which I arrived to a method containing two freezing steps. I have shown that I can achieve a 84.5% classification accuracy ratio compared to the 79% of the discriminant analysis based model, if I use refined crossing with adding extra genes. My results were published in the ICAI conference held in 2010 [2].

Thesis 3: Model based on feedforward neural network

I examined the feedforward multilayer neural networks and created six models with different configurations. I showed that my own model containing two hidden layers and made by 6-5 networks produced a better classification accuracy than the traditional models.

In the wake of international publications I started to experiment with neuron networks. I presented the international results and the theory of multi-layer feedforward networks in chapters 3.3.1 and 3.3.2. This was followed by the presentation of the model developed by us and its specialties in Chapter 3.3.3. My results are presented in the form of a comparative analysis in Chapter 3.3.4. I showed that I can achieve a classification accuracy of 82.7% with a double-layer feedforward network using 6-5 neuron networks opposed to the 79% of the traditional model. I presented my publication related to this thesis on the IRFIX conference held in 2009 and on the ICAI Conference held in 2010 [3].

Thesis 4: SVM based model

I extended the SVM algorithm and placed it on a special base to create an efficient and well useable method in default forecasting. I studied and implemented the method of sequential minimal optimization (SMO), and built it into my own model. The specialties of the model and the uniqueness of the implementation guaranteed a quick and efficient process, of which I showed that can achieve a classification accuracy of 78% by using the Gaussian kernel function.

The experience and motivating results I gathered by studying neuron networks prompted me to further focus my research on the new methods related to neural networks. I show the results of my experiments with support vector machines and my own unique set of solutions accordingly in Chapter 3.4. I examined and implemented the algorithm of sequential minimal optimization (SMO) and integrated it to my own model. The results of the model created by me were measured using my concrete implementation, which specializes in the signal-slot mechanism provided by the QT language. This model was presented in Serbia on the international SISY conference held in 2011, the results were presented in my conference publication referred by IEEE [4].

Thesis 5: ESN based model

No significant results are known in international publications demonstrating the application of echo state networks in default forecasting, so I performed groundbreaking work in this area. I examined the echo state networks, the Jordan and Elman networks, with the help of which I implemented an echo state network based default forecasting model, and proved that by training it with a Wiener-Hopf teaching algorithm it can achieve a classification accuracy of 71%.

The results of my research with echo state networks were discussed in detail in Chapter 3.5, comparing the results with the results of the discriminant analysis. I examined the Hopfield networks and Elman, Jordan networks. I created my echo state network based default forecasting model based on the experience gained there. I presented my results on the international LINDI conference in Slovakia held in 2012, and published them in the conference publication referred by IEEE [5].

Thesis 6: SOM based model

Among the unsupervised teaching neuron networks I showed that my research conducted in the Kohonen's self-organizing maps could achieve a classification accuracy of 97% by using SOM.

Last but not least, I close the range of classical methods related to neuron networks with the results of my research regarding the Kohonen's self-organizing maps. Among the unsupervised neuron networks producible by teaching methods the international results of the Kohonen's self-organizing maps were the most promising in default forecasting, so this method has also became the center of my researches.

The results related to them are presented in detail in Chapter 3.6, where besides the 97% of classification accuracy of the model measured on the teacher-set, the test sample set gave a result of 73.5%. I presented my results linked to this on the international SISY conference in Serbia held in 2012, and published them in the conference publication referred by IEEE [6].

These results and experience has shown me - and the international results further confirmed it - that it is worth to create models based on new hybrid methods by combining classical methods and emphasizing their good qualities, and to examine their performance and results. Under the aegis of this, my research with the so-called hybrid artificial intelligence methods were placed in Chapter 4.

Thesis 7: Neural network combined with a decision tree

I developed and implemented two types of decision tree based hybrid methods. The first is the so called brute force model, on which I based my refined thin model, which is an improved version of the former. These were compared to each other and to the results of the classic methods. I showed that the refined thin model can achieve a 83.97% of classification accuracy.

In Chapter 4.1 I presented the neural network based model combined with a decision tree, and compared its results to the results of the traditional methods. In the course of my research I have built two models here: the results of the first model, which I call brute force model, are described in detail in Chapter 4.1.1, while the results of the second model, which is referred to as the refined thin model, are presented in Chapter 4.1.2. The two models - with a classification accuracy of 84.29% and 83.97% respectively - both surpass the 79% result of the discriminant analysis. These results were presented for the first time in 2010 on the CSCS conference held in Szeged, and were published in 2011 in the journal Acta Periodica Polytechnic [7].

Thesis 8: ANFIS model

I developed an adaptive neural network based fuzzy inference system, in which I used a five-layer network. I have shown that with the help of the Gaussian type inclusion function and 17 financial indicators I can create a model with a classification accuracy of 84%.

I carried out experiments with fuzzy systems among hybrid models, which are the main field of the research of my supervisor, Róbert Fullér. I have built several hybrid models; one of them is a neuro-fuzzy system of which the results are described in detail in Chapter 4.2. These results were presented on the international CINTI conference held in 2010, and published them in the journal publication referred by IEEE [8].

Thesis 9: FSVM model

The fuzzy SVM-based model showed a classification accuracy between 82% and 93% by using the Gaussian type kernel function. Contrary to the expectations the polynomial kernel function based FSVM model has achieved only a 39% of classification accuracy.

The outstanding results of the ANFIS model prompted us to focus my research on another hybrid fuzzy model, namely FSVM. The FSVM model is a fuzzy support vector machine based approach, which is described in more detail in Chapter 4.3 with the related results. I have shown that I can achieve a classification accuracy of 82% and 93% with the FSVM model, which is an outstanding result.

These results were presented and published on the international CINTI conference held in 2011 [9].

My further results concerning my research with hybrid Fuzzy models were presented and published on the international CINTI conference held in 2012 [10].

References

[1] Bozsik J., Körmendi G.: Decision tree-based credit decision support system, *Logistics and Industrial Informatics (LINDI), 2011 3rd IEEE International Symposium,* (2011), 189-194.

[2] Bozsik J.: Genetic algorithm in default forecast, *Proceedings of the 8th International Conference on Applied Informatics Eger, Hungary*, **1**, (2010), 41-51.

[3] Bozsik J.: Artificial Neural Networks in Default Forecast, *Proceedings of the 8th International Conference on Applied Informatics Eger, Hungary*, **1** (2010), 31–39.

[4] Bozsik J., Kozma M.: Neural network based Support Vector Machine in financial default forecast, *Intelligent Systems and Informatics (SISY), 2011 IEEE 9th International Symposium*, (2011), 163-167.

[5] Bozsik J., Ilonczai Zs.: Echo state network-based credit rating system, *Logistics and Industrial Informatics (LINDI), 2012 4th IEEE International Symposium*, (2012), 185-190.

[6] Bozsik J., Maksai D., Ilonczai Zs.: Creditworthiness decision-making system based on self-organising maps, *Intelligent Systems and Informatics (SISY), 2012 IEEE 10th Jubilee International Symposium*, (2012), 171-175.

[7] Bozsik J.: Decision tree combined with neural networks for financial forecast, *Periodica Polytechnica Electrical Engineering*, **55**/3-4 (2011), 95-101.

[8] Bozsik, J: Neural fuzzy system for default forecasts, *Computational Intelligence and Informatics (CINTI), 2010 11th International Symposium*, (2010), 69-74.

[9] Bozsik J.: Fuzzy support vector machine in credit rating systems, *Computational Intelligence and Informatics (CINTI), 2011 IEEE 12th International Symposium*, (2011), 313-318.

[10] Bozsik J., Fullér R.: A neural fuzzy system for asset valuation, *Logistics and Industrial Informatics (LINDI), 2012 4th IEEE International Symposium,* (2012), 197-200.