The classification of Iranian banks based on Artificial Neural Network (ANN)

Mahdi Salehi (corresponding author)
Faculty of Economics and Administrative Sciences, Ferdowsi University of Mashhad, Iran,
E-mail:Mahdi_salehi54@yahoo.com

Elahe Torabi
Department of Accounting, Science and Research Branch,
Islamic Azad University, Nyshabour, Iran.

Jalil Khaksar

Department of Accounting, Mashhad Branch,
Islamic Azad University, Mashhad, Iran.

Abstract

In this study, it is attempted to examine the banking practice in Iran based on new scientific methods. It is used the financial ratios demonstrating healthy or non-healthy of banks to assess the financial health of listed banks in Tehran Stock Exchange. The assessment of these ratios with use of Neural Network (NN) as a non-parametric method for modeling is recommended for presenting this model. Information about the financial health of banks could be effective on the decisions of different groups of banks' financial reports users, including shareholders, auditors, stock exchange, central bank and etc.

Keywords: Neural Network(NN), Artificial Neural Network(ANN), Multi-Layer Perceptron (MLP), classification

Introduction

The role of banks in providing services to governments and society is not only undeniable but also was very significant and crucial and can be examined from different point view. Managing the peoples' life affairs and countries' economic affairs requires banks. It is possible to have a good influence on the decisions of users according to the importance of banks in modern societies with classifying the banks based on their financial health. According to this of banks in society is as the same as blood in vessel, there for it seems impossible to live without banks.

On the other hand ,Neural Network (NN) has emerged over the years and has made remarkable contribution to the advancement of various fields of endeavor. The NN has been ray effective in the field of engineering, agriculture, economy, financial issues.

Based on this efficiency we attempts to examine the financial health of banks based on 10 significant and crucial variables in NN.

In this study, financial health classification modeling of banks has been considered. Here, it is possible to identify effective factors in determining the financial health of banks with utilizing the data mining tools, including the analysis of each independent variable.

Neural Network-based method is selected to categorize because of the complexity of this issue. Although the Neural Network-based methods have been used in various fields, they have been rarely used for classification of banks.

The structure of this article continuation is organized as follows:

In the section 2, the proposed method steps for classification are described. The process of data mining involves data collecting, preprocessing and the quality of extracted factors affecting on the classification of banks are described in section 3. Subsequently, it involves the classification results of banks and discussion on the results of section 4 and ends with collecting and presenting the suggestions in section 5.

Ranking with Classification Models

Ranking examples using classification models, such as rule based models, is not a novel undertaking. There have been many attempts in the past, including those by researchers in the fields of expert systems, e.g., MYCIN (Buchanan and Shortliffe (eds.) 1984) and fuzzy logic (Wang L.X. and Mendel J.M. 1992). geometric methods(Alvarez I. & Bernard S. 2005), hybrid trees including: the Perceptron Tree(Utgoff, P. 1988) and NBTree (Kohavi R (1996), and ensembles of tree(Provost, F., Fawcett, T., & Kohavi, R. 1998). Recently, Loterman et.al (Loterman G., Brown, I., Martens, D., Mues, C., and Baesens, B.2012) has reported in the International Journal of Forecasting on a trend of using nonlinear techniques that perform significantly better than more traditional linear techniques in modeling financial markets. They also advocate the use of comprehensible model components. The research presented in this paper contributes to this trend. There are also several existing studies in Machine Learning e.g., (Zhang J & Michalski, R.S.1995). These studies generally focused on methods for generating partial matching, whereby the scores for individual examples are computed based on how well they match the rules. In these approaches, examples that satisfy all conditions of a rule share the same score. Extensive studies, on the other hand, have been dedicated to the incorporation of ranking capabilities into the decision tree learning paradigm. Related work generally falls into four groups of methods: learning probability estimation trees (Provost F. J. & Domingos, P. 2003),

Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) is an information processing paradigm that is nspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurones) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurones. This is true of ANNs as well. (Oludele Awodele and Olawale Jegede ,2009)

One of the most important applications of neural network is predicting and forecasting of special variable by some inputs (John, 2003). This means that a series of input data is used to predict output variable by specific process. In some applications, neural network models get a series of

data as input and output variable, and then such nonlinear regression estimate the relation between input and output variable (David, 2006). Figure 1 demonstrates this relation by multilayer perceptron neural network two input, two output and two hidden layers.

In this model, all data are divided into 2 groups including training and test data. The test data eliminates and training data is used to carry out model fitting process. After model fitting, the accuracy of model is assessed by test data. Most researchers consider 10 to 20% of all data as test data (Ward System Group, 1995),
even though in some neural network models, the size of test data reaches 25%. It is important to notice that in time series model, the test data must be selected from latter data but in ordinary model, it must be selected randomly. The basic goal of test data is to forecast solutions for these data and then compare them with actual solutions. It is obvious that the differences between forecasted solutions and actual solutions for dependent variable demonstrate error quantity which is calculated by absolute error or mean squared error. Sometimes, test data set is defined as product set (Ward System Group, 1995). Figure1 demonstrates one of the important neural networks called multi layer perceptron.

As observed, this network has an input, one output layer and one hidden layer. In this network, x1 and x2 demonstrate input vector while y1 and y2 is output vector.

Then this network including 2 input and 2 output variable and w jk i specify the weight of j variable in k neuron in layer i. Any neurons in the hidden layer include a transformation function thus, the output of one of the hidden layer as an example can be represented as follows:

$$Y^H = g(\sum_{i=1}^{2} w_{jk}^{i} x_j) + w_{jb}$$

Knowing hidden layer function quantity, output layer quantity can be calculated as follows:

$$Y^O = \sum_{h=1}^{2} w_{kj}^{2} \left(g\left(\sum w_{jk}^{i} x_j \right) + w_{jb}^{1} \right) + w_{jb}^{2}, k = 1,2$$

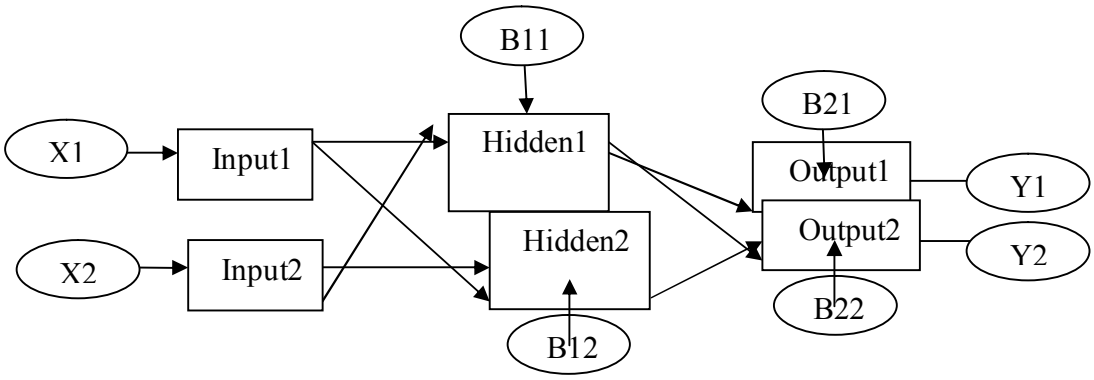

Figure 1. Multilayer perceptron neural network with two input and two output variables.

Architecture of Neural Networks

Neural networks are not only different in their learning processes but also different in their structures or topology. Bose (1996) has broadly classified neural networks into recurrent (involving feedback) and nonrecurrent (without feedback) ones. In a little more details, Haykin has divided the network architectures into the following three classes:

Feed-forward Networks

Feed-forward ANNs allow signals to travel one way only; from input to output. There is no feedback (loops) i.e. the output of any layer does not affect that same layer. Feed-forward ANNs tend to be straight forward networks that associate inputs with outputs. They are extensively used in pattern recognition. This type of organization is also referred to as bottom-up or top-down. Single-layer perceptrons and Multi-layer perceptrons are classes of feed forward networks.

Single-layer perceptrons (feed forward networks)

The single-layer perceptrons was among the first and simplest learning machines that are trainable .In Haykin's book (1999), perceptron denotes the class of two-layer feed forward networks,1) whose first-layer units have fixed function with fixed connection weights from the inputs, and 2) whose connection weights linking this first layer to the second layer of outputs are learnable.

Multi-layer perceptrons (feed forward networks)

Multi-layer feed forward structures are characterized by directed layered graphs and are the generalization of those earlier single layer structures (Bose, 1996).

Structure and features of MLP

Multi-layer perceptron (MLP) networks are feed forward nets with one of more layers of nodes between the input and output nodes. The structure of an unadorned multilayer perception network is shown in Figure 2.

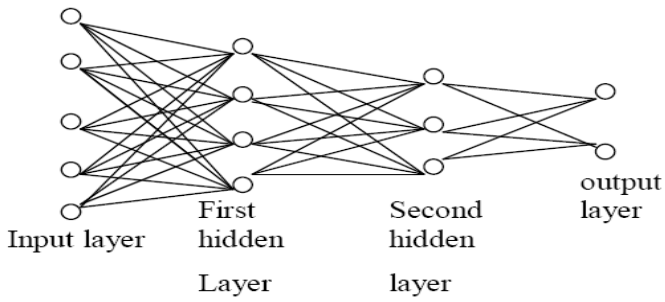

Figure 2. Feed forward multi-layer perceptron architecture
(Pandya & Macy, 1996, p.74)

The capabilities of multi-layer perception stem from the nonlinearities used within nodes.

Feedback networks

Feedback networks (Figure 2) can have signals traveling in both directions by introducing loops in the network. Feedback architectures are also referred to as interactive or recurrent, although the latter term is often used to denote feedback connections in single-layer organizations. In the neural network literature, neural networks with one or more feedback loops are referred to as recurrent networks. A recurrent network distinguishes itself from a feed forward neural network in that it has at least one feedback loop.

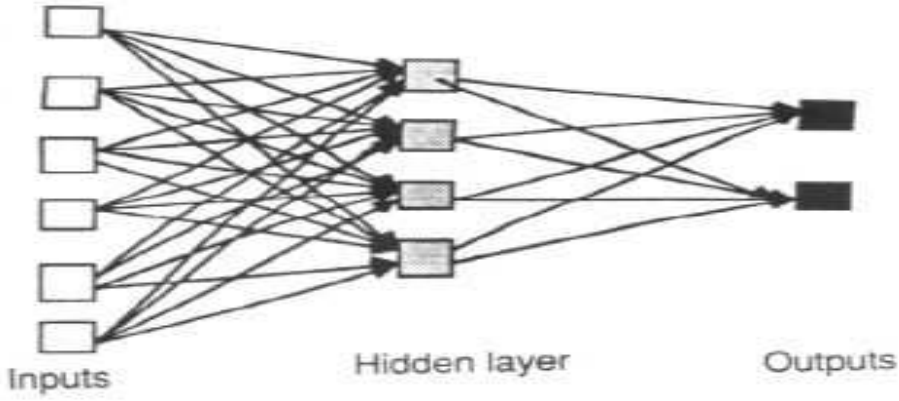

Figure 3. An example of a simple feed forward network
(Stergiou & Siganos, 1996)

Network Layers

The commonest type of artificial neural network consists of three groups, or layers, of units: a layer of "input" units is connected to a layer of "hidden" units, which is connected to a layer of "output" units. The activity of the input units represents the raw information that is fed into the network .The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units .The behavior of the output units depends on the activity of the hidden units and the weights between the hidden and output units.

Training a Neural Network

For the most part, a network is trained by changing the weights of the connections between nodes. These weights can be randomly chosen or individually chosen. Usually, a computer program randomly generates values for connection weights. Then, the network is given an input, and it is allowed to process the information through its nodes to produce an output.

The learning process

The memorization of patterns and the subsequent response of the network can be categorized into two general paradigms:

- **Associative mapping** in which the network learns to produce a particular pattern on the set of input units whenever another particular pattern is applied on the set of input units.
- **Auto-association**: an input pattern is associated with itself and the states of input and output units coincide.

All learning methods used for adaptive neural networks can be classified into two major categories:

- ***Supervised learning*** which incorporates an external teacher, so that each output unit is told what its desired response to input signals ought to be.

- ***Unsupervised learning*** uses no external teacher and is based upon only local information. It is also referred to as self-organization, in the sense that it self-organizes data presented to the network and detects their emergent collective properties.
-

Data analysis method

At first, data analysis method begins with the financial ratios calculation using the Excel software and then banks are classified by Artificial Neural Network modeling to classify the health of the banks. We used MLP method for data processing.

On the other hand, the research has faced with limitations due to the lack of access to all financial statements of banks between the years of 2008 - 2011 and lack of sufficient data on financial statements and attached notes of financial statements for calculating the ratios.

Hence, the problem provides us with a useful model by "checking the effective variables in classifying the banks". Ten variables are above the minimum amount of importance according to the Table 1.

Table 1. Variables of the study

Number	Description of the Ratios	Status
1	**Total operating expense to total operating income**	Lower is better
2	**Return on average total assets**	higher is better
3	**Return on equity**	higher is better
4	**Return on total assets**	higher is better
5	**Personnel expense to total operating income**	Lower is better
6	**Total interest expense to total operating income**	Lower is better
7	**Earnings assets to total assets**	higher is better
8	**Total operating income to total assets**	higher is better
9	**Nonperforming loans to total assets**	higher is better
10	**Allowance for loan and lease loss to net loans and leases**	higher is better

Table 2. List of networks Cluster outputs and inputs.

BANK	YEAR	Total operating expense to total operating income	Return on average total assets	Return on equity	Return on total assets	Personnel expense to total operating income	Total interest expense to total operating income	Earnings assets to total assets	Total operating income to total assets	Nonperforming loans to total assets	Allowance for loan and lease loss to net loans and leases	Cluster (out put)
EGHTESAD NOVIN	86	0.154	0.014	0.329	0.014	0.022	0.194	0.008	0.121	0.005	0.117	2
PARSIYAN	86	0.387	0.020	0.315	0.020	0.066	0.387	0.014	0.037	0.034	0.128	2
KARAFAIN	86	0.168	0.029	0.036	0.029	0.044	0.173	0.024	0.117	0.077	0.105	1
MELAT	86	0.423	0.005	0.121	0.005	0.178	0.519	0.001	0.069	0.001	0.063	3
PASARGAD	86	0.183	0.029	0.229	0.029	0.038	0.185	0.018	0.083	0.001	0.078	2
POST BANK	86	1.028	0.000	0.003	0.000	0.549	1.062	0.003	0.047	0.000	0.047	3
TEJARAT	86	0.410	0.011	0.184	0.011	0.160	0.494	0.002	0.070	0.000	0.066	3
SINA	86	0.099	0.012	0.156	0.012	0.060	0.099	0.016	0.145	0.016	0.144	1
SADERAT	86	0.731	0.005	0.046	0.005	0.104	0.733	0.010	0.083	0.000	0.067	3
EGHTESAD NOVIN	87	0.152	0.018	0.383	0.018	0.023	0.186	0.005	0.150	0.003	0.143	1
PARSIYAN	87	0.422	0.173	0.258	0.173	0.084	0.438	0.039	0.364	0.297	1.452	1
KARAFAIN	87	0.133	0.035	0.464	0.035	0.041	0.140	0.013	0.138	0.061	0.124	1
MELAT	87	0.415	0.007	0.154	0.007	0.181	0.524	0.002	0.077	0.021	0.071	3
PASARGAD	87	0.144	0.025	0.232	0.025	0.030	0.161	0.011	0.124	0.008	0.111	1
POST BANK	87	0.876	0.008	0.132	0.008	0.471	0.911	0.004	0.073	0.001	0.065	3
TEJARAT	87	0.363	0.008	0.145	0.008	0.158	0.474	0.004	0.075	0.000	0.067	3

SINA	87	0.124	0.018	0.193	0.018	0.071	0.124	0.009	0.152	0.008	0.124	2
SADERAT	87	0.558	0.018	0.174	0.018	0.095	0.562	0.010	0.081	0.000	0.067	2
EGHTESAD NOVIN	88	0.156	0.019	0.301	0.019	0.026	0.163	0.008	0.157	0.003	0.152	1
PARSIYAN	88	0.144	0.018	0.258	0.018	0.021	0.147	0.007	0.157	0.013	0.149	1
KARAFAIN	88	0.122	0.045	0.394	0.045	0.045	0.124	0.023	0.146	0.005	0.136	1
MELAT	88	0.413	0.007	0.007	0.007	0.185	0.425	0.003	0.074	0.007	0.068	3
PASARGAD	88	0.162	0.027	0.273	0.027	0.003	0.169	0.018	0.143	0.003	0.131	1
POST BANK	88	0.330	0.004	0.082	0.004	0.138	0.751	0.012	0.085	0.000	0.051	3
TEJARAT	88	0.088	0.009	0.152	0.009	0.232	0.405	0.009	0.081	0.002	0.066	3
SINA	88	0.161	0.020	0.020	0.020	0.010	0.206	0.015	0.144	0.001	0.142	2
SADERAT	88	0.490	0.013	0.152	0.013	0.120	0.492	0.010	0.081	0.000	0.075	2
EGHTESAD NOVIN	89	0.173	0.022	0.022	0.022	0.082	0.175	0.009	0.135	0.004	0.129	2
PARSIYAN	89	0.101	0.021	0.272	0.021	0.402	0.107	0.010	0.131	0.014	0.122	2
KARAFAIN	89	0.158	0.043	0.311	0.043	0.061	0.158	0.029	0.120	0.000	0.109	1
MELAT	89	0.379	0.009	0.242	0.009	0.186	0.441	0.039	0.078	0.012	0.071	2
PASARGAD	89	0.135	0.034	0.193	0.034	0.003	0.147	0.025	0.113	0.002	0.101	1
POST BANK	89	0.499	0.012	0.255	0.012	0.291	0.693	0.005	0.122	0.006	0.080	2
TEJARAT	89	0.310	0.010	0.169	0.010	0.245	0.352	0.004	0.085	0.014	0.063	3
SINA	89	0.202	0.025	0.025	0.025	0.067	0.231	0.019	0.140	0.001	0.136	1
SADERAT	89	0.482	0.030	0.229	0.030	0.098	0.487	0.013	0.128	0.000	0.119	2

In this study , we trained 80 percent of data and tested 20 percent of data and then after examination and training many times, the best results represented in the following:

correctnessPercentageTrain =

85.7143

correctnessPercentageTest =

87.5000

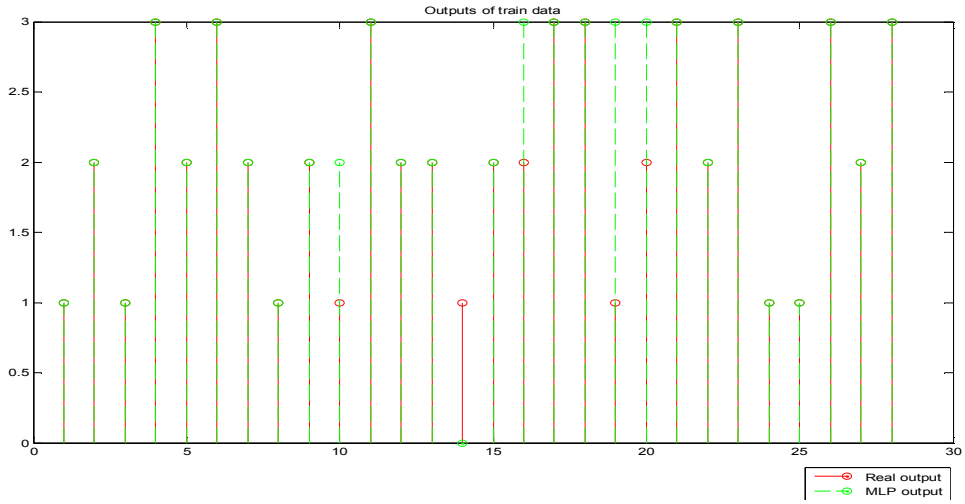

Figure4:out put of train data

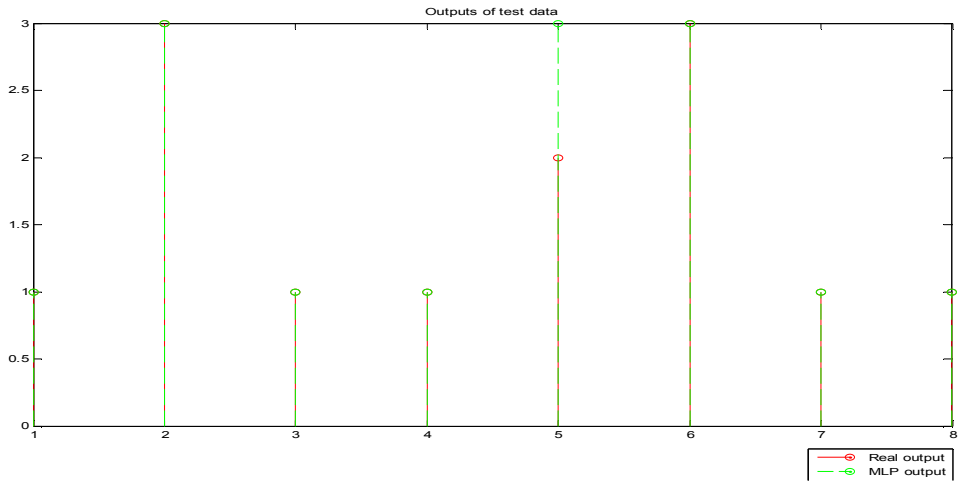

Figure5:out put of test data

Implementation results

The result of this study have been shown briefly in table3 .according to NN method applied in this study ,we have achieved the following classification with 87.5000 percent reliability .as we shown in the table 3 ,some banks have fluctuated and changed their classification during this 4 years. The resulted classification of based on the rules derived from **Artificial Neural Network** is presented in Table 3.

Table 3. Classification of sample society

Cluster Number:	**Cluser 1**				**Cluster 2**				**Cluster 3**			
Year:	*2008*	*2009*	*2010*	*2011*	*2008*	*2009*	*2010*	*2011*	*2008*	*2009*	*2010*	*2011*
EGHTESAD NOVIN		□	□		□			□				
PARSIYAN		□	□		□			□				
KARAFAIN	□	□	□	□								
MELAT								□	□	□	□	
PASARGAD		□	□	□	□							
POST BANK								□	□	□	□	
TEJARAT									□	□	□	□
SINA	□			□		□	□					
SADERAT						□	□	□	□			

References:

-Alvarez I . & Bernard S. (2005). Ranking Cases with Decision Trees: a Geometric Method that Preserves Intelligibility. IJCAI, 635-640.

-Bose, N. K.; & Liang P. (1996). Neural network fundamentals with graphs, algorithms, and applications. McGraw-Hill.

-Buchanan and Shortliffe (eds.) (1984). Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming project.

-David W (2006). Neural network credits scoring models, Comput. Operat. Res., 27: 1181-52.

-John C, Cooper B (2003). Artificial neural network versus multivariate statistics: an pplication from Economics, J Appl. Statistics, 26 (8):909-921.

-Kohavi R (1996). Scaling up the accuracy of naive-bayes classifiers: a decision-tree hybrid. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining.

-Loterman G., Brown, I., Martens, D., Mues, C., and Baesens, B. Benchamarking Regression Algorithms for Loss Given Default Modeling. International Journal of orecasting. Volume 28, Issue 1, January–March 2012, Pages 161–170.

-Oludele Awodele and Olawale Jegede (2009), Neural Networks and Its Application in Engineering, Proceedings of Informing Science & IT Education Conference (InSITE) ,pp.83-95

-Provost, F., Fawcett, T., & Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. In Proceedings of the 15th International Conference on Machine Learning, pp. 445-453.

-Provost F. J. & Domingos, P. (2003). Tree induction for probabilitybased ranking . Machine Learning, 52(30).

-Utgoff, P. (1988). Perceptron Trees – A case study in hybrid concept representation. Proceedings of the seventh national conference on artificial intelligence, 601-606.

-Wang L.X. and Mendel J.M. (1992), Generating Fuzzy Rules by Learning from Examples, IEEE Transactions on Systems, Man and Cybernetics, Volume 22, No 6., pp. 1414-1427.

Ward System Group (WSG) (1995). Neuroshell users guide, Frederick, MD: pp.101-108.

- Zhang J & Michalski, R.S. (1995). An Integration of Rule Induction and Exemplar -Based Learning for Graded Concepts. Machine Learning 21(3): 235-267