Hybrid Neuro-Genetic Principle Components Analysis as networks in Corporate Financial Evaluation

Matsatsinis Nikolaos [(a)], Loukeris Nikolaos [(b)]

Technical University of Crete
Department of Production Engineering and Management
Akrotiri Campus, 73100 Chania, Crete, Greece

(a)Associate Professor Techn. Univ. of Crete, (b)PhD student University of Essex

Abstract:-Portfolio managers, auditors and financial analysts process accounting data and financial time series of companies to determine their economic heath. Volatility in stock prices is formed in a significant proportion by accounting statements, the managerial quality in the companies and the trends in the markets. The extended amount of accounting data and the complexity of financial indices demand advanced methods of econometrics and artificial intelligence to provide the hidden information to analysts, managers and investors. Hybrid systems of neural networks with genetic algorithms optimization are able to support efficiently decisions on portfolio management, corporate management, and financial accounting. Principle Components Analysis with a Neural Network and Genetic Algorithms optimizations provides acceptable results of corporate financial classification.

Key-words: Financial Markets, Hybrid Systems, Neural Networks, Genetic Algorithms, Principle Components Analysis

1 Introduction

Corporate Financial Analysis investigates the performance of companies through their financial data, outsourced by accounting statements and stock price volatility. Corporate financial evaluation and bankruptcy prediction achieved significant results through advanced methods of artificial intelligence, data mining, and operations research, **[1]**, seeking new more effective analysis methods. Data mining from the late 1980's includes a new area of Artificial Intelligence, expressing non linear regressions in formations that simulate biological neurons, and genetics, acquiring extensive support. In the next paragraph a description of Principal Components Analysis is presented. In paragraph 3 it is described the hybrid for of PCA networks optimized by the Genetic Algorithms. In paragraph 4 the Financial Analysis indices are presented. In paragraph 5 are the results, and finally in paragraph 6 are the conclusion regarding the implementation of hybrid neuro-genetic PCA networks in Corporate Financial Analysis.

2 Principal Component Analysis with neural networks

The fundamental problem in pattern recognition is to define data features that are important for the classification, and to extract them. Transformation of input samples into a new feature space, where information about the samples is retained, but dimensionality is reduced, advancing the classification procedure.

Principal component analysis -PCA also called Karhunen-Loeve transform of Singular Value Decomposition (SVD) is such a technique. PCA aims to identify patterns in data, expressing the data in a way as to determine their similarities and differences. Data of high dimension, face difficulties to find patterns, whilst graphical representation is not available, PCA consists a forceful statistical technique for data analysis, applied in fields such as face recognition and image compression, being a common method for finding patterns in data of high dimension.

PCA finds an orthogonal set of directions in the input space, providing a way of finding the projections into these directions in an ordered way. The first principal component is the one that has the largest projection. The orthogonal directions, called the eigenvectors of the correlation matrix of input vector, and the projections the corresponding eigenvalues. PCA orders the projections, thus a reduce of dimensionality is possible, by truncating the projections to a given order. The reconstruction error is equal to the sum of the projections (eigenvalues) left out. The features in projection space, which is linear, become the eigenvalues.

PCA is normally done by analytically solving an eigenvalue problem of the input correlation function. However, **[2]** and **[3]**, demonstrated that PCA can be accomplished by a single layer linear neural network trained with a modified Hebbian learning rule (1). PCA network has p inputs (the 16 financial indices and a 17th classification index) and m<p linear output PEs, in our case the elements of confusion matrix.

$$\Delta w_{ji}(n) = \eta[y_j(n)x_i(n)-y_j(n)\sum_{k=0}^{j} w_{ki}(n)y_k(n)] \quad (1)$$

Where $i=0,1,..p-1$
$j=0,1,..m-1$
η: is the step size.

The output is given by (2)

$$y_i(n) = \sum_{i=0}^{p-1} w_{ij}(n)x_i(n) \quad (2),$$

Where $j=0,1..m-1$

PCA networks can be used for data compression, providing the best m linear features. They can also be used for data reduction in conjunction with multilayer perceptron classifiers. In this case, the separability of classes is not always guaranteed, whilst if data clusters are sufficiently separated, the network will proceed.

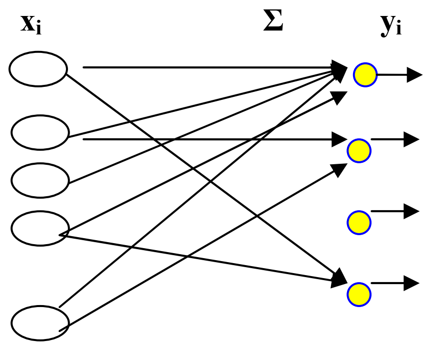

Figure 1. Principal Components Analysis network

But if the classes are on top of each other, PCA will receive the largest projections, although separability can be in some of the other projections. Linear PCA networks do not perform satisfactorily in outlying data points. Outliers will distort the estimation of eigenvectors and create skewed data projections. Nonlinear networks are better able to handle this case. The importance of Principal Components Analysis is that the number of inputs for the MLP classifier can be reduced a lot, which positively impacts the number of required training patterns, and the training times of the classifier. It is quite common a hybrid PCA/MLP net for digit recognition, where the unsupervised layer PCA at the input acts as a feature extractor and the supervised layer MLP behaves as a classifier

3 Principle Components Analysis with Genetic Algorithms Optimization

The weights of hybrid neuro-genetic PCA were chosen to be updated through On-Line learning, after the presentation of each exemplar. In contrast, Batch learning updates the weights after the presentation of the entire training set, and it was rejected. Accumulation of the gradient contributions for all data points in the training set before updating the weights is referred as batch learning, **[4], [5]**. In online learning, the weights are updated immediately after seeing each data point. The gradient for a single data point can be considered a noisy approximation to the overall gradient *G*, called stochastic (noisy) gradient descent. Online learning has a number of advantages, **[4], [5]**:

1. it is often much faster, especially when the training set is redundant,
2. it can be used when there is no fixed training set,
3. it is better at tracking non-stationary environments, the noise in the gradient can help to escape from local minima. Each one of the 16 financial indices, that are the inputs of the PCA, is not predefined hence Genetic Algorithms are used to select the significant inputs.

A Genetic Algorithm was used in
a) Processing Elements,
b) Step Size, and
c) Momentum Rate

solving the sub-problem of the optimal values for these three parameters. This form of optimization requires that the network be trained multiple times in order to find the settings that produce the lowest error. Output layer was chosen to implement Genetic Algorithms optimizing the value of the Step size and the Momentum.

4 Financial Analysis

In this research 4 different topologies of neural networks were used: Principal Component Analysis networks-PCA. Data came by 1411 companies from the loan department of a Greek commercial bank, with the following 16 financial indices:

1) EBIT/Total Assets**,**
2) Net Income/Net Worth,
3) Sales/Total Assets,
4) Gross Profit/Total Assets,
5) Net Income/Working Capital,
6) Net Worth/Total Liabilities,
7) Total Liabilities/Total assets,
8) Long Term Liabilities / (Long Term Liabilities + Net Worth),
9) Quick Assets/Current Liabilities,
10) (Quick Assets - Inventories) / Curent Liabilities,
11) Floating Assets/Current Liabilities,
12) Current Liabilities/Net Worth,
13) Cash Flow/Total Assets,
14) Total Liabilities/Working Capital,
15) Working Capital/Total Assets,
16) Inventories/Quick Assets,

and a 17th index with initial classification, done by bank executives.

Test set was 50% of overall data, and training set 50% as well.

5 Results

Genetic Algorithms were chosen to be used in each intermediate step of solutions that neural networks produced, because they offered the optimal choice of solution genes in neural network, consuming significant time periods. The method of data representation on each genetic session of offspring was on-line, because it provides always the optimal neuron on in each generation, although sometimes it is exposed to the higher risk of falling in local minima/maxima. Batch representation was rejected because confusion matrix changes on each new chromosome, without adjusting instantly, whilst it has lower risk of getting trapped in local minima/maxima. In all neural networks 500 epochs were chosen for each generation of solutions that Genetic Algorithms used to, aiming to give the necessary time for convergence in the optimal offspring set, without wasting surplus time when the optimal set of solutions was found in each repeat. In the following table 1, are presented the results of each different neural network. PCA Neural networks are presented in different topologies, where in each topology hidden layers are increased by 1. Thus the effect of hidden layers number can be presented in each different neural network's architecture.

Hence for PCA neural networks overall time follows the rule 2^i, where i, is the

number of hidden layers and the outcome is in hours.
The 10 different neural network topologies were examined thoroughly to evaluate their performance with the same data set of 1411 companies and the training set 50% of overall data. In each topology of PCA neural networks we implied initially 1 hidden layer and after convergence and results, we increased the number of hidden layers, noticing that while a change in the number of hidden layers the results such as confusion matrix, MSE varied significantly. This is expected since neural networks' performance is a 'black box' to users and neural nets cannot produce the same output, given the same input. We noticed that as the number of hidden layers was different in the same architecture, results in the confusion matrix were different for each topology in the same architecture. Analytically the performance of each neural network's architecture is in the table 1.
PCA networks were deployed in 4 different topologies with 1 layer, 2, 3 and 4 layers in each different network. Three networks produced an acceptable confusion matrix that converged, out of 14 networks. PCA nets produced very low MSE: 0.153 for 1 layer net, 0.191 in 2 layers, 0.22 with 3 layers, 0.299 in 4 layers.

6 Conclusions - Future Research.

Optimal performance was achieved by PCA with 1 hidden layer, providing a successful classification of healthy companies of mark 0 to 0 in a proportion 100% of the training set, whilst companies initially characterized by bank experts as in distress of mark 1 had a classification to 1 in a proportion of 66.66%, misclassifying companies in distress to healthy companies in a rate 33.33%, the processing offered a significantly low MSE at 0.153 with a significant value of correlation coefficient r in 0.664 indicating a medium covariation, and it converged in 2 hours and 37 minutes.

PCA hybrid networks with MLP of 2 layers and Genetic Algorithms optimization performed quite well classifying healthy companies in their rank with a proportion 100%, but companies in distress performed worst at a successful classifying rate of 33.33%, with a misclassification proportion at 66.66% for the distressed companies put falsely to healthy. Its MSE was 0.191, and correlation coefficient was low at 0.526 hence the fit of the model to the data was medium. PCA with 3 layers MLP and Genetic optimization had identical confusion matrix to PCA of 2 hidden layers, whilst the MSE was slightly higher at 0.222 and correlation coefficient r performed even worst at 0.397 with a low fitness to the data. Finally Hybrid neuro-genetic PCA with 4 hidden layers on the MLP classified correctly the healthy companies to their rank with a rate 100%, but failed completely to classify companies in distress and had put them as healthy companies. Its MSE was 0.299 and correlation coefficient was negative at -0.204, indicating a low negative correlation between the variables, as they vary in opposite ways with a low fitness to of the model to the data.

The overall performance of hybrid Principle Components Analysis with a Multi Layer Perceptron and Genetic Algorithms optimization was acceptable only for networks of 1 hidden layer, but it did not presented any interesting result. Thus different architectures of neural networks may provide more optimal performance. Hence other networks and hybrid systems should be examined thoroughly for corporate financial evaluation.

Table 1. Overall results in Neural Networks per architecture

Neural Network		Active Confusion Matrix				Performance						Time
	Layers	0->0	0->1	1->0	1->1	MSE	NMSE	r	%error	AIC	MDL	
PCA	1	100	0	33.33	66.66	0.153	0.598	0.664	17102534	95.48	69.78	2 h 37’39
PCA	2	100	0	66.67	33.33	0.191	0.743	0.526	7672387.5	2203.3	1807.6	4 h 42’00
PCA	3	100	0	66.67	33.33	0.222	0.863	0.397	18657844	2045.38	1678.16	8 h 15’00
PCA	4	100	0	100	0	0.299	1.164	-0.204	20810044	2443.46	2007.58	16 h 49’00

Bibliography

[1] Loukeris N., Matsatsinis N., ‘Corporate Financial Evaluation and Bankruptcy Prediction implementing Artificial Intelligence methods’, Journal of Business and Economics, Issue 4, Volume 3, April 2006, WSEAS Transactions

[2] Sanger T. (1989)-"Optimal unsupervised learning in a single layer linear feedforward network." Neural Networks 12, 459-473

[3] Oja E. (1982)-"A simplified neuron modeled as a principal component analyzer." J. of mathematical biology 15, 267- 273

[4] Orr G. and Leen T. (1994)- ‘Momentum and Optimal Stochastic Search’, Proceedings of the 1993 Connectionist Models Summer School, M. C. Mozer, P. Smolensky, D. S. Touretzky, J. L. Elman, & A. S. Weigend (Eds.)

[5] Leen T. and Orr G., (1994)- ‘*Optimal Stochastic Search and Adaptive Momentum*’-, Advances in Neural Information Processing Systems 6