Portfolio Selection with Predicted Returns Using Neural Networks

FÁBIO D. DE FREITAS
Secretaria da Receita Federal and
Departamento de Engenharia Elétrica
Universidade Federal do Espírito Santo
Pietrângelo de Biase, 56/308
29010-190 - Vitória E.S. - Brazil
email: freitas@computer.org

ALBERTO F. DE SOUZA e FRANCISCO J. N. GOMES
Departamento de Informática
Universidade Federal do Espírito Santo
Fernando Ferrari, S/N
29060-970 - Vitória E.S. - Brazil
email: alberto@inf.ufes.br

AILSON R. DE ALMEIDA
Departamento de Engenharia Elétrica
Universidade Federal do Espírito Santo
Fernando Ferrari, S/N
29060-970 - Vitória E.S. - Brasil
email: ailson@ele.ufes.br

ABSTRACT
The Markowitz's Portfolio Selection Model defines the return and risk variables as first-order statistical measurements, which have made this model to be known as mean-variance model. This work presents a modified version of the Markowitz's Model that uses time series prediction instead of first order statistical measurements. We have used a neural network predictor for providing an estimate of future returns, which were used as expected returns on the Markowitz's Model. The resulting new model was named prediction-quadratic deviation model. We carried out investment simulations using real data with the Markowitz's model and our model. These simulations shown that the prediction-quadratic deviation model can achieve a return 12.39% higher than the mean-variance model.

KEY WORDS
Stock Market, Neural Networks, Portfolio Optimization

1 Introduction

Portfolio Selection is one of main procedures in financial engineering and a fundamental approach for the investment strategy of diversification. The optimal portfolio selection can be defined as follows: given an universe of M securities, stocks for example, with theirs respective expected returns R and risks v, select the proportions $X_i (i = 1, \ldots, M)$ of the available wealth and invest on each stock i, minimizing the portfolio risk v_p (i.e. investment risk) for a portfolio expected return R_p (i.e. investment return).

In the original portfolio selection model proposed by Markowitz [1], estimates of the expected return and risk of each stock are taken and the percentage of each stock in the final portfolio is computed by solving the minimization problem described above. The mentioned expected return and risk variables are defined in the model as first order statistical measures and, hence, the model of Markowitz became known as *mean-variance model*. For each stock, the expected return is the mean of the historical series of returns - where the return is the variation of the stock price computed between two consecutive samples - and the risk is the variance of this series.

This work introduces a modified version of the model of Markowitz that uses time series prediction to forecast the future returns of stocks. A neural network predictor was used to obtain an approximation of future prices of stocks and these values were used to calculate future returns. These, named *predicted returns*, were utilized as the expected returns of the model of Markowitz. The quadratic deviations of the historical series of returns around the respective predicted returns where used as the associated risk of each stock. This adapted model was named *prediction-quadratic deviation model*. The use of neural networks to predict future stock returns is not new and, although there are some controversy with respect to its usefulness in this application [2, 3], our results show that it can produce better estimates for future returns than the time series mean value. We carried out simulations with 66 stocks from Brazilian BOVESPA stock exchange using both models. In these simulations, the prediction-quadratic deviation model achieved an investment return 12.39% higher than the mean-variance model with the evaluating parameters used.

This paper is organized in five sections. After this introduction, Section 2 describes the Markowitz's model and the modified Markowitz's model with predicted returns. Following, in Section3 are the experimental environment, and the procedures and data utilized. Section 4 presents the results and compares the mean-variance model with the prediction-quadratic deviation model. Section 5 closes this paper presenting our conclusions and perspectives for future work.

2 Portfolio Selection with Predicted Returns

2.1 The Markowitz's Model for Portfolio Selection

The Markowitz's model is based on the return and risk measures of a linear combination of stocks, which are obtained from the individual return and risk measures of each stock in this combination. The original and widely used expected return measure of one stock is the arithmetic mean of the series of returns of this stock, which is defined in Equation 1,

$$\bar{R} = \frac{1}{N}\sum_{t=1}^{N} R_t \,, \tag{1}$$

where $\bar{R}$ is the expected return of the stock in time $t+1$, N is the number of past times and R_t is the realized return in time t, which is defined in Equation 2.

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \,, \tag{2}$$

where P_t is the price of the stock in time t and P_{t-1}, in time $t-1$.

The risk measure proposed by Markowitz tries to reflect the uncertainties of realization of the expected return through the measurement of its time series dispersion - the dispersion of the time series around its mean value, the statistical variance, was used as risk measure. The variance of the time series of returns is defined as:

$$v = \sigma^2 = \frac{1}{N}\sum_{t=1}^{N}(R_t - \bar{R})^2 \,, \tag{3}$$

where v is the risk, which is equal to the variance σ^2 of the series of returns, N is the number of observed past times, R_t is the return in time t and $\bar{R}$ is the mean value of the series of returns.

Other measures of risk, like semi-variance [4], mean-absolute deviation [5] and downside risk [6], have been examined as alternatives to the traditional approach.

The portfolio return is obtained as:

$$\mathcal{R} = \sum_{i=1}^{M} X_i \bar{R}_i \,, \tag{4}$$

where $\mathcal{R}$ is the portfolio return, M is the number of stocks in the portfolio, X_i is the percentage of stock i ($X_i \geq 0$ and $\sum_{i=1}^{M} X_i = 1$, details below) and $\bar{R}_i$ is the expected return of stock i.

The portfolio risk is then defined as the variance of a linear combination of stocks[1] and is presented in Equation 5. The first sum represents the individual risk components for each stock and the second group of sums represents the combined risks of each pair of stocks.

$$\mathcal{V} = \sigma_P^2 = \sum_{i=1}^{M} {X_i}^2 v_i + \sum_{i=1}^{M}\sum_{j=1, i\neq j}^{M} X_i X_j \gamma_{ij} \tag{5}$$

[1] This development is detailed in [4]

In Equation 5, $\mathcal{V}$ is the portfolio risk, which is equal to the portfolio variance σ_P^2, X_i is the percentage of stock i in the portfolio, and v_i is the risk of stock i.

The term γ_{ij} in Equation 5 is the covariance of stocks i and j, which is defined as:

$$\gamma_{ij} = \frac{1}{N}\sum_{t=1}^{N}(R_{i_t} - \bar{R}_i)(R_{j_t} - \bar{R}_j) \,, \tag{6}$$

where R_{i_t} and R_{j_t} are the returns of stocks i and j in the time t, $\bar{R}_i$ and $\bar{R}_j$ are the expected returns of stocks i and j, and N is the number of observed past times. This covariance is the measure of the *interactive risk* between two stocks, and its effect on the portfolio risk $\mathcal{V}$ is crucial for obtaining a portfolio with risk lower than the individual risks of its stocks: the foundation of the *Markowitz's Portfolio Optimization Model* is this interactive risk [1]. Markowitz's optimization model minimizes the portfolio risk $\mathcal{V}$ for a desired portfolio return $\mathcal{R}$ by solving the quadratic programming problem formulated as:

$$Minimize\ \mathcal{V} = \sum_{i=1}^{M} {X_i}^2 v_i + \sum_{i=1}^{M}\sum_{j=1, i\neq j}^{M} X_i X_j \gamma_{ij} \tag{7}$$

$$Subject\ to \quad \sum_{i=1}^{M} X_i \bar{R}_i = \mathcal{R} \,, \tag{8}$$

$$\sum_{i=1}^{M} X_i = 1 \ \ and \tag{9}$$

$$X_i \geq 0, i = 1, \cdots, M \tag{10}$$

Equation 8 is the desired return constraint $\mathcal{R}$, Equation 9 guarantees total resource allocation and Equation 10 restricts the model for purchasing trades only.

The set of all portfolios with minimum risk for the desired return is obtained via parameterizing the model for $\mathcal{R}$. These portfolios are named *Efficient Portfolios* and this investment strategy is named *Efficient Diversification*. The bidimensional *locus* of the return-risk space where lies all possible efficient portfolios is denoted *Efficient Frontier*. Each set of stocks has their own efficient frontier, that depends only on the individual expected returns and risks of each stock and its time series correlation (i.e. covariances) matrix.

2.2 The Prediction-Quadratic Deviation Portfolio Selection Model

The predicted stock return $\widehat{R}$, computed with a neural network predictor (described in Section 3), is used as expected return in our model, while the portfolio return and risk are calculated using equations analogous to the equations 4 and 5, respectively.

The risk of a predicted return is quantified through the quadratic deviations of the time series of returns around the predicted return, defined as:

$$\hat{v} = \frac{1}{N} \sum_{t=1}^{N} (R_t - \widehat{R})^2 \,, \tag{11}$$

where $\hat{v}$ is the risk of the predicted return, N is the number of past times, R_t is the return in the observed time t, and $\widehat{R}$ is the predicted return.

The measure of the interactive risk $\widehat{\gamma}$ is defined as:

$$\widehat{\gamma}_{ij} = \frac{1}{N} \sum_{t=1}^{N} (R_{i_t} - \widehat{R}_i)(R_{j_t} - \widehat{R}_j) \,, \tag{12}$$

where $\widehat{\gamma}_{ij}$ is the analogous of the covariance of stocks i and j (see Equation 6) for the predicted returns, R_{i_t} and R_{j_t} are the returns of stocks i and j in time t, $\widehat{R}_i$ and $\widehat{R}_j$ are the predicted returns of the stocks i and j, and N is the number of observed past times.

We can now, after the definition of all variables, formulate the *Portfolio Selection Model with Predicted Returns* as:

$$Minimize\ \widehat{\mathcal{V}} = \sum_{i=1}^{M} {X_i}^2 \hat{v}_i + \sum_{i=1}^{M} \sum_{j=1, i \neq j}^{M} X_i X_j \widehat{\gamma}_{ij} \tag{13}$$

$$Subject\ to \quad \sum_{i=1}^{M} X_i \widehat{R}_i = \mathcal{R} \,, \tag{14}$$

$$\sum_{i=1}^{M} X_i = 1 \ \ and \tag{15}$$

$$X_i \geq 0, i = 1, ..., M \tag{16}$$

Equation 14 is the constraint of desired return $\mathcal{R}$, Equation 15 ensures total resource allocation, and Equation 16 restricts the model for purchasing trades only.

3 Methods

3.1 Predicted Returns

We have chosen to predict future stock prices and to calculate future returns indirectly using the following equation:

$$\widehat{R} = \frac{\widehat{P} - P_t}{P_t} \,, \tag{17}$$

where $\widehat{R}$ is the predicted return of the stock in time $t+1$, $\widehat{P}$ is the predicted price of the stock in time $t+1$, and P_t is the stock price in time t.

3.2 Neural Network Predictor

We have trained one neural network for each stock. The prediction model used was the *Autoregressive Model* [7] of fourth order - *Ar(4)*, implemented with a *feedforward* neural network using the *backpropagation* training algorithm [8]. In this model, the inputs are the historical prices $P_{t-3}, P_{t-2}, P_{t-1}, P_t$ and the output is the future price P_{t+1}. We have used a fixed three layers, $4-15-1$, *fully connected* network topology. The neuron's transfer function is sigmoidal with output in the $(-1, 1)$ interval. All training were conducted over 25.000 epochs, with a 0.0009 learning rate and 0.95 inertia. These training parameters were empirically determined. The details can be found in [9].

3.3 Error Measures

Two error measures were utilized. For the evaluation of the training convergence of the neural predictor we have used the *Root Mean Squared Error - RMSE*, described as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (target_i - out_i)^2} \,, \tag{18}$$

where $target_i$ is the desired value, out_i is the predicted value and N is the number of training pairs.

For the evaluation of the quality of the predictions, we used the standard *Mean Absolute Percent Error - MAPE*, defined as:

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{target_i - out_i}{target_i} \right| , \tag{19}$$

where $target_i$ is the desired value and out_i the predicted value.

3.4 Time Series

We used 66 representative stocks of the Brazilian BOVESPA stock exchange. Their time series were created with their weekly adjusted prices, observed on Wednesdays to isolate begin of week and end of week fluctuations. The training set had the initial date 1990-01-03 and final date 1995-12-27, totalizing 313 weeks. The investment simulation had initial date 1996-01-03 and final date 1996-05-22, totalizing 21 weeks.

4 Experiments and Results

4.1 Prediction of Returns

The statistics of the training performance of the 66 neural network predictors are presented in Table 1. On the top of the table we have the maximum, minimum and average values of $RMSE$ and $MAPE$ errors for all 66 neural network predictors. On the bottom, we have the variance and standard deviation of these errors. As Table 1 shows, the worst performing network presented outputs with values 38.9989% distant from their *targets* on average, while the best performing network presented outputs only 7.3023% distant from their targets on average. Together, the neural network predictors have shown an average prediction

Number of time series = 66			
Error	Max	Min	Avg
MAPE	0.389989	0.073023	0.167851
RMSE	0.081223	0.025872	0.046226
Error	Variance	Standard deviation	
MAPE	0.004531	0.067311	
RMSE	0.000112	0.010602	

Table 1. Performance statistics of 66 neural network predictors

error of 16.7851%, with standard deviation of 0.067311, suggesting that most predictors performed satisfactorily.

The predictor performance for the stock FAP4-Cofap PN of the BOVESPA stock exchange is shown in graph form in Figure 1 as an example. The time series of FAP4 prices is in solid line, the neural network prediction is in dashed line and the predicted future price is the diamond. Figure 2 shows the time series of returns in solid line; the mean return is the horizontal dashed line, while the predicted return is the diamond. With the help of the graph of Figure 1 we can verify that the neural network achieved a very good fit for this stock. The large distance between the predicted return and the mean return (Figure 2) is analyzed below.

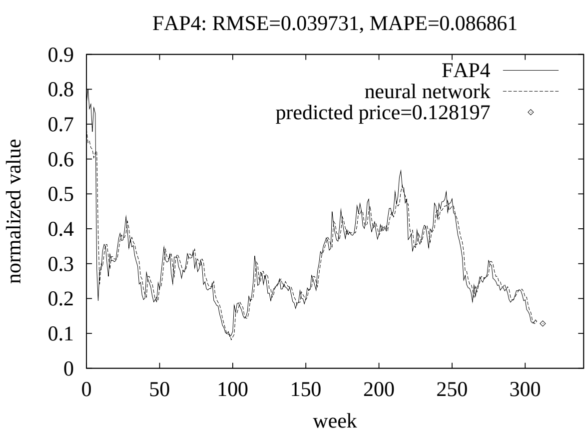

Figure 1. Neural network predictor performance for FAP4

4.2 Efficient Frontiers

Efficient frontiers were generated for the mean-variance and prediction-quadratic deviation models via solving the equations 7, 8, 9, 10 and 13, 14, 15, 16 for different values of $\mathcal{R}$ and are presented in Figure 3. Careful observation of Figure 3 shows that the prediction-quadratic deviation model allowed portfolios with return levels much higher than the mean-variance model. We believe that this is due to the fact that the return estimation method used by the prediction-quadratic deviation model allowed predicted return values far away from the mean return, as is the case in Figure 2, which resulted in better portfolios.

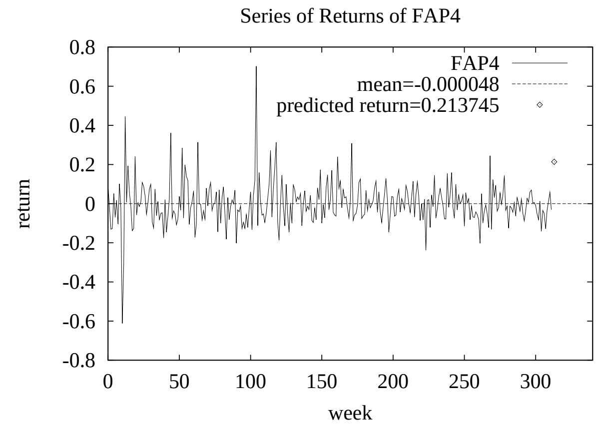

Figure 2. Series of returns and predicted return for FAP4

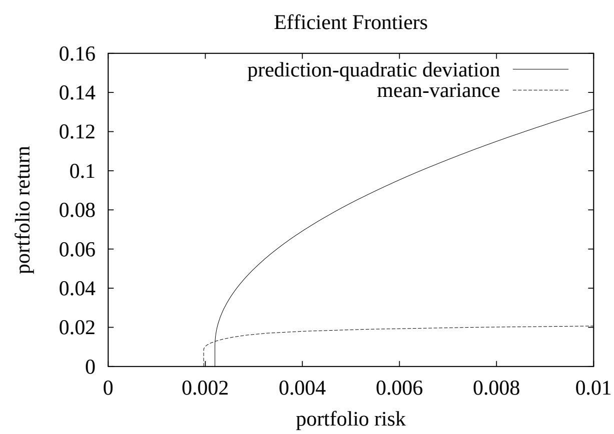

Figure 3. Efficient Frontiers for the prediction-quadratic deviation and mean-variance models

4.3 Benchmarking

On the last step of our experiments, we have chosen one representative efficient portfolio for each model in the risk level near the intersection of their efficient frontiers. Hence, we selected portfolios with risks $\mathcal{V} \simeq \widehat{\mathcal{V}} \simeq 0.002$. We evaluated the simulated investment for 21 weeks between 1996-01-03 and 1996-05-22, ranking the portfolios returns. In these simulations, the prediction-quadratic deviation portfolio showed a better performance in 19 of the 21 weeks, achieving a return 12.39% higher than the mean-variance portfolio on average. Figure 4 shows the evolution of the returns of the mean-variance and prediction-quadratic deviation portfolios. The similarity of the two curves of Figure 4 is due to the fact that both models selected the same 46 stocks from the 66 available, with the prediction-quadratic deviation model allocating higher proportions of the stocks with higher predicted returns. If we assume that the above cited similarity of stocks used on the solutions was dictated by the risk level chosen, we can conclude that these stocks have strong first order components in their time series of returns. This fact also suggests that, for higher levels of risk, the solutions (i.e. stocks used) will differ because the prediction-quadratic deviation model will choose stocks with higher order components that are not captured by the mean-variance model. This can be appreciated with the help of Figure 3, which shows that our model produces increasing returns for increasing levels of risks, while the mean-variance model almost stop producing better returns for risks higher than 0.003.

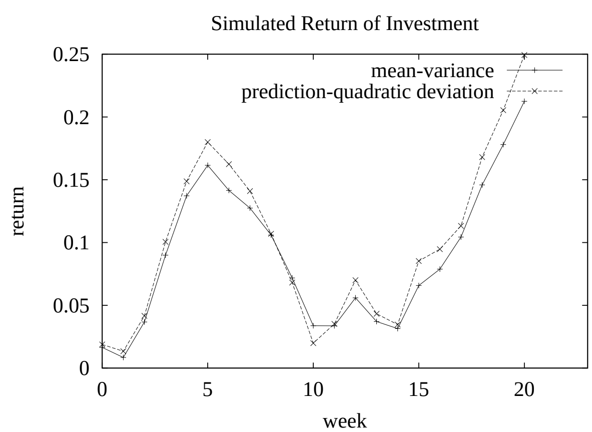

Figure 4. Simulation of Investment for the prediction-quadratic deviation model and mean-variance model

5 Conclusions and Future Work

In this paper we focus on one of the fundamental problems in financial markets, Portfolio Selection, proposing a variant of the Markowitz's model that uses predicted returns as expected returns. Our experimental evaluation of this new model, presented in this paper, has been developed in an application oriented approach, analyzing the time series of 66 representative stocks of the Brazilian BOVESPA stock exchange. Our experiments show that our model, named prediction-quadratic deviation model, achieves better results than the mean-variance model for the same level of risk. The prediction-quadratic deviation model achieved an average return 12.39% higher than the mean-variance model, defeating the classical model in 19 of the 21 weeks evaluated. This happens because the the prediction-quadratic deviation model selects higher proportions of stocks with predicted returns higher than the mean returns used in the original model, and also because it can pick solutions on regions of the return-risk space that are unknown for the classical model.

The main contributions of this work are introducing an alternative approach for the expected return estimation - the use of artificial neural network predictors - and incorporating predicted returns to the portfolio selection model of Markowitz. Future research includes the comparison of our model with other portfolio selection models [10]; exploration of more sophisticate neural network training paradigms, like Genetic Algorithms [11], Genetic Programming [12, 13] and Particle Swarm Optimization [14]; and the use of multivariate modeling (i.e. to add extra inputs to the neural network) to enhance predictor performance.

References

[1] H. Markowitz. Portfolio selection. *The Journal of Finance*, VII(1):77–91, Mar. 1952.

[2] J. Moody. Economic forecasting: Challenges and neural solutions. In *International Symposium on Artificial Neural Networks*, Hsinchu, Taiwan, 1995. Keynote Talk.

[3] R. Sitte and J. Sitte. Analysis of the predictive ability of time delay neural networks applied to the s&p500 time series. *IEEE Transactions on Systems, Man and Cybernetics - Part C*, 30(4):568–572, 2000.

[4] H. Markowitz. *Portfolio Selection: Efficient Diversification on Investment*. John Willey & Sons, New York, 2 edition, 1991.

[5] H. Konno and H. Yamazaki. Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science*, 37(5):519–531, May 1991.

[6] A. M. Duarte jr. A comparative study of downside risk and volatility in asset allocation. *Investigación Operativa*, 4(3):213–228, 1994.

[7] H. White. Economic prediction using neural networks: The case of IBM daily stock returns. In *Proceedings of The IEEE International Conference on Neural Networks*, 1988.

[8] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, (323):533–536, 1986.

[9] F. D. Freitas. Computational environment for portfolio selection using predicted returns. Master's thesis, Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo - PPGI/UFES, 1997.

[10] M. Gilli and E. Këllezi. Heuristic approaches for portfolio optimization. *VI International Conference on Computing in Economics and Finance*, 2000.

[11] A. Likartsis, I. Vlachavas, and L. H. Tsoukalas. A new hybrid neural-genetic methodology for improving learning. *Proceedings of the 9th International Conference on Tools with Artificial Intelligence (ICTAI '97)*, 1997.

[12] H. Iba and T. Sasaki. Using genetic programming to predict financial data. In *Proceedings of the Congress of Evolutionary Computation*, 1999.

[13] V. Maniezzo. Genetic evolution of the topology and weight distribution of neural networks. *IEEE Transactions on Neural Networks*, 5(1), january 1994.

[14] J. Salerno. Using the particle swarm optmization technique to train recurrent neural model. *Proceedings of the 9th International Conference on Tools with Artificial Intelligence (ICTAI '97)*, 1997.