Application of Artificial Intelligence Approach to Portfolio Selection and Management

Vesna Bogojevic Arsic

Faculty of organizational sciences, University of Belgrade
154 Jove Illica, 11000 Belgrade, Serbia
bogojevic@fon.bg.ac.yu

Abstract. This paper is trying to explain the portfolio selection and management problem as a one of the most important problem in finance. Modern portfolio theory and Markowitz efficient frontier start with a set of assets (securities) and generate an optimal weight combination for the optimal risky portfolio that lies on the efficient frontier. The first step is to identify which assets (securities) should be selected from a pool of available assets. The second step is to predict the expected returns for a better utilisation of the Markowitz Efficient Frontier. Artificial Intelligence (AI) techniques are widely used in various fields of finance, which motivated the use of these techniques to find a quantitative and systematic method to construct an optimal portfolio. The Genetic Algorithms Technique (GAs) is one of the AI techniques being successfully used to solve complex optimisation problems. GAs are deployed in this research to select the optimal portfolio based on maximising a composite objective function that maximises return, minimises risk and minimises cross-correlation between assets in the candidate portfolio. GAs are tested on two stock markets, the US stock market, represented by a pool of 40 US companies. In this market the generated optimal portfolio based on Genetic Algorithms was able to provide higher risk adjusted returns than the market index, in both the training period and the testing period. The Neural Networks technique is used to provide a better estimate for expected returns than the conventional historical average. It was found that, even in bearish market periods, the optimally selected portfolio, which was weekly managed using Neural Networks, was able to generate positive returns utilising the Markowitz Efficient Frontier. The research result has demonstrated the usefulness of applying the proposed AI approach represented by Genetic Algorithms and Neural Networks in active portfolio selection and management. In: Shebani K (ed). Proceedings of the 1st International Conference on Applied Operational Research – ICAOR (2008), pp 77–86. Lecture Notes in Management Science Vol. 1. ISSN 2008-0030.

Keywords: asset allocation, portfolio selection, artificial intelligence, genetic algorithms, neural networks.
1 Security Selection and Asset Allocation

Security selection and asset allocation (Bernstein, 1995) refers to the process of determining optimal allocations for the broad categories of assets (such as stocks, bonds, cash, real estate etc.) that suit the investor’s time horizon and risk tolerance. Simply stated, asset allocation is an investment strategy where the investor decides on dividing money among stocks, bonds or other asset classes. When spreading the money among a number of different types of investments, this can reduce overall risk and improve the returns of the investment portfolio. The goal of asset allocation is to achieve the highest return for the acceptable level of risk or, alternatively, the lowest risk for a needed rate of return. By combining assets with different characteristics in a portfolio, an investor can achieve higher returns with lower risk over the long-term. Adding high-risk asset classes and investments to a portfolio may seem risky, but the likely net effect will be to both increase returns and lower the risk of the portfolio.

The investment process, which determines how the portfolio is managed, is described by an investment policy that consists of the following major sets of decisions, which form a hierarchy:

- Choice of asset classes in which to invest.
- Choice of normal asset class weights that remain unchanged over time. The weights are often determined by an optimisation procedure designed to generate an expected risk (variance) and return appropriate to the circumstances of the particular investor (choice of asset mix).
- Security selection and market timing. The choice of individual securities within each asset class and adjusting the asset class weights from their normal values on a short-term basis.

The focus of the paper is on the third point, i.e. security selection for the stocks asset class. Following a top down investment strategy, reaching to the industry level, and then picking stocks in those industries, needs a quantitative technique to provide guidance so as to obtain the best possible combination of assets that maximises the investment objective – maximum return, minimum risk and minimum cross correlation between assets in the selected portfolio.

1.1 Problem definition

The construction of investment portfolios poses a problem that involves multiple objectives. The main goal is to select a set of stocks that are expected to generate profits in order to form the investment portfolio by considering a great number of possibilities and parameters. Once the stocks are chosen, there is the problem of defining the percentage that must be invested in each asset, which is also called the asset allocation problem. When defining the vector of weights, both the expected return on each asset and the risk that the investor is willing to take on the
investment are considered. After the stocks and the percentage of investment have been defined, it becomes necessary to manage the portfolio by means of a periodic evaluation of its performance and of the alteration of the weight values.

Exhaustive portfolio selection is not practically feasible and requires huge combinations. Assuming 50 assets to start with, selecting a portfolio of 10 assets requires \( C^{50}_{10} = (3.727 \times 10^{16}) \) computations. An effective, efficient and practical optimisation technique is required to solve this complex combinatorial optimisation problem and this is where the genetic algorithms play a role. In order to effectively manage the selected optimal portfolio using the Markowitz efficient frontier, a reasonable estimate is required for the asset returns comprising the portfolio. Here is where the neural network part comes into play. This technique also needs to be tested on real market data to verify its validity. The problem definition can be summarised by the following needs:

- the need to select a pre-specified number of assets comprising a portfolio from a pool of a large number of available assets (genetic algorithms)
- the need to find the optimal weight allocations for each asset based on a risk adjusted basis (Markowitz efficient frontier)
- The need to forecast, with reasonable accuracy, the asset returns in the portfolio for proper portfolio management (neural networks).

1.2 Objective and methodology of the paper

The main purpose of this research paper is to investigate the use of artificial intelligence (AI) techniques in the portfolio selection and management problem. Two well known AI techniques are used, namely genetic algorithm (GA), to identify an optimal set of assets (securities) that will comprise the portfolio, starting from a universal set of assets and neural networks (NN) for the prediction of the assets in the portfolio to provide a better estimate of the assets expected returns. Then the Markowitz efficient frontier module (MEF) will be used to find the optimal set of weights for the optimal portfolio selection defined through the GA.

Another key objective of the research is to apply these techniques to real market data and, in this regard, the US (Y2000–Y2002) stock market data are used for training and testing the proposed GA-based portfolio selection model.

The neural network predictor model is only applied to the US Stock Market, since it requires large data for training. Five year weekly historical returns (1998–2002) are used to train the NN models.
2 AI Techniques as Applied to Finance and Investment

AI (Haykin, 1998) can be defined as the branch of computer science that is concerned with the automation of intelligent behaviour, as well as the study of how to make computers do things that, at the moment, people do better. Research areas that fall under the AI category include (Haykin, 1998); artificial neural networks (simply neural networks), genetic algorithms, fuzzy logic, natural language processing and understanding, computer vision and speech recognition and synthesis.

Recently, AI has been successfully applied in various fields of finance and investment, namely; stock market prediction, financial decision making, assessing the fundamentals for financial analysts, analysing the financial health of firms, bankruptcy prediction, bond rating, risk assessment, predicting consumer credit performance and credit scoring, stock picking, testability of arbitrage pricing theory, pricing initial public offerings, commodity trading, predicting options volatility, managerial forecasting and portfolio selection and management.

The AI tools that were extensively applied in finance are mainly NN and GA. There has been an international interest in applying those AI techniques in portfolio management and stock market prediction.

Noever and Baskaran (1994) investigated the problem of predicting trends and prices in financial time series, conducting experiments on the S&P500 stock market. Mahfoud and Mani (1996) addressed the general issue of predicting the future performance of individual stocks. Their work is particularly relevant, as they compare GAs and NNs as applied to financial forecasting. On the other hand, Kimoto and Asakawa (1990) used NN technology to develop a TOPIX buying and selling prediction system. TOPIX is a price-weighted index of all stocks listed on the first section of the Tokyo Stock Exchange.

Frieslebeln (1992) used a simple feed-forward NN using past and present data to predict the value of the FAZ-Index, which is the German equivalent of the DJIA. However, Chapcot (1992) used GAs to track the FTSE100 UK index. This project was concerned with passive portfolio selection using GAs and quadratic programming techniques. Landt (1997), on the other hand, used standard feed-forward NN for stock prediction of selected companies from the Amsterdam Stock Exchange.

Chan et al. (1996) used GAs in a multi-stage portfolio optimisation system applied to the Shenzhen stock market of China and Armano et al. (2002) presented a hybrid approach to Italian stock market forecasting (COMITA).

Stopiglia et al. (1996) have described a NN-based aid to the financial analysis of companies, which is in current use for portfolio management with a view to long-term investments, within the Groupe Caisse des Dépôts, Lazo et al. (2000), in their work, presented the development of a hybrid system based on GAs and NNs for the selection of stocks and portfolio management applied to the Bovespa Index.
3 Conceptual Framework and Research Design

The following block diagram (Figure 1) shows the conceptual framework of the overall proposed model.

![Diagram](image)

**Fig. 1.** Block diagram of the proposed genetic – neural model

The proposed system comprises three key modules:

1. GA for the selection of the assets (stocks) that will form the investment portfolio
2. Markowitz efficient frontier optimiser, to find the optimal weight allocation for each asset on a risk adjusted basis
3. NN model for the prediction of asset returns, to provide a better estimate for the expected returns vs. the historical arithmetic average.
Due to the fact that NNs require large historical data for proper training and accurate historical returns, so they were applied to US market stocks. Five year weekly historical returns for the optimal assets obtained from the GA module (240 sample points for each) are used to train the NN model.

4 Applying Genetic Algorithms to Portfolio Selection

The key question to which the paper is trying to give the answer is; what is the optimal set of assets to be selected in a portfolio among a pool of market assets? Given a pool of various securities (40 companies in the case of the US stock market), what portfolio combination is optimal in some sense. Quantifying the term ‘optimal’ should be accompanied with a definition of an objective function to be maximised.

The objective function selected is the one that maximises returns, minimises risks and minimises cross correlation between different assets selected in the portfolio and is given by:

$$f(\text{portfolio}) = \prod_{i=1}^{n} \prod_{j=1}^{n} R_i (1 - \sigma_i)(1 - \rho_{ij})$$

where:
- $R_i$ - return on the asset at ith gene of the chromosome (monthly average)
- $\sigma_i$ - inherent risk of the asset at ith gene of the chromosome
- $\rho_{ij}$ - correlation coefficient of the asset at ith gene with the asset at jth gene of the chromosome
- $n$ - portfolio size (number of assets in the portfolio).

This makes it possible to include the most promising assets in the chromosome, where each chromosome is a certain portfolio candidate and the gene value represents a particular security (asset).

5 Application to the US Stock Market

Table 1 shows the optimal portfolio selection of securities based on the GA module.
Table 1. GA optimal portfolio: US market case

<table>
<thead>
<tr>
<th>Company</th>
<th>Industry</th>
<th>Average return</th>
<th>Risk</th>
<th>Beta VS S&amp;P500</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>Information technology</td>
<td>0.2452%</td>
<td>5.48%</td>
<td>1.0627</td>
</tr>
<tr>
<td>Philip Morris (Altria)</td>
<td>Consumer products</td>
<td>0.889%</td>
<td>5.20%</td>
<td>0.3528</td>
</tr>
<tr>
<td>American Electric Power Company (AEP)</td>
<td>Large public utility holding company</td>
<td>0.1574%</td>
<td>4.54%</td>
<td>0.3702</td>
</tr>
<tr>
<td>Merrill Lynch</td>
<td>Investment banking</td>
<td>0.6261%</td>
<td>6.79%</td>
<td>1.516</td>
</tr>
</tbody>
</table>

The optimal portfolio selected shows diversification in various industries. The selected portfolio has the best possible 5-asset combination in terms of the composite measure (objective function) that maximises return, minimises risk and minimises cross correlation between the 5 assets in the portfolio.

Using Mean Variance Optimisation (MVO) and Markowitz Efficient Frontier (MEF), the optimal risky portfolio can be obtained, based on investor preference, by identifying a point on the efficient frontier. MEF for the five selected assets portfolio is shown in Figure 2.

Fig. 2. Markowitz efficient frontier: the US stock market case
This important figure shows the plot of the five assets in the GA optimal portfolio for the training period Y2000-Y2001. In addition, three major US market indices are plotted in this period. Table 2 shows Weight distribution for the US optimal portfolio.

<table>
<thead>
<tr>
<th>Company</th>
<th>Optimal Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>14.7%</td>
</tr>
<tr>
<td>Philip Morris (ALtria)</td>
<td>32.8%</td>
</tr>
<tr>
<td>Proctor and Gamble</td>
<td>0.68%</td>
</tr>
<tr>
<td>American Electric Power Company</td>
<td>31.42%</td>
</tr>
<tr>
<td>Merrill Lynch</td>
<td>20.36%</td>
</tr>
</tbody>
</table>

### 6 Summary

Investigating the application of AI techniques in portfolio selection and management was the main subject of this paper. The paper mainly addresses the problem of portfolio selection, complemented by a prediction tool to provide more reliable estimates for the assets expected returns in order to manage the portfolio using the celebrated Markowitz efficient frontier.

One of the AI techniques that has been deployed in the current research is Genetic Algorithms (GAs), which has been used by many researchers in various disciplines as a powerful optimisation tool capable of solving complex optimisation problems of a combinatorial nature.

Weekly returns are selected as the basis of computations, which represent a compromise between the instantaneous nature of daily returns and the coarse nature of monthly or quarterly returns, so as to reasonably represent the variability of returns and provide meaningful correlation values. One market has been tested, the mature US stock market. A success criterion is that the optimal portfolio outperforms the market in the testing period while showing some positive correlation with the market, which means that the portfolio is a good representative of the market but with better average returns on a risk adjusted basis. The success of the efficient optimal portfolio in the testing period is highly dependent on the extent to which the expected returns match actual returns. If there is a 100% match, which is practically never the case, then the efficient optimal portfolio in the training period will enjoy more or less similar performance in the testing period.

Accordingly, having a prediction tool that would give some reasonable prediction better that the conventional histriovichal averages would indeed improve the portfolio performance if NN were adopted for that purpose.
For the US market data, 40 companies from the Large-Mid Caps Blue Chips are used to represent the pool of available stocks. The GA algorithm has optimally selected the following companies to be included in the portfolio: IBM, Philip Morris (now Altria), American Electric Power (AEP), Proctor and Gamble and, Merrill Lynch. For the training period, the portfolio outperformed the market index, where the average returns for the optimal portfolio in this period was 0.6% and −0.1517% for the market index. For the testing period, the average optimal portfolio returns were −0.37% at 3.77% risk, whereas the market returns averages −0.5% at 2.7% risk. The NN based portfolio, which was weekly managed based on forecasted returns, showed a positive return of 0.0525% in this bearish market at 2.73% risk. The results show that the NN-based technique can outperform the market in a bearish market period and produce positive returns.

The several problems should be researched. First, the transaction costs, the costs that had not been taken into considerations and could play a role, particularly when weekly managing (changing weights) of the optimal portfolio. Taking that into consideration would give a more realistic picture regarding the portfolio’s performance. Second, the benchmarking of the proposed optimal portfolio, not only to the market index but also to other practical portfolios constructed by professional asset management firms, will validate the usefulness and power of the proposed approach. Third, the approach should be tested on more data and different periods to validate the consistency of performance. Also, testing daily, monthly and possibly quarterly returns vs. the selected weekly returns would provide an insight into the effect of the sampling window on the validity of the overall approach. Fourth, the different types of technical indicators as inputs to NNs have to be investigate as well as the effect of these new inputs on the performance of the NNs and, hence, the overall performance of the optimal portfolio. Finally, for the GA module, different objective functions could be tested and compared with well-known composite portfolio performance measures, like Sharpe, Treynor and Jensen (Reilly and Brown, 1997). Also, adding another dimension to the optimisation by including terms that maximise the correlation with the market index would be beneficial for passive portfolio management purposes.

References


