Investment portfolio management and forecasting the return on assets based on artificial intelligence methods (neural analysis and genetic algorithm)

Sunchalin A.M.
Financial university under the government of the Russian Federation,
Bauman Moscow state technical university
Moscow, Russia
AndreySunchalin@mail.ru

Sunchalina A.L.
Bauman Moscow state technical university,
Moscow, Russia
AndreySunchalin@mail.ru

Ivanyuk V.A.
Financial university under the government of the Russian Federation,
Moscow, Russia
ivaver6@gmail.com

Abstract — The purpose of the study is the methodology of construction and management of the investment portfolio. Due to the fact that the transition to market economy in Russia has not been made until the end of the XX century, the Russian equity market is still emerging, distinguished by a high degree of volatility and a relatively short time horizon of historical information. Taking into account the crises that have happened in our country over the period from 1991 to 2019 (economic crisis of 1998, financial and economic crisis of 2008, and monetary crisis of 2014), the Russian equity market is characterized by instability and therefore requires a deeper analysis. In this regard, this study appears to be relevant, as it discusses the models for constructing an investment portfolio, as well as its practical realization and software implementation in MS Excel. The paper also reviews the methods of artificial intelligence (neural networks and genetic algorithm), on the basis of which the model of forecasting the return on assets is built.

Keywords — digital technologies, neural networks, genetic algorithm, investment portfolio management

I. INTRODUCTION

Currently, the term "portfolio" is widely used and, depending on the context, may be applied to "any set of something with similar properties", and it is not necessarily related to finance (for instance, a project portfolio or a business portfolio).

In this paper, the portfolio of securities refers to all securities held by the investor. Typically, a portfolio consists of several parts: bonds, ordinary shares, preferred shares, etc. Investment objectives and investor profile (aggressive, conservative, etc.) determine the composition of the portfolio.

When putting together a portfolio, there is an increased focus on such important characteristics as risk, return and holding period[1].

Portfolio risk quantitatively reflects the probability of deviation of actual performance from expected values, that is, estimates the amount of extraordinary loss (income).

Portfolio return reflects the ratio of the relative change in the portfolio value over the given period to the portfolio value at the beginning of the period.

Another characteristic of the portfolio is the holding period – the period of time during which the portfolio remains unchanged. This characteristic is constant for a particular portfolio.

When considering this set of portfolio characteristics, one can see that it is only the holding period of a particular portfolio that depends on the investor, while the risk and the return are objective indicators that depend solely on the market situation.

In analyzing the investment process, the following five stages of the investor's activity are highlighted: choice of investment policy, stock market analysis, investment portfolio construction, revision of the constructed portfolio, and evaluation of its performance.

In the first stage – when choosing the investment policy – the goals and scope of investment are determined. This is where a global assessment of the return-risk ratio of the expected assets takes place. After a thorough study of the preferred industry, its position in the stock market and the subsequent selection of a number of potential shares for the future investment portfolio, one moves to the next stage.

The second stage of the investment process is the evaluation of the selected stocks. The purpose of this analysis is to find undervalued or overvalued assets, as they are more likely to undergo changes in the near future in reference to their
quotations. At this stage, the two most powerful tools of the stock market analysts come into operation: fundamental analysis and technical analysis.

Technical analysis is generally concerned with the study of historical data on the prices of stocks of certain companies, analyzing possible cyclical patterns, changes in quotations, overall dynamics and other indicators. In this manner, the current trend is identified, using which it is possible to predict the expected behaviour of quotations for the nearest period. The fundamental analysis, for its part, is concerned with the search for incorrectly valued securities of a certain company, with the risk assessment, as well as with forecasting future dividends and income. A competent investor always relies on the entire body of results obtained from both approaches, without neglecting either of them.

Further, in the next stage, it is necessary to construct a portfolio of valued assets. This stage is the main one and involves a number of certain difficulties for the investor. Selectivity, or macroforecasting, is the first of them [2]. The investor needs to predict future stock prices on the basis of their own analysis in order to correctly include the asset in the portfolio. Choosing the right moment, or macroforecasting, is the next difficulty [5]. Finally, there is diversification – minimizing the risk of the investment portfolio by constructing it in a specific way.

After this stage, the management of the investment portfolio begins. Now the investor can either maintain the constructed portfolio and keep it unchanged or reform it according to the changes in investment goals, the situation on the stock market or eventual loss of the portfolio optimality. In the latter case, the investment process is moving into the fourth stage – the revision of the securities portfolio. In this stage, the investor often has to repeat the first three steps so that the new portfolio has a return not lower than the previous one.

Last, but not least comes the evaluation of the investment performance which by no means takes place at the last moment and should be carried out on a regular basis after the portfolio has been constructed. Periodic assessment of the return and risk of the portfolio allows to avoid unnecessary losses, as well as to optimize and change the pool of assets in time, if necessary.

It is the most important and crucial stage for the investor because it is at this moment that the results obtained are being analyzed and compared with the intended goals. Based on the returns available for each asset, the investor can calculate the current rate of return on the entire portfolio. Next, the method of calculating the parameters of the investment portfolio proposed by Harry Markowitz – the founder of modern portfolio theory – is used [9]. For example, for a portfolio of N securities we have the following rule of return formula (1):

\[ r_p = \sum_{i=1}^{N} X_i r_i = X_1 r_1 + \cdots + X_N r_N, \]  

where \( r_p \) – the calculated return on the investment portfolio, \( r_i \) – the return on the \( i \)-th asset, \( X_i \) – the share of the investments in the \( i \)-th asset in the aggregate portfolio investments. The same formula can be used to calculate the expected return of the portfolio by taking for each asset its expected return.

Similarly as with the return, the risk of the securities portfolio can be calculated, provided the data on the risk of each asset are available. Thus, for a portfolio consisting of N assets, the formula (2) is valid:

\[ \sigma_p = \left[ \sum_{i=1}^{N} \sum_{j=1}^{N} X_i X_j \sigma_{ij} \right]^{1/2} \]  

where \( \sigma_p \) – the calculated risk of the investment portfolio, \( X_i \) – the share of the investments in the \( i \)-th asset in the aggregate portfolio investment, \( \sigma_{ij} \) – the covariance of returns on the \( i \)-th asset and the \( j \)-th asset, calculated as (3)

\[ \sigma_{ij} = \rho_{ij} \sigma_i \sigma_j, \]

where \( \rho_{ij} \) and \( \sigma_j \) – the risks of the \( i \)-th and the \( j \)-th assets, respectively, \( \rho_{ij} \) – the coefficient of correlation between the returns on the \( i \)-th asset and the \( j \)-th asset, confined in the interval from -1 to 1.

The following formula (4) is used to calculate the coefficient of correlation between the returns on two assets, by convention x and y, over the period of N years:

\[ \rho_{xy} = \frac{\sum_{i=1}^{N} X_i y_i - \left(\sum_{i=1}^{N} X_i \right) \left(\sum_{i=1}^{N} y_i \right)}{\sqrt{\sum_{i=1}^{N} X_i^2 - \left(\sum_{i=1}^{N} X_i \right)^2} \sqrt{\sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i \right)^2}} \]  

where \( x_i \) and \( y_i \) – the returns on securities x and y respectively, at the time \( i \), \( i=1, \ldots, N \).

This completes the calculations regarding the existing investment portfolio, at which point it is necessary to draw the right conclusions for taking follow-up actions on the portfolio management based on the obtained results. It is essential to recognize whether the portfolio is optimal and, if not, how to achieve its optimization [3,4].

According to the Markowitz portfolio theory (MPT), in order to optimize the securities portfolio, it is necessary to solve the target optimization problem that maximizes the expected return on the entire portfolio or, in the framework of the inverse problem, minimizes the overall risk given the constraints. Then the problem is divided into two types of quadratic programming problems.

1. Maximum return portfolio – system (5):

\[
\begin{align*}
\{ r_p &= \sum_{i=1}^{N} X_i r_i \to \text{max} ; \\
\sigma_p &\leq \sigma_{max} ; \\
\sum_{i=1}^{N} X_i &= 1, \ X_i \geq 0.
\end{align*}
\]

II. RESEARCH METHODOLOGY

Neural networks have recently gained in popularity among scholars. In particular, neural networks are widely used in forecasting, classification, clusterization, and modelling [11]. Mainly, this is attributable to one remarkable property of neural network-based models – they are able to "see" non-linear relationships as opposed to many models, which mostly have linear relationships only.

An artificial neural network differs from any other program in that it can be trained, and it will produce a result that depends not only on the input data of the current experience but also on the past experiences [10].

**Artificial intelligence methods based on neural networks**

A structural unit of an artificial neural network is an artificial neuron, the working principle of which generally resembles that of a biological neuron. Let us consider the McCulloch-Pitts neuron, which was proposed in 1943.

The **artificial neuron** has the following three main components, which are graphically represented in Fig. 1[6].

1. The input signal is represented by synapses (or connections), each of which is characterized by its own weight or strength. In particular, the signal \( x_j \) at the input of a synapse \( j \) associated with the neuron \( k \) is multiplied by the weight \( w_{kj} \). Unlike synapses of a biological neuron in the brain, the synaptic weight of an artificial neuron can be both positive and negative.

2. By an adder, the input signals weighted relative to the corresponding synapses of the neuron are combined. In other words, the values \( x_i \) are multiplied by the corresponding weight \( w_{ki} \), and the results are added up. This operation can be described as a linear combination.

3. The activation function limits the amplitude of the neuron output signal. This function is called the compression function. Usually, the normalized range of amplitudes of the neuron output lies within the interval \([0,1]\) or \([-1,1]\).
From a mathematical point of view, a neural network is a multiple superposition of a polynomial of sigmoid-like functions \[ F = k_0 \sigma(k_1 \sigma(k_2 \sigma(k_4 x_{ij})) + k_5 \sigma(k_6 \sigma(k_7 k_8 x_{ij}))) + \ldots \]

Here is an example of building a model for one neuron (Figure 3).

The simplest way to train a neural network is programming by examples: neural networks select coefficients in such a way that the result output is in line or close to the output of empirical data with the corresponding input sets. The coefficients are selected in such a way that as a result of the model run, the least possible errors for the "inputs-output" set are obtained.

There are several approaches to the neural network training: with a teacher (when the values that the network should give out are known), without a teacher and a mixed approach.

Neural network approach in forecasting has its pros and cons.

The advantages of using neural networks include:

- Versatility. Neural networks do not depend on the properties of the input data, for them, there is no requirement for a certain type of distribution of the source data (in many models there is an assumption about the normality of the distribution of the source data, while in neural network models this assumption is not required). There is either no requirement for linearity of target functions (this is the main advantage of neural networks – they can "see" non-linear relationships, unlike most standard models based on linear relationships, for example, CAPM which is used to assess market risk)[12]

- Simplicity. The use of a neural network does not require specialized training for practical application.

- There is no "curse of dimensionality" problem. They are able to model dependencies in the case of a large number of variables.

- The process of finding dependence is accelerated by processing of data simultaneously by all neurons.

Nevertheless, neural networks have proven their effectiveness in forecasting tasks. Using the example of the US dollar, we will build an autocorrelation neural forecast. Suppose that tomorrow's dollar rate depends on the previous five values. Let's build a time series with a 1-day shift and estimate autocorrelation (Figures 4-5).

Let us calculate the coefficient boundaries of the neuron as \[ \left[ \frac{-3}{x_{max}}, \frac{3}{x_{max}} \right] \] and \[ \left[ x_{min}, 6x_{max} \right] \] (Figure 6).
Let us determine the initial values of the coefficients and create the formula for the neural network (Figure 7).

![Fig. 6 Neuron’s coefficient boundaries](image)

Let us calculate the network error as the sum of squared deviations (Figure 8).

![Fig. 7 The initial coefficient values](image)

Let's continue the neural network formula with one value forward, and generate the forecast data as a network output taking into account the root-mean-square deviation (Figure 9).

![Fig. 8 Network error calculation](image)

Next, we obtain the forecast and assess its quality by constructing a graph (Figure 11).

![Fig. 9 Data for the forecast](image)

![Fig. 10 Selection of the neural network coefficients](image)

![Fig. 11 Forecast schedule](image)
Artificial intelligence methods based on the genetic algorithm (GA)

Artificial intelligence methods based on the genetic algorithm (GA) – an algorithm designed to optimize the functions of several variables. It should be borne in mind that a number of functions of several variables may contain several extrema, in other terms – a set of solutions. Such functions are called polymodal.

Optimization problems of this kind of functions can be solved with the help of two approaches:

The first approach is a method of exhaustive search of function values for all possible existing or valid argument values. This method is applicable for a limited time period only outside the set of rational numbers.

The second approach is applicable to the set of rational numbers (R arguments). It involves a limited number of comparisons, as well as function values for a random selection of arguments and is inherently stochastic. At the same time, the larger the selection of arguments, the higher the probability of finding the correct solution (or the desired optimum). This method was improved by J. Holland in 1975 based on the evolutionary processes occurring in nature and became known as the genetic algorithm.

The essence of the genetic algorithm is to simulate the process of evolution of individuals of the same species of animals in a confined space: the result is the survival of the fittest. The advantages of the genetic algorithm include: 1) high speed of finding a solution in comparison with other stochastic methods; 2) rather low probability of finding an incorrect solution (falling into a local optimum); 3) clear boundedness in time depending on the required accuracy of calculations.

The genetic algorithm has proved its effectiveness in optimization problems. We shall now consider the example of the genetic algorithm as in the case of constructing an investment portfolio based on the method of evolutionary search in MS Excel.

III. RESEARCH RESULTS. INVESTMENT PORTFOLIO MANAGEMENT BASED ON THE GENETIC ALGORITHM

Using the MS Excel Solver add-in, it is necessary to calculate the profitability-optimal composition of the investment portfolio for the specified parameters:
- portfolio value: 99,500 – 100,000 RUB;
- portfolio profitability: not lower than 9.51%;
- portfolio risk: not higher than 9.99%.

At the first stage, we create a table for calculating value, profitability and risk of the portfolio (Table 2).

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security</td>
<td>Price</td>
<td>Risk</td>
<td>Profit</td>
<td>Max in the portfolio</td>
<td>Max in the portfolio</td>
<td>Share</td>
<td>Risk</td>
<td>Profit</td>
</tr>
<tr>
<td>Company 1</td>
<td>81,800.24 RUB.</td>
<td>0.83%</td>
<td>1.22%</td>
<td>1</td>
<td>81,800.24 RUB.</td>
<td>0.83%</td>
<td>1.22%</td>
<td></td>
</tr>
<tr>
<td>Company 2</td>
<td>1,782.96 RUB.</td>
<td>4.13%</td>
<td>4.11%</td>
<td>1</td>
<td>1,782.96 RUB.</td>
<td>4.13%</td>
<td>4.11%</td>
<td></td>
</tr>
<tr>
<td>Company 3</td>
<td>420.71 RUB.</td>
<td>11.24%</td>
<td>11.23%</td>
<td>1</td>
<td>420.71 RUB.</td>
<td>11.24%</td>
<td>11.23%</td>
<td></td>
</tr>
<tr>
<td>Company 4</td>
<td>3,249.53 RUB.</td>
<td>16.24%</td>
<td>16.17%</td>
<td>1</td>
<td>3,249.53 RUB.</td>
<td>16.24%</td>
<td>16.17%</td>
<td></td>
</tr>
<tr>
<td>Company 5</td>
<td>308.11 RUB.</td>
<td>4.08%</td>
<td>4.09%</td>
<td>1</td>
<td>308.11 RUB.</td>
<td>4.08%</td>
<td>4.09%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>99,954.69 RUB.</td>
<td>9.99%</td>
<td>10.65%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

At the second stage, we open the Solver add-in and select the evolutionary method, then specify the optimization parameters, and set the parameters of the variable constraints.

Table 3 shows the result of the calculations

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security</td>
<td>Price</td>
<td>Risk</td>
<td>Profit</td>
<td>Max in the portfolio</td>
<td>Max in the portfolio</td>
<td>Share</td>
<td>Risk</td>
<td>Profit</td>
</tr>
<tr>
<td>Company 1</td>
<td>81,800.24 RUB.</td>
<td>0.83%</td>
<td>1.22%</td>
<td>1</td>
<td>81,800.24 RUB.</td>
<td>0.83%</td>
<td>1.22%</td>
<td></td>
</tr>
<tr>
<td>Company 2</td>
<td>1,782.96 RUB.</td>
<td>4.13%</td>
<td>4.11%</td>
<td>1</td>
<td>1,782.96 RUB.</td>
<td>4.13%</td>
<td>4.11%</td>
<td></td>
</tr>
<tr>
<td>Company 3</td>
<td>420.71 RUB.</td>
<td>11.24%</td>
<td>11.23%</td>
<td>1</td>
<td>420.71 RUB.</td>
<td>11.24%</td>
<td>11.23%</td>
<td></td>
</tr>
<tr>
<td>Company 4</td>
<td>3,249.53 RUB.</td>
<td>16.24%</td>
<td>16.17%</td>
<td>1</td>
<td>3,249.53 RUB.</td>
<td>16.24%</td>
<td>16.17%</td>
<td></td>
</tr>
<tr>
<td>Company 5</td>
<td>308.11 RUB.</td>
<td>4.08%</td>
<td>4.09%</td>
<td>1</td>
<td>308.11 RUB.</td>
<td>4.08%</td>
<td>4.09%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>99,954.69 RUB.</td>
<td>9.99%</td>
<td>10.65%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IV. SUMMARY

Thus, the software implementation for constructing an investment portfolio based on the method of evolutionary search (genetic algorithm) is proposed.

Based on the optimization algorithm, a model of the risk and return management of the investment portfolio and the calculation of the asset share of the portfolio is proposed.

The paper also presents a forecasting model based on the neural network for calculating the return of portfolio assets.

References

Conference "Management of Large-Scale System Development", MLSD 2018


