Application of Machine Learning in Financial Asset Price Prediction and Allocation

Xinli Guo [1]

[1] College of Economics, Bohai University, Jinzhou City, Liaoning Province, China

Keywords: Machine learning, Financial assets, Price prediction, Asset allocation

Abstract: In today's era, the wave of artificial intelligence has swept across, injecting unparalleled vitality into various industries. The financial field is no exception. Artificial intelligence has provided this industry with new and colorful technologies and brought new opportunities for its further development. This article introduces several basic machine learning algorithms from the perspective of underlying machine learning algorithms, reveals the underlying algorithms used by more popular technologies, and explains the application of machine learning in financial asset price prediction and allocation.

1. Introduction

Compared with machines, brain capacity and computing power have severe limitations on human behavior and thinking. The machine learning algorithm is used to design a method for the computer, so that the computer can automatically make decisions on similar problems after learning by itself, and it has a speed far surpassing human beings [1]. Therefore, in the financial field, some tasks with huge data volume and needle-finding in a haystack are suitable for using machine learning to complete.

Machine learning algorithms can be divided into two categories: supervised learning and unsupervised learning. Supervised learning is particularly widely used in the financial field. The general steps are as follows. In the learning and training stage, the machine learns a class of problems to train a model [1]. At this stage, a test set with answers is first needed, and samples with correct decisions have been made for the computer to learn. In the testing phase, the test set is used to test the model that has been trained, and to further modify the model to get better results.

2Machine learning algorithm

1.1 Decision Tree Algorithm and Random Forest

The decision tree algorithm is a classification algorithm based on the ‘tree’ data structure [2]. A ‘tree’ is composed of nodes and branches, where nodes are used to store information and data, and branches are used to represent the relationship between information and data.

In this algorithm, how to construct a decision tree scientifically and reasonably is the key to this algorithm. Consider the information gain when selecting the features of non-leaf nodes and the feature division method. From the top to the bottom of the decision tree, select the feature with the largest information gain and the feature division method as non-leaf nodes, so that the diversity of features can be utilized as much as possible [2].

1.2 Neural Networks

Neuron is the most basic structure of neural network, it can also be said that it is the basic unit of neural network. Its design inspiration comes from the information transmission mechanism of neuron in biology. Our classmates who have studied biology know that neurons have two states: excited and inhibited. Under normal circumstances, most neurons are in an inhibited state, but once a neuron is stimulated, causing its potential to exceed a threshold, then this neuron will be activated, in an “excited” state, and then to other neurons [3]. Of neurons transmit chemicals (in fact, information).

  DOI: 10.25236/mmmce.2021.023

1.3 Support Vector Machines

Support vector machine is a supervised nonlinear classifier. His algorithmic idea is to find the global optimal solution of the separating hyperplane that can correctly divide the training data set. Support vector machine algorithm has the following characteristics, which makes it can be well applied in the financial industry [4]. First of all, the classification of support vector machines is not absolute, it can tolerate small errors to a certain extent, which is called soft interval. The setting of this soft interval helps to improve the generalization ability of this algorithm. At the same time, the support vector machine uses a kernel function to ensure the existence of the classification hyperplane. In complex realistic tasks, the linear plane may not be able to complete the classification well.

3The application of mechanical learning in the prediction and allocation of financial assets

In recent years, through the influence of various financial institutions and experts, ordinary people have become familiar with the term “asset allocation”, and even feel it a bit cliché. But in fact, many people do not know the basic principles and specific applications of asset allocation [5]. In fact, there are three parts here:

(1) Market analysis

(2) Configuration theory application

(3) Rebalancing method (Rebalancing means that when the asset portfolio deviates from the original state due to market fluctuations, the portfolio is rebalanced by adjusting the asset ratio in time. Yale investment master David Swenson rebalances the asset allocation described as “the pie falling from the sky”.

1.4 Lazy Asset Forecast Portfolio (1/n)

In this combination method, assuming 4 investment targets, each allocation is 25%, and has 3 ways to adjust positions that lazy combinations often match:

(a) Buy-and-hold Strategy

(b) Constant-mix Strategy

(c) CPPI

Let's take the method (b) constant mixing strategy as an example to see how the trigger adjustment for rebalancing is performed when the ratio is different from the original configuration after a certain period of time [6]. First of all, suppose we allocate 25% in the US stock market, Chinese stock market, bond market, and gold, and set a quarter to make adjustments, as shown in Table 1.

Table 1 Fund Allocation Of U.s. Stock Market, Chinese Stock Market, Bond Market, and Gold

2020/1/1	U.S. stocks	A shares	Bond market	Gold
ETF Price	10	10	10	10
Configuration amount	2500	2500	2500	2500
Configuration ratio	25%	25%	25%	25%

After the first quarter, if we find that the U.S. stocks and A-shares are rising, and the gold and bond markets are falling, as shown in Table 2 below:

Table 2 Adjustments to The Allocation of Funds in the U.s. Stock Market, China Stock Market, Bond Market, and Gold

2020/1/1	U.S. stocks	A shares	Bond market	Gold
ETF Price	18	14	12	6
Configuration amount	4500	3500	3000	1500
Configuration ratio	36%	28%	24%	12%

Then, when rebalancing, U.S. stocks and A-shares need to be reduced to 25%, while the bond market and gold need to be increased to 25%, and then adjusted quarterly to complete the rebalancing process.

1.5 Risk Parity Portfolio

Risk parity is a kind of asset allocation investment. Assets with the same risk weight can be allocated to investment portfolios to maintain and increase the value of assets. Stocks and commodity investments have higher risks, and bonds have lower risks. Therefore, high-risk assets should be reduced during the distribution process [7]. Configure so that the risks contributed by it are the same. Under the assumption that the asset correlation is equal, we can calculate a certain type of asset i by risk-parity, and its configuration weight is expressed as follows, but because Risk-Parity does not consider Return, only consider the volatility (risk).

A deep learning model based on big data has been developed to predict the volatility of risk. For example, scholars at Stanford University and Google use Google search trend data and recursive neuron learning methods to predict stock market fluctuations. Its prediction effect is better than the traditional linear model or GARCH model and has been improved by more than 31% [8]. Therefore, as long as it is used properly, the risk-parity approach can provide better configuration results, as shown in Figure 1.

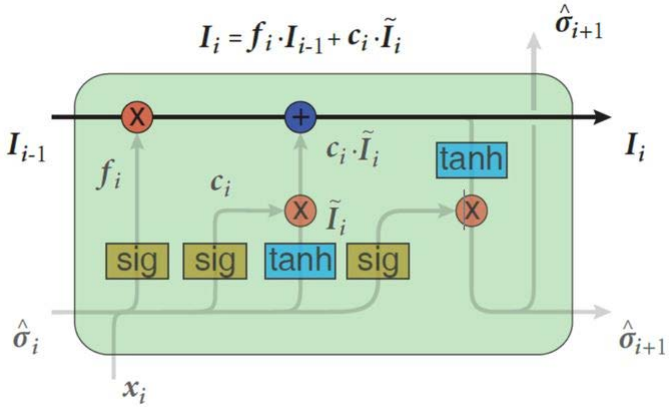


Fig.1 Schematic Diagram of the Learning Method of Xiong, Nichola & Shen Using Deep Learning

1.6 Markowitz's Mean Variance-Efficient Frontier Combination or Black-Litterment Price Prediction Model

The effective frontier portfolio is a configuration method that many students who have studied finance are familiar with. In a portfolio with n assets, it is the investment weight of each asset and the corresponding rate of return [9].

The investment weight obtained by the Lagrangian multiplication method is the optimal investment plan of the model. If the optimized investment portfolio is drawn in a two-dimensional plane with the standard deviation (volatility) as the abscissa and the expected return rate as the ordinate, a curve is formed, as shown in Figure 2.

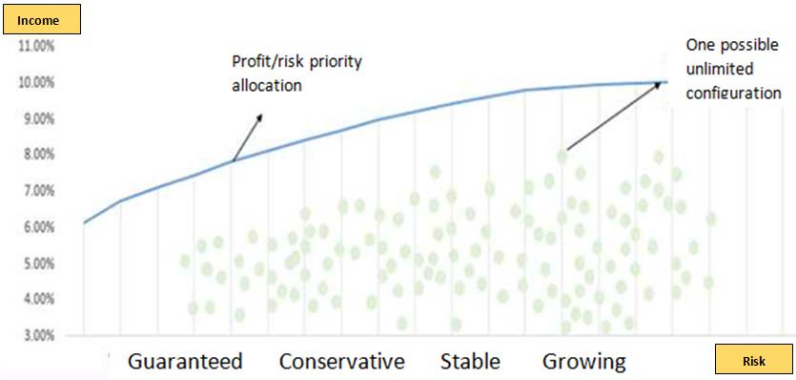


Fig.2 Investment Portfolio (Green Dot) and Effective Frontier (Blue Line)

The more to the right of this curve, the higher the risk of the portfolio, but the higher the relative return, which is also in line with the general perception of high risk corresponding to high return [9]. With this feature, we can map the risk choices to different customer attributes. Configuration. Therefore, the effective frontier portfolio has become the theoretical basis of asset allocation.

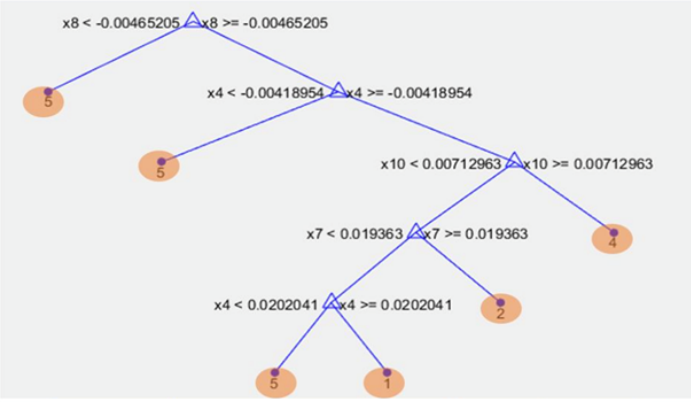
1.7 Three Key Points for Constructing Asset Allocation Portfolio

Now we determine to use Markowitz's mean variance-efficient frontier theory as the basis of asset allocation portfolio. In addition, derive three key points of the construction model: (1) how to estimate the risk, (2) how to estimate the return, and (3) how to correctly classify user attributes.

(1) Estimated risk

(2) Estimated income

For the second part of the income estimation part, the benefits are: lower input data form restrictions; linear/non-linear learning; self-evolution, correction, etc. Machine learning can make supervised learning of model factors [10]. The selection of factors can include fundamental, technical, and bargaining data, and make corresponding estimates of market returns. Figure 3 below shows a case in which the U.S. stock market uses the decision tree factor model to analyze returns.



X8: line growth rate this year; **X4:** Expected corporate revenue growth rate; **X10:** Short-term moving average multiplying and leaving; **X7:** Concentration growth rate

Fig.3 a Case of Using Decision Tree Factor Analysis in the U.s. Stock Market

There are many ways to estimate earnings. We put the asset returns calculated by several different return estimation models as input parameters into Markowitz's effective frontier model or Black-Litterment model to calculate, and analyze its impact on the investment portfolio, that is, the return on investment. which performed.

Below we use three methods to estimate earnings, one is the support vector machine regression model (SVM), the other is the linear regression model, the analysis factors of these two models include interest rate, price-earnings ratio, price-to-book ratio, dividend ratio, enterprise Earnings, transaction value, implied volatility, MACD, KD, etc [10]. And the third is the most commonly used by ordinary people, directly using the market trend in the past few years to predict the future market. Please see Figure 4:

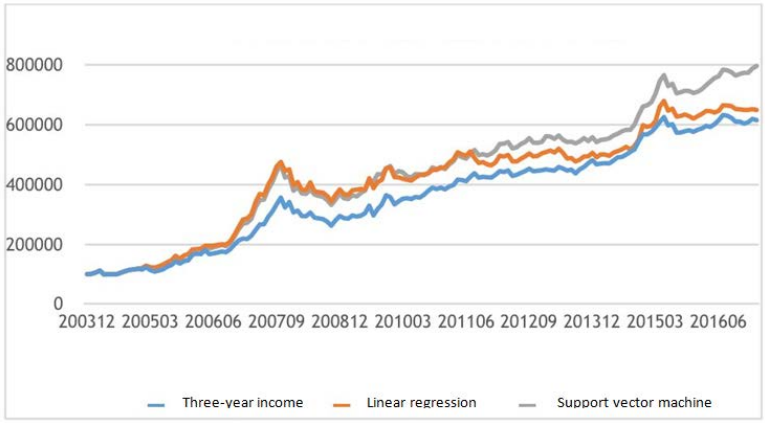


Fig.4 Markowitz Allocation Income Forecast

- Gray curve: SVM Regression – Use Support Vector Machine (SVM) to predict the portfolio performance of asset returns

- Orange curve: Linear Regression-use linear regression to predict the portfolio performance of

asset returns

- Blue curve: directly use the average yield of the past three years to predict the portfolio performance of asset returns.

- Markets covered: the United States, Eurozone, Japan, Australia, Latin America, emerging Europe, Asia except Japan, China, commodities, REITs, overseas investment grade bonds, overseas emerging market bonds and overseas high-yield bonds.

2. Conclusion

In summary, the application of machine learning for asset prediction and allocation combination has been very rich. For financial institutions and asset management companies, strengthening the cultivation of financial technology R&D talents may be an important task in the next few years.

References

[1] Yan, W.M., Li, D.G. and Wu, W.M.: Application of machine learning in the financial field [J]. Computer Education, 08 (08), pp.87-89, 2014.

[2] Chen, H.T.: Research on the application of machine learning in home finance [J]. Fintech Era, 11 (08), pp.11-13, 2015.

[3] Wu, Y.X.: Application of machine learning algorithms in financial market risk prediction [J]. Information System Engineering, 13 (02), pp. 114-116, 2019.

[4] Li, Y.N.: Application of neural network model in anti-fraud in banking Internet finance [J]. The era of financial technology, 11 (08), pp.24-28, 2018.

[5] Wang, Y.J.: Research on bank personal customer credit scoring model [J]. Modern Business Industry, 36 (19), pp.64-65, 2015.

[6] Ju, Zh.L.: Application of machine learning algorithms in financial market risk prediction [J], Financial Economics, 09 (32), pp.22-24, 2017.

[7] Zhao, H.T.: A review of research on the application of machine learning in financial asset price prediction and allocation [J], Journal of Management, 09 (21), pp.216-219, 2018.

[8] Zhang, K.M.: The application of machine learning in the financial field [J], the era of wisdom and wealth, 12, pp.120-123, 2019.

[9] Li, Zh.Y.: Application of machine learning algorithms in financial market risk prediction [J], Financial Economics, 13 (03), pp.400-402, 2019.

[10] Wang, G.Q.: Application of machine learning algorithms in financial market risk prediction [J], Zhifu Times, 10 (18), pp.272-275, 2017.