

A Comparative Analysis of Multiple Linear Regression Models and Neural Networks for Stock Price Prediction - Take BYD as an Example

Yixuan Zhao

Macquarie Business School, Macquarie University, Sydney, NSW 2109, Australia

*Corresponding author. Email: 13726290053@163.com

ABSTRACT

In the context of global efforts to combat the climate warming, energy saving and emission reduction, China has also put forward the goal of achieving carbon peak by 2030 and carbon neutrality by 2060, and the new energy industry has received key support and development, and is also favored by capital, so it is important to analyze and study the factors influencing stock prices, both for the industry and for investors. BYD is the leading company in the new energy industry, it is typical. Therefore, this paper chooses to take BYD as an example to analyze and study the influencing factors of stock price through BP neural network and multiple linear regression model, and finally finds that the company's stock price has a statistically significant correlation with USD/CNH exchange rate, CPI, P/E ratio and Dow Jones Industrial Index. The BP neural network has better explanation ability than the multiple linear regression model.

Keywords: Neural network, multiple regression model, P/E ratio, CPI, Dow Jones Industrial Average

1. INTRODUCTION

The COVID-19 in early 2020 caused strong shocks in global stock markets, with the U.S. Nasdaq index falling by about 6% for the whole year of 2020, and China's Shanghai and Shenzhen indices and Hong Kong's Hang Seng index falling sharply during the Chinese New Year when they were in a favorable market, and the epidemic black swan event had a significant impact on global stock markets and macroeconomics. With the introduction of anti-epidemic policies and vaccination, the epidemic was gradually brought under control, economic fundamentals and investor confidence improved, and the stock market showed a shocking upward trend. However, the economic situation under the epidemic was generally grim, and countries reached a consensus in the UN General Assembly to combat warming through energy saving and emission reduction. In this context, China proposed the goal of achieving carbon peak by 2030 and carbon neutrality by 2060, which in turn made the new energy industry sought after by the capital, and stock prices showed a counter-trend upward trend. It is clear that the stock market is characterized by numerous influential factors and has a complex nature.

After more than a decade of development, China's stock market is playing an increasingly important role in the capital market and economic and financial fields. First of all, the stock market acts as a bridge between fundraisers and investors in the capital market. Fundraisers issue shares through the stock market to raise capital and absorb social capital into the business of enterprises, thus contributing to the rapid development of enterprises. At the same time, stocks can be bought and sold flexibly in the secondary market, and the sufficient liquidity provides investors with a convenient way to invest, and different investors can flexibly allocate their investments according to their own risk preferences and investment ability, and have their own portfolio strategies to take specific risks and obtain corresponding expected returns. Therefore, equity investment, as an important investment channel for investors, has a wide and important significance for investors. In addition, the stock market plays an important role in resource allocation, as listed companies belong to different industries and have differences in management, product competitiveness and capital strength, which ultimately affect the growth potential and investment value of enterprises. The existence of the stock market allows capital to flow

freely and at a lower cost from industries with stagnant growth to industries with better growth prospects, and from poorly managed companies to companies with outstanding performance and higher returns on investment for investors.

Therefore, it is important to analyze in depth the factors influencing stock prices and to price stocks correctly, both from the perspective of investors and financial markets. However, the stock market is inherently complex, with many dimensions affecting stock prices, such as economic cycles, investor expectations, macroeconomic events or company fundamentals. In particular, the complexity of the stock market is more prominent in China, where the proportion of retail investors is large, the regulatory system is not perfect, and the capital market is not deep enough, so it is difficult for the traditional factor model based on the efficient market hypothesis to effectively deal with the complex stock market. Therefore, this paper combines the quantitative analysis to find the influencing factors of stocks from massive data, and compare them with the output of traditional pricing models in order to optimize the explanatory power of the models and find a more effective stock pricing model. Stock pricing model can identify the variables which affect the stock price. In addition, the model tends to help us to forecast price changes when the value of variables changed.

In the context of carbon peaking and carbon neutral, China is vigorously developing new energy industries, including photovoltaic, wind power, energy storage and new energy vehicles, in order to systematically reduce CO₂ emissions and achieve emission reduction targets. On the application side, the new energy industry has a wide range of application scenarios, with the potential to generate a large number of industry opportunities. In addition, the upstream of the industry involves new materials, traditional materials, mining, precision processing and many other industries, so the development of the new energy industry can drive the development of many upstream industries. Considering the downstream application scenarios and upstream industry opportunities, the new energy industry has good investment prospects and investment value. BYD has an early layout in the new energy industry, and the new energy vehicles it produces are widely recognized by the market. It is the leading new energy company in China. BYD is planning to spin off its lithium battery business for listing, which will further promote BYD's development in the direction of specialization after its success, and is more typical among new energy companies, with research significance and value.

2. LITERATURE REVIEW

Every country reached a consensus on combating climate warming, energy saving and emission reduction

to develop a green economy in the Paris Conference, and China has also proposed an overall strategy to achieve carbon peaking by 2030 and carbon neutrality by 2060. The development of the new energy industry is the main path to achieve this goal. The new energy industry has also been sought after by capital, with strong performance in the secondary market.

YongKang[1], in his study of the new energy vehicle industry, selected strategic emerging industries in China, accelerating the development of the new energy vehicle industry will largely help China achieve high-quality economic development. China's new energy vehicle industry is in a rapid development stage. However, there is still a significant gap compared to developed countries. It is necessary for China to take comprehensive measures to help the development of the new energy vehicle industry.

According to TianLei[2], China's motor vehicle energy saving regulations play an important role in promoting skillful emission reduction technologies and reducing CO₂ emissions, and the overall goal of China is to achieve carbon peaking by 2030 and carbon neutrality by 2060, and in the future, China should prioritize optimizing industrial development patterns and improving motor vehicle energy efficiency and emission reduction standards to achieve this goal.

YangkaD[3] refers to the widely accepted term "carbon neutral" as meaning balancing carbon dioxide emissions into the atmosphere with the inputs and outputs of economic activity, which requires the quantification and tracking of carbon emissions with the ultimate goal of offsetting them. Achieving carbon neutrality requires more efficient technologies and the adoption of electrified economic development models and transportation modes. The relationship between carbon dioxide emissions and economic growth, energy use and social welfare are close but complex.

ZhuangZhuang[4] said that the new energy vehicle industry has successfully become an indispensable part of China's sustainable economic development. In order to maintain socio-economic growth and better face the problems of energy shortage and ecological pollution, China needs to seize the development opportunities of the new energy vehicle industry, face market competition with a positive attitude, solve the technical bottlenecks in power batteries and adhere to an innovation-driven development model.

In an empirical study, Samarakoon[5] finds that the dividend policy of a firm is not affected by the liquidity of the stock market, but by the liquidity of the firm itself. Therefore, the dividend policy varies depending on the nature of the firm itself.

In an empirical study, YAOWANG[6] finds that the two traditional classical models of CAPM, FAMA and FRENCH three factors have poor utility in the Chinese

brokerage stock sector, market capitalization does not have a significant impact on returns, and book-to-market ratio and P/E ratio have some impact on lagged one-period returns but the explanatory power is less than 20%, testifying that the Chinese capital market is not mature enough and has obvious the herding effect and speculative nature of the Chinese capital market are evident.

In an empirical study, Chellaswamy [7] finds that prohibitions on fraudulent and unfair trading practices have a significant positive impact on stocks. Chellaswamy [7] also finds that large mergers and acquisitions, and requirements for debt disclosure have a significant positive effect on stock returns. Besides, Chellaswamy [7] found that trading volume reflects the general condition of the stock market. Besides, the study of stock market reforms and trading volume will help investors to better interpret the trends in the stock market and its changes.

3. RESEARCH METHOD

Multiple linear regression model: Multiple linear regression is a mathematical method of constructing regression equations for two and more independent variables to find linear relationships between multiple independent variables and the dependent variable. The main purpose of this method is to determine whether there is a correlation between multiple variables and to predict the dependent variable by constructing a model with the general expression:

$$Y = K_1X_1 + K_2X_2 + \dots + K_nX_n \quad (1)$$

where Y is the dependent variable, X₁, X₂X_n are the independent variables, and E is the residual term.

Where Y is the independent variable, X₁, X₂..... X_n is the independent variable, and E is the residual term.

In the multiple linear regression equation, the following assumptions are implied.

- (1) the expectation of the residual term is 0
- (2) The variance of the residual term corresponding to any one observation is certain
- (3) There is a linear relationship between the independent variable and the dependent variable
- (4) The residuals are normally distributed
- (5) The correlation coefficient between the residuals is 0

Multiple linear regression models are based on the least squares method to minimize the sum of squares of residuals and fit the scatter points on a straight line, so they are only suitable for use between variables with linear correlations, but cannot effectively fit variables with nonlinear relationships, such as variables that present an approximate exponential relationship or more

complex relationships. However, there are many factors influencing stock prices and complex relationships among variables, and the use of simple linear equations for such a large number of variables with complex interactions is inevitably biased. To cope with this situation, this paper introduces a neural network model, which is widely used to describe the relationships among complex nonlinear variables.

Neural networks work in a similar way to how neurons in the human brain process information, and can be used for supervised learning (labeling data) or unsupervised learning (no labeling of data, the machine itself identifies and classifies the features of the data). Generally speaking, neural networks are divided into three layers: the first layer is the input layer, which can be understood as the neuron that receives the original data, and then the received data is transmitted to the next layer of neurons with the corresponding weighted average processing, which is generally called the hidden layer. The data initially processed by the hidden layer will be processed at a deeper level by the activation function, and finally the deeply processed data will be transmitted to the neuron in the latter stage and the result will be output, which is called the output layer. The simplest neural network is divided into the above three layers, but the hidden layer can also be extended from one layer to multiple layers. The neural network with multiple hidden layers is called deep learning, which can cope with more complex relationships between variables and make up for the lack of linear equations.

4. INFLUENCING FACTORS

4.1. Dow Jones Industrial Average

The Dow Jones Industrial Average is based on the top 30 most influential publicly traded companies in the U.S. The final price index is an arithmetic average obtained by dividing the value of all stocks included by the total number of shares. As the index reflects the share price level of the most influential companies as a whole, it largely reflects the overall expectations of investors on the stock market, and when the index is at a high level, it means that investors' expectations on the stock market are generally positive and investment enthusiasm is high, which enables the stock market to maintain a high level. On the contrary, if the index is at a low level, it reflects investors' pessimistic expectations of the stock market to a certain extent and their enthusiasm for investment is low. Therefore, the level of the index is a response to market sentiment, reflecting investors' expectations and attitudes towards the market.

4.2. Consumer Price Index

Consumer Price Index, or CPI for short, is a

macroeconomic indicator that reflects changes in the price level of consumer goods and services purchased by households. It is a statistical survey of the final prices of products and services, which is closely linked to the life of the public and reflects to some extent the degree of inflation. In general, the higher the CPI, the higher the level of inflation and the overheated macroeconomic development. Conversely, if the CPI index is lower, the lower the level of inflation, generally speaking, the economic development is more depressed. Generally speaking, when inflation is high and the economy is overheated, investment demand is strong, which is reflected in the upward trend of the stock market, and conversely, when inflation is low or even deflationary, the economy is underdeveloped and lacks vitality, which is reflected in the stock market, where investment demand is low and stock prices are down.

4.3. USD/RMB Exchange Rate

USD to RMB exchange rate refers to the rate of conversion of USD to RMB, which is determined by the foreign exchange market. If the USD to RMB exchange rate is decreasing, it means the USD depreciates and the goods and services exported from the U.S. to China become cheaper for Chinese residents; Conversely, if the exchange rate of the USD rises against the RMB, which represents a depreciation of the RMB, then U.S. exports of goods and services to China become more expensive for Chinese residents, who will tend to buy domestic goods rather than imported goods. Therefore, a depreciation of a country's currency will have a positive stimulating effect on domestic demand, on domestic firms, or on the stock market.

4.4. P/E Ratio

The price-to-earnings ratio (P/E) is the result of dividing the share price by the earnings per share, which indicates how many years it will take to recover the investment cost with the same earnings per share.

Generally speaking, the lower the P/E ratio, the lower the share price relative to the company's profitability, and therefore the shorter the payback period and the lower the investment risk. The P/E ratio reflects the valuation of the secondary market and the company's fundamentals, and is a commonly used valuation indicator.

5. EMPIRICAL ANALYSIS

In this paper, the closing price of BYD Auto at the end of each month from December 2016 to November 2021 is selected as the explanatory variable and set as Y_i , i denotes the month corresponding to the closing price. The Dow Jones Industrial Index, the U.S. dollar to RMB exchange rate, the consumer price index and the purchasing managers' index from 2016 to November 2021 were selected as the explanatory variables, set as X_i for the Dow Jones Industrial Index, Z_i for the U.S. dollar to RMB exchange rate, C_i for the consumer price index and P_i for the purchasing managers' index, and used to construct the multiple linear regression equation as follows:

$$Y_i = b_0 + b_1 X_i + b_2 Z_i + b_3 C_i + b_4 P_i + e \tag{2}$$

Where b_0 is the intercept of the equation, b_1 is the correlation coefficient between the Dow Jones Industrial Index and the month-end closing price, b_2 is the correlation coefficient between the USD/RMB exchange rate and the month-end closing price, b_3 is the correlation coefficient between the consumer price index and the month-end closing price, b_4 is the correlation coefficient between the purchasing managers' index and the month-end closing price, and e is the residual term of the regression equation. The month-end closing prices of BYD and Dow Jones Industrial Index data are obtained from the wind data terminal, and the CPI and PMI data are obtained from the public macroeconomic data published by the Yingwei Talent Financial website.

Table 1. Regression results of BYD influencing factors

Influencing Factors	BETA
Constants	220.023
USD to RMB exchange rate	-53.088
Dow Jones Industrial Index	0.08
P/E Ratio	0.481
CPI	-11.448

The empirical analysis shows that the share price of the company is negatively correlated with the USD-RMB exchange rate with a correlation coefficient of -53.088, which is statistically significant, meaning that when the USD-RMB exchange rate rises and the RMB depreciates, BYD's share price tends to fall. The above

preliminary analysis suggests that when the RMB depreciates, imported goods will become more expensive and favorable for export, thus boosting domestic demand and therefore the stock price, which does not match the results obtained in our preliminary analysis. This may be due to BYD's high international

market share. Since overseas markets occupy an important position in the company's overall market structure, and the US dollar plays an important role in the settlement system of overseas countries, the appreciation of the US dollar has increased the prices of the company's products, which has a negative impact on the overseas sales of BYD's products.

The correlation coefficient between the company's stock price and the Dow Jones Industrial Index is 0.08, which is positively correlated. The article mentioned in the preliminary analysis above that the Dow Jones Industrial Index is an overall response to investor confidence and its stock market sentiment, so BYD's stock should tend to show a positive correlation with the Dow Jones Industrial Index, which is consistent with the preliminary analysis, although BYD is not a traditional manufacturing industry, but belongs to the new energy track with broad prospects and pursued by capital, which does not fully match with the Dow Jones Industrial Index that cannot highlight its industry advantages. However, the companies selected for the Dow Jones Industrial Average are large companies, largely reflecting the economic boom and investors' confidence and attitude towards the stock market, and the direction of individual stocks and the index is broadly in line with expectations.

The correlation coefficient between the company's share price and P/E ratio is 0.481, showing a positive

correlation, which means that the company's share price tends to move in the same direction as the P/E ratio. The high P/E ratio represents the high valuation of the company's shares in the market, and the same directional movement can be interpreted as investors' positive outlook on the company and believe that the company has good growth potential. BYD's operating results and its industry show that the company has good profitability and is in a period of explosive growth in the new energy industry, presenting a Davis win-win situation for the company's stock. The company's share price and P/E ratio show a positive correlation in line with expectations.

The BP neural network selected 50 data as the training set and 8 data as the test set. In general, the prediction samples showed a better fit between the predicted and true values, and the output results are shown in the following figure. In order to compare the prediction ability of BP neural network and multiple linear regression model, the output results of both are presented in the following table. Through comparison, it is found that BP neural network has more accurate prediction of stock price with smaller residuals and better fit, and in this case, the explanatory ability of the model is better than that of multiple linear regression model.

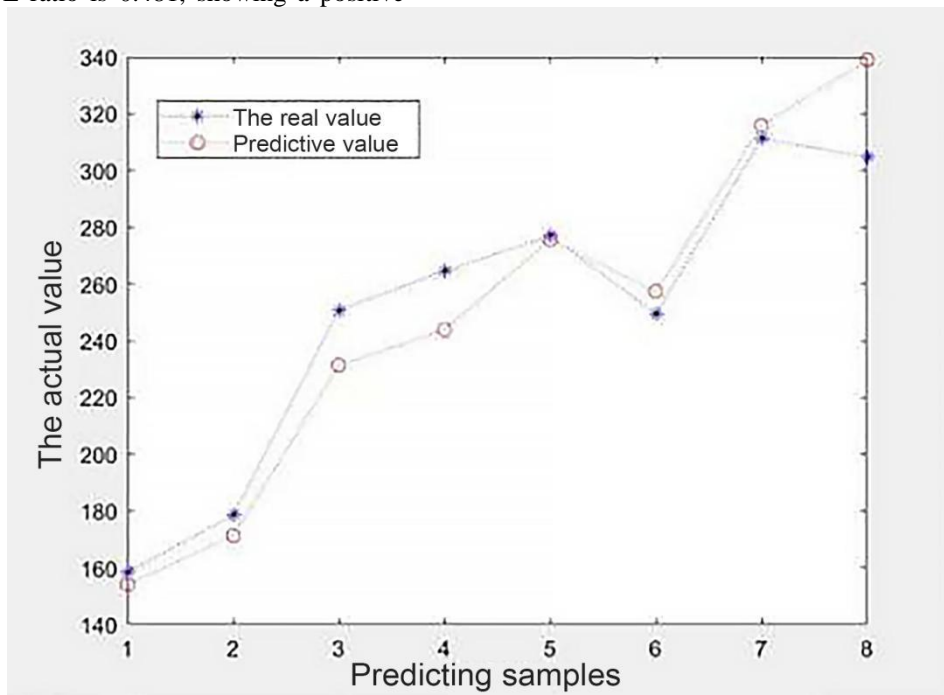


Figure 1. The real value and predicted value of BP neural network

Table 2. Comparison of BP neural network and SPSS results

Real Stock Price	BP Neural Network	SPSS
158.4906	154.0334	186.7589

178.6105	171.3517	197.1394
250.8743	231.3537	218.4236
264.5774	243.8893	230.957
277	275.6863	249.8241
249.51	257.3899	229.9791
311.31	315.8873	283.151
304.7	339.048	260.052

6. CONCLUSION

In summary, in this case, the BP neural network is more accurate than the multiple linear regression model in predicting the stock price, and the explanatory power of the model is stronger. BP's neural network does not have the inherent limitation that the multiple linear regression model can only deal with the linear relationship between variables, and can better cope with the complexity of the stock price influencing factors themselves, and adapt to the non-linear and more complex relationship between multiple variables. The data selected for the test set in this paper are small and only P/E ratio can reflect unsystematic risks. This may limit the explanatory power of the model to some extent. In the future, we will extend the dataset and add other variables reflecting individual risks into the model to improve the model's explanatory power.

REFERENCES

- [1] Liang Y, Zhang W. Research on the High-Quality Development Path of New Energy Vehicle Industry in China[J]. IOP Conference Series Materials Science and Engineering, 2019, 677:032090.
- [2] Tianlei, Zheng Zhao, Wang Shaohui, Liu Xiang, Bao Zhichao, Liu Mengxue, Ji. (2021). The development trend and prospect of automobile energy-saving standard system under the goal of peak carbon dioxide emissions. E3S Web of Conferences. 271. 02006. 10.1051/e3sconf/202127102006.
- [3] Yangka D, Rauland V , Newman P . Carbon Neutral Bhutan: sustaining carbon neutral status under growth pressures. 2020.
- [4] Hao, Zhuangzhuang Li, Zhiyang Ni, Hongjun Lv, Shuaishuai Wang, Xingxing Zhu, Yu. (2021). Present situation and prospect of new energy vehicle industry in China. IOP Conference Series: Earth and Environmental Science. 791. 012153. 10.1088/1755-131
- [5] Hettiarachchi, M Samarakoon, S.M.R.K. (2020). Stock market liquidity and firm dividend policy: Evidence from a Frontier Market.
- [6] Wang Y, Yang KL. Study on the applicability of CAPM model and factors influencing returns of listed brokerage stocks[J]. Journal of Chongqing University of Technology (Natural Sciences), 2019.
- [7] Chellaswamy K P, Natchimuthu N, Faniband M . Stock Market Reforms and Stock Market Performance[J]. International Journal of Financial Research, 2021, 12.