Application of Mathematical Model using Artificial Neural Network in Photovoltaic Electrical Power Generation Building

Shengfeng Meng*, Zeyu Pan and Jiayi Meng

Residential college, Shanghai University, Shanghai, China

*Corresponding author: 1826566439@shu.edu.cn

Abstract. By counting the stock data of PV building integration-related companies in Shanghai and Shenzhen stock markets, a sector index model of PV building integration was established, moving averages were calculated and moving averages were plotted, and the prediction results of the model were used to compare with known data to optimize the model, and data from April 1, 2019 to May 6, 2021 were used as the training set and data from May 6, 2021 to May 28, 2021 were used as the test set by recurrent neural network and LSTM long short-term memory artificial neural network to optimize the model. as the training set and the data from May 6, 2021 to May 28, 2021 as the test set to train the optimized model to obtain the prediction results for the next month. Using the data from April 1, 2019 to May 28, 2021, a time period of 2 months is used to analyze the correlation between the SSE index and the PV building integration sector index for each period using Pearson correlation analysis. Finally, the time series forecasting model analysis is combined to assess the investment risk of individual stocks in the PV building integration sector.

Keywords: LSTM long and short period memory model, RNN recurrent neural network, Pearson correlation analysis, PV building integration.

1. Introduction

Recently, the goals of "carbon peaking" and "carbon neutrality" have been included in the government work report for the 2021 National People's Congress and the subsequent outline of the 14th Five-Year Plan. These two parts have been hotly debated by many people. The related concept sector has become a hot sector in the market recently. The core of achieving the goal of "carbon peaking and carbon neutral" is the implementation of energy saving and carbon reduction in the building sector [1].

According to the stock data of 37 PV building integration-related companies in Shanghai and Shenzhen stock markets, the stocks of these 37 companies are considered as a whole and called the PV building integration sector. As an emerging segment, the PV building integration segment can predict the development trend of the corresponding industry by forecasting the development trend of the segment index.
2. Data pre-processing
This paper counts the changes in market capitalization of 37 stocks of PV building integration-related companies in Shanghai and Shenzhen stock markets from April 1, 2019 to May 28, 2021, due to the fact that the 37 stocks studied to this paper have different listing dates and trading times, some of which have been listed and traded since the 1990s, while some of them have only been listed and traded in recent years. Therefore, it is necessary to preprocess the data and analyze only the data needed.

Southern Energy stock (003035) opens on January 19, 2021, so in the statistical timeframe, Southern Energy does not participate in the calculation until January 19, 2021. There are only 36 stocks that make up the PV building integration sector in the previous period, while there are 37 before January 19, 2021, so pay special attention to the difference data here when modeling.

Since the stock market trades on weekdays, i.e., not on weekends and holidays. Therefore, when using the date representation, there are many gaps in between. These gaps have no trading data and cannot be filled with zeros, so when plotting the image we should maintain some consistency in the pre and post holiday trading data. After pre-processing the data, the image is plotted as shown in Figure 1.

![Figure 1. Data pre-processing effect](image)

3. Sector Index Model
The moving average (5-day, 10-day, 15-day, etc.) model of the PV building integration sector index is established through the historical data of 37 PV building integration-related enterprise stocks in Shanghai and Shenzhen stock markets, i.e., how 37 individual stocks constitute the sector by weighting, and the index of the sector is calculated. After obtaining the index calculation model, this paper can calculate the moving averages and plot the moving averages. This paper refers to the index calculation methods specified in the "Rules for Calculating and Maintaining Stock Indices of China Securities Index Co.", which are applicable to the calculation and modeling of this sector index.

3.1. Model Building
The calculation is performed using the Piech's weighting method and is generally based on the following formula: Reporting period index = Adjusted market value of the sample for the reporting period / divisor x basis point, where Adjusted market value = Σ (security price x number of adjusted equity shares x weighting factor x exchange rate). The number of adjusted equity is obtained by adjusting the sample equity according to the graded shift method. In order to calculate the adjusted equity, both the free float and the graduation factor must be confirmed.
To reflect the movement of actual outstanding shares in the market, this index excludes restricted shares and basic non-floatable shares held for strategic holdings and other reasons from the listed company's share capital. The remaining share capital is referred to as free float, or free float volume.

Unless otherwise specified, China Securities Index Co., Ltd. uses a graded reliance on the gearing method in calculating the index, i.e., the total share capital of the class is given a certain weighted percentage based on the proportion of free float to the total share capital of the sample (i.e., the proportion of free float), so as to ensure the relative stability of the share capital for calculating the index.

Free float ratio = free float / total sample equity.
Adjusted equity = Total equity of the sample × Weighted ratio.

3.2. Plotting graphs with moving averages
With the index calculation model, the graph is plotted with moving averages as shown in Figure 2, where light orange is the daily closing price of the index, orange is the MA5, yellow is the MA10 and green is the MA20.

4. An exponential prediction model based on LSTM long and short-term memory artificial neural network
This paper analyzes the error of the proposed model based on the data from May 6 to May 28, 2021, and corrects the model. Based on the revised model, the future trend of the sector is predicted, and the daily moving average for 20 trading days, weekly moving average for 3 weeks and monthly moving average for 2 months after May 28 are given.

In this paper, Recurrent Neural Networks (RN) and LSTM fully known as Long-Short Term Memory (LSTM) artificial neural networks are used. The data from April 1, 2019 to May 6, 2021 are used as the training set and the data from May 6, 2021 to May 28, 2021 are used as the test set. The optimization model is trained to obtain the forecast results for the next month.

The data from April 1, 2019 to May 6, 2021 is used as the training set and the data from May 6, 2021 to May 28, 2021 is used as the test set to train the optimization model and obtain the prediction results for the next month.

The number of hidden layers of LSTM is 128, the number of stacked layers is 2, the learning rate is 0.001, and the training data accounts for 95% of the total data, and the predicted trend and average for the next month is shown in Figure 3.
5. Correlation analysis between SSE index and PV building integration sector index
Based on the data from April 1, 2019 to May 28, 2021, a time period of 2 months, the correlation between the SSE index and the PV building integration sector index was analyzed separately for each period using Pearson correlation analysis.

The results of the correlation analysis between the SSE index and the PV building integration sector index for each segment through the Pearson correlation coefficient calculation formula are shown in Figure 4.

6. Photovoltaic building integration sector stocks to assess the investment risk
The maximum retracement A, maximum increase B and maximum decrease C from April 1 to May 28, 2019 were counted and then analyzed in conjunction with the time series forecasting model to obtain the uptrend and downtrend D (positive numbers indicate an increase, normalized between 1 and -1). Among the 20 trading days of the prediction model, the percentage of days exceeding MA5 and MA10 are E and F, respectively, and the percentage of days exceeding MA20 is G.

The RNN and LSTM models above are still used to predict the movements of individual stocks over the next 20 trading days.

Considering that the greater the maximum retracement, the greater the maximum increase or decrease, the greater the change trend, the more severe the deviation from the moving average, the greater the risk. In addition, the deviation and change trend of the moving average significantly predict short-term (one-month or two-month) risk, and the maximum retracement rate, maximum up and down rate significantly predict long-term (more than six months of investment) risk. Considering that our investments are short-term investments within one month, short-term factors account for 65% of the impact effect and long-term factors account for 35% of the impact effect. After normalization and linear superposition, we can build the following risk assessment model.
Get the results of risk ranking, taking the highest and lowest risks.

Based on the predicted trend of each stock, determine the appropriate time to trade. We choose some of these stocks with less risk rather than the least risk to focus on and consider them in conjunction with their actual market, using the same RNN+LSTM model as in the second question to make predictions, and the results of the prediction for some of the best stocks are shown in Figure 5.

\[
\text{RiskIndex} = 0.65(0.25D + 0.25E + 0.25F + 0.25G) + 0.35(0.4A + 0.3B + 0.3C)
\]

**Figure 5.** Some of the best stock trends are expected

7. Conclusion

This paper uses the index calculation reference method of CSI, which is in line with the actual application scenario. This paper uses RNN and LSTM for stock market prediction, which is relatively in line with the prediction method for this application scenario. This model considers both long-term and short-term data effects, and the prediction effect for the stock market is relatively accurate. This paper innovatively establishes the predictive effect of each indicator of the stock for risk, using normalized preprocessing and linear superposition, which is intuitive and effective.

References

[1] Lu Yufa. The world's largest single-scale thin-film photovoltaic building integration demonstration project submits an eye-catching report card of "carbon neutrality"[J]. China Building Materials,2021(04):84.

[2] Wang Dong, Zhang Kejia, Zhang Yang. The development status of BIPV-related standards at home and abroad[J]. Solar Energy,2021(05):12-19.

[3] Fang JJ. Photovoltaic building integration concept and PV system design points [J]. Technology and Industry,2021,21(05):251-254.

[4] Zhang J. Research on K-line sequence prediction based on LSTM neural network[J]. Information Technology and Informatization,2021(04):75-78.

[5] Li WJ,Wang XIAO. Time series prediction method based on simplified LSTM neural network[J]. Journal of Beijing University of Technology,2021,47(05):480-488.

[6] Wu, X. F., Lin, X. Y. Y., Jin, Y. N.. An integrated portfolio forecasting model based on time series model selection[J]. Statistics and Decision Making,2021,37(09):5-8.

[7] Xie J, Hu N, Gao B, Luo Tien T. Research on the prediction of market excess returns by illiquidity in Chinese stock market--based on ARFIMA model[J]. Journal of Guizhou University of Finance and Economics,2021(02):31-40.