

Application and optimization strategy of artificial intelligence in financial data analysis

Zhiying Liang Lixin Jiao*

Zhuhai University of Science and Technology, Zhuhai 519040, China.

Abstract: This paper first briefly introduces the artificial intelligence technology commonly used in financial data analysis, including data preprocessing, machine learning algorithm and deep learning technology. Then, the application of artificial intelligence in financial data analysis is discussed in detail. Then, the optimization strategy of artificial intelligence in financial data analysis is proposed, including data quality improvement, feature engineering optimization, model selection and parameter adjustment, model fusion and enhancement, online learning and real-time prediction. Finally, the challenges and future development directions of artificial intelligence in financial data analysis are prospected.

Keywords: Financial Data Analysis; Artificial Intelligence Technology; Optimization Strategy

1. Introduction of artificial intelligence technology in financial data analysis

1.1 Data Preprocessing

Data preprocessing is an important step in financial data analysis that aims to clean and transform raw data to improve its quality and availability. Common problems include missing values, outliers, and noise. Data preprocessing includes missing value filling, outlier processing and noise filtering. In addition, there are steps such as feature selection, feature transformation, and data standardization to extract useful information and reduce dimensions.

1.2 Machine learning algorithms

Machine learning algorithms make predictions and decisions by learning patterns and rules from data. In finance, machine learning algorithms can be used for tasks such as market trend prediction, risk assessment, credit evaluation, and portfolio optimization. Common machine learning algorithms include linear regression, logistic regression, decision trees, support vector machines, and random forests.

1.3 Deep learning technology

Deep learning technology is a machine learning method based on neural networks, which has powerful pattern recognition and feature extraction capabilities. In financial data analysis, deep learning techniques can process complex financial data, including stock prices, market indexes, and trading volumes, through multi-layer neural network models. Applications of deep learning technology in the financial sector include market trend prediction, risk assessment and optimization of trading strategies.

2. Application of artificial intelligence in financial data analysis

2.1 Market trend forecast

By mining historical data, machine learning and deep learning algorithms identify potential patterns and patterns, and make accurate predictions about future price movements, helping investors to formulate smart investment strategies. The prediction methods include time series analysis, regression analysis and neural network model. Time series analysis predicts the development trend based on time series data; Regression analysis to determine the interdependence between variables; The neural network model can find the potential nonlinear relationship from a large number of complex financial data and improve the prediction accuracy.

2.2 Risk assessment and credit assessment

Common risk assessment and credit assessment methods include logistic regression, support vector machine and random forest. Logis-

tic regression is a statistical method widely used in risk assessment, which can determine the probability of customer default by learning historical data. Support vector machine is a classification method based on the principle of maximum interval classification, which can help to find the best separated hyperplane for customer classification. Random forest is an ensemble learning algorithm that improves model generalization by building multiple decision trees to more accurately assess customer credit risk.

2.3 Portfolio optimization

With the help of artificial intelligence technology, financial data analysis can be used for portfolio optimization. Machine learning algorithms analyze historical data on assets to determine the optimal allocation ratio to maximize portfolio returns and minimize risk. Markowitz model, genetic algorithm and particle swarm optimization are common methods for portfolio optimization. The Markowitz model is based on mean-variance analysis to help find the portfolio with the highest expected return and the lowest risk. Genetic algorithm and particle swarm optimization are based on natural heuristic algorithms to find the best solution by simulating the natural evolution process to maximize the return and minimize the risk of the investment portfolio.

2.4 Automatic trading and quantitative trading

Artificial intelligence is applied to automated trading and quantitative trading, and by monitoring and analyzing market data in real time, machine learning and deep learning algorithms can automatically execute trading strategies and adjust them according to market conditions. Automated trading improves trading efficiency and accuracy, and reduces the impact of human error and emotional factors. Common methods include rule-based trading systems, neural network trading systems, and high-frequency trading algorithms. Rules-based trading systems make decisions based on preset rules and avoid emotional influences. Neural network trading systems use neural networks to learn the best strategies to maximize returns. High-frequency trading algorithms use high-speed computers and complex algorithms to carry out high-frequency trading, capture market opportunities, and achieve ultra-high-frequency trading profits.

3. Optimization strategy of artificial intelligence in financial data analysis

3.1 Data quality improvement

In the process of applying artificial intelligence to financial data analysis, the improvement of data quality is an important optimization strategy. First, a comprehensive data cleaning is needed to eliminate outliers and missing values and ensure the accuracy and consistency of the data. Second, the data can be noise filtered to reduce noise interference with the analysis results. In addition, data enhancement techniques can be used to expand the data set by synthesizing new data samples, thereby improving the generalization ability and robustness of the model.

3.2 Optimization of feature engineering

First, statistical methods can be used for feature selection, for example, the most important features can be selected by calculating the correlation between features and target variables. Second, model-based feature selection methods, such as recursive feature elimination, can be used to determine the best feature combination through the change of model performance. Third, deep learning techniques can also be used for feature extraction, such as using autoencoders to learn effective feature representations.

3.3 Model selection and parameter adjustment

In terms of model selection and parameter adjustment, models can be evaluated and optimized through methods such as cross-validation, grid search, and Bayesian optimization. Cross-validation is a statistical method that divides the data set into several subsets, then trains and tests the model on each subset, and finally synthesizes the results of all subsets to evaluate the model performance. Common cross-validation methods include K-fold cross-validation and leave-one cross-validation. Grid search is an exhaustive search method that evaluates the performance of each combination by performing various possible combinations of each parameter value within a pre-set range and cross-validation. Bayesian optimization is a method based on Bayesian statistics that uses Bayesian prior distributions to guide the search process to find the optimal combination of parameters more efficiently. For example, for K-fold cross-validation, the following formula can be used to calculate the average Accuracy: $AVG\ Accuracy = (1/k) * \sum(Accuracy_i)$, where AVG accuracy is the average accuracy, k is the number of

folds, and Accuracy_i is the accuracy of the I-fold.

3.4 Model fusion and enhancement

To improve the accuracy and stability of financial data analysis, AI can optimize strategies through model fusion and augmentation techniques. In terms of model fusion, ensemble learning method can be used to combine the prediction results of multiple models to obtain more accurate predictions. In terms of model enhancement, data enhancement can be used to expand the data set and improve the generalization ability of the model, and model enhancement techniques, such as knowledge distillation, can also be used to improve the learning ability and generalization ability of the model.

3.5 Online learning and real-time prediction

Online learning can receive new market data in real time and update model parameters to adapt to changes in the market. Real-time forecasting can monitor market data in real time and adjust forecast results based on the latest data, thus improving the accuracy and timeliness of forecasts. Through online learning and real-time forecasting, financial data analysis can be more agile and accurate, and better respond to changes in the market.

4. Case: Application of artificial intelligence in the field of investment advisory

An investment company uses artificial intelligence technology to optimize the work flow of investment advisers, obtains the risk tolerance of customers through questionnaires, takes it as a model parameter, combines the ability to obtain multiple data, and compares the investment return rate and risk fluctuation of various funds. The use of deep learning algorithms to build data models to improve business efficiency and decision quality, especially in risk assessment and portfolio management. The data result takes the customer's risk tolerance, the fund's investment rate of return and risk fluctuation as parameters to input the model, and the output optimizes the investment portfolio. Under the premise of meeting the customer's risk tolerance, the investment rate of return is increased as much as possible. Related formula: In this case, the main use of Markowitz's portfolio theory. According to this theory, the expected return of a portfolio can be calculated using the following formula: $E(R_p) = \sum W_i * E(R_i)$; Where $E(R_p)$ is the expected return of the portfolio, W_i is the weight of asset i in the portfolio, and $E(R_i)$ is the expected return of asset i .

5. Conclusion

By summarizing the application of artificial intelligence technology in financial data analysis and its optimization strategy, this paper reveals the great potential of artificial intelligence technology in the financial field. Although there have been successful application cases, there are still many problems to be solved, such as data quality, feature engineering, model selection, etc.

References

- [1] Guan Peng, Huang Qin, Mu LAN. Exploration and practice of PBL teaching model of "Financial Data Mining" under the background of Artificial intelligence [J]. Science and Education Literature (Next ten-day issue), 2021, (27): 144-146.
- [2] Yu Fei. Research on the application of Artificial Intelligence in Data Analysis [J]. Information Recording Materials, 2019, 21(12): 179-180. (in Chinese)
- [3] Wang Lipeng. Thinking about the application of artificial intelligence technology in the financial field -- Taking deep Learning as an example [J]. Times Finance, 2018, (06): 31+35.