

Optimization of Quantitative Investment Strategies in the Financial Big Data Environment

Jinhong Wang

Xiamen University, Xiamen, China

Abstract: With the rapid progress of technology, the financial field is rapidly shifting towards a data-driven environment, with "financial big data" becoming its core component. This transformation has brought about profound changes in the financial market, especially with a significant impact on algorithmic quantitative investment strategies. This article delves into the optimization of quantitative investment strategies in the financial big data environment, analyzes how the characteristics of big data affect quantitative investment, and how to utilize these characteristics to optimize investment strategies. The article first outlines the definition, characteristics, and sources of financial big data, with a focus on describing its core characteristics such as capacity, speed, diversity, and authenticity. This article highlights the application of machine learning and deep learning in financial analysis, as well as the new perspective provided by unstructured data for quantitative strategies. Finally, the conclusion section of the article summarizes the opportunities and challenges brought by the integration of financial big data and quantitative strategies, calling on financial institutions and investors to have a deeper understanding and utilization of these two major trends.

Keywords: Financial big data, Quantitative investment strategies, Strategy optimization, Data-driven.

1. Introduction

In the information age of the 21st century, data has become a new "currency". Especially in the financial field, with the rapid progress of computer technology, network technology and storage technology, financial data has shown explosive growth. This large-scale, diversified and high-speed data environment is called "financial big data". In this data background, traditional financial investment strategies are facing unprecedented challenges and opportunities, especially those based on mathematical models and computer algorithms. Financial market has always been an information-intensive market [1]. However, in the past, due to the limitation of technology and equipment, a large number of financial data were omitted or underutilized. Today, every stock transaction and every economic comment on social media are captured, stored and analyzed, which not only comes from the traditional financial trading system, but also includes emerging technology platforms such as the Internet, mobile devices and the Internet of Things [2]. Quantitative investment is not a new concept, but in the era of big data, it presents new vitality. Quantitative investment strategy refers to the strategy of selecting and trading financial assets through mathematical models, which attempts to find the laws and trends of the market by analyzing a large number of data, so as to obtain excess returns [3]. In the context of big data, quantitative investment has a wider range of data sources, more complex analytical tools and higher computing power. With the in-depth application of big data technology in the financial field, quantitative investment strategy is facing a transformation. Traditional models and algorithms may no longer be applicable, but new data-driven methods are emerging. For example, advanced data analysis technologies such as machine learning and deep learning are being applied to stock selection, risk assessment, market prediction and other fields. At the same time, unstructured data, such as text, voice and images, also provide a new analytical dimension for quantitative investment [4-5]. In view of the above background, this paper aims to explore how to optimize the

quantitative investment strategy in the financial big data environment. We will first outline the characteristics and sources of financial big data, and then discuss in depth how big data technology affects and promotes the development of quantitative investment strategies. We will also discuss the significance and challenges of this transformation to investors, financial institutions and regulators [6]. In short, financial big data and quantitative investment strategy are two important development trends in the field of financial technology. Their integration will not only reshape the pattern of financial markets, but also provide investors with more accurate, efficient and innovative investment tools [7]. This paper hopes to provide readers with a comprehensive perspective and understand the intersection of these two trends, as well as the opportunities and challenges they bring to the financial field.

2. The Impact of Financial Big Data on Quantitative Investment Strategies

2.1. Data volume and data quality

With the rapid growth of data in the financial field, big data has become a reality that financial institutions and investors must face. In this context, understanding how financial big data affects quantitative investment strategies has become particularly important. Although financial big data has brought unprecedented opportunities to the financial field, it has also brought many challenges. The capacity and diversity of big data make data cleaning, validation, and integration more difficult. Errors, omissions, and inconsistencies in data may lead to misleading analysis results. With the growth of data volume, data security and privacy issues have become increasingly important. Financial institutions need to ensure the integrity, confidentiality, and availability of data, while also complying with relevant privacy laws and regulations. Processing and analyzing big data requires advanced computer hardware, software, and algorithms, as well as professional talents capable of using these technologies[8]. But at the same time, financial big data has also brought

unprecedented opportunities to the financial sector. It enables financial institutions to better understand market dynamics, predict market trends, optimize investment strategies, improve risk management capabilities, and provide more personalized financial services. New market participants, trading patterns, and risk factors may emerge, which poses new challenges for financial regulation. With the increasing amount of data generated by exchanges, brokers, and other financial institutions every day, ensuring data integrity has become even more difficult[9]. Any errors or omissions in the process of data collection, storage, or transmission may lead to incomplete data and mislead quantitative strategies. Ensuring data accuracy is a challenge in the big data environment. Financial big data refers to the large amount of data generated and collected in the financial field, which has a high degree of diversity, speed, and accuracy, and its scale exceeds the processing capabilities of traditional database applications [10]. The characteristics of financial big data are shown in Figure 1.

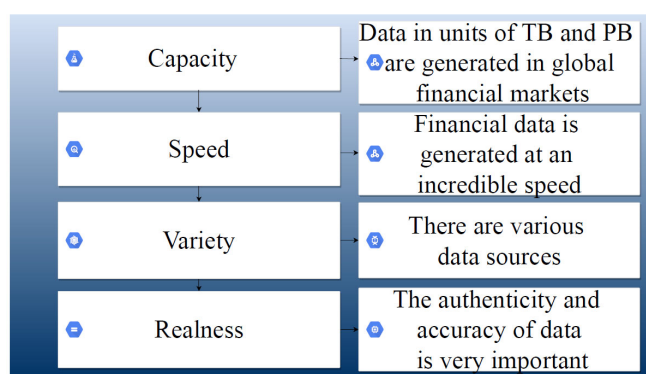


Figure 1. Characteristics of Financial Big Data

Incorrect data input, system errors, or external factors such as fraudulent behavior can all affect data quality. Investors must use various data cleansing and validation techniques, such as anomaly detection and data consistency check, to ensure the accuracy of the data. In high-frequency trading and algorithmic trading, data latency may lead to significant economic losses. Therefore, financial institutions are looking for faster data processing and transmission technologies, such as low latency networks and parallel processing technologies.

2.2. Data diversity

Financial big data is not only traditional structured data, such as stock price and trading volume. It also includes unstructured data, such as news reports, social media posts and online search queries. With the development of financial big data, quantitative strategies are increasingly diversified and personalized. Both institutional and individual investors can choose or customize the most suitable quantitative strategy according to their risk preferences, investment objectives and market views. By combining multiple forecasting factors, a more robust and effective investment strategy can be constructed. According to the change of market environment, the strategy parameters and weights are automatically adjusted. Customized strategies for each investor are optimized according to their personal risk preferences, investment objectives and market views. Natural language processing technology is used to analyze news reports and social media posts to extract emotions and intelligence related to financial markets. Investing based on

the response to a specific event. Big data technology allows investors to identify and respond to these events more quickly and accurately.

2.3. The challenge of computing power

With the increase of data, the traditional data processing and analysis methods are no longer applicable. Investors need more powerful computing power and more efficient algorithms to process, analyze and interpret these data. Use multiple compute nodes to process a large amount of data in parallel. For example, frameworks such as Hadoop and Spark allow users to run complex analysis tasks on large data sets. Analyze the data as it enters the system to gain real-time market insight. For example, stream processing technologies such as Kafka and Storm can process and analyze real-time data streams. Using advanced algorithmic models, such as neural networks and random forests, patterns and relationships are extracted from big data, thus improving investment strategies. Financial big data brings opportunities and challenges for quantitative investment strategies. Investors need to understand these influences and choose appropriate data processing and analysis methods according to their own needs and abilities.

3. Application of Big Data Technology in Quantitative Investment Strategy Optimization

3.1. Data cleaning and preprocessing

In the context of financial big data, the traditional quantitative investment strategy may not be able to fully utilize the existing huge data resources. Therefore, the introduction of big data technology is of decisive significance for policy optimization. With the wide application of big data, its threat to personal privacy has also increased. When financial institutions use these data to optimize their strategies, they must ensure that their behaviors conform to ethical standards and respect consumers' privacy rights. Transparency, fairness and non-discrimination of data become key considerations. Financial institutions should disclose their data sources, data processing methods and how to use these data. It should be ensured that the collection, processing and use of data will not lead to any form of discrimination or prejudice. Financial institutions should take measures to protect consumers' privacy, such as using data desensitization, encryption and anonymization technologies. For example, there are often abnormal values in financial data, such as "unconventional" trading behavior. These outliers are detected and processed by statistical methods and machine learning methods. The missing data may be due to various reasons. Methods such as linear interpolation, nearest neighbor filling, using regression model or filling method based on deep learning such as self-encoder. In order to ensure the effective comparison and integration between different data sources, the commonly used methods are Min-Max standardization, Z-score standardization and Decimal Scaling.

3.2. Characteristic engineering

Use statistical method or machine learning method to select the features that are most relevant to the predicted target. Carry out feature transformation: such as principal component analysis or linear discriminant analysis, to transform the original features into more informative forms.

Extracting information from unstructured data, for example, using TF-IDF, Word2Vec and other technologies to extract keywords or semantic information from news texts.

3.3. Machine Learning and Deep Learning

Using historical data to train models to predict future financial market changes, such as linear regression, decision trees, support vector machines, random forests, neural networks, etc. Utilize potential structures in the data for investment decisions, such as K-means, hierarchical clustering, DBSCAN, or self-organizing maps. The application of deep reinforcement learning in investment, such as Deep Q-Learning or Actor-Critic methods, learns the optimal strategy through continuous interaction with the environment.

3.4. Backtesting and simulation

Conduct policy backtesting on consecutive time windows to evaluate the performance of the policy at different time periods. Evaluate the performance of strategies across multiple asset classes, such as stocks, bonds, foreign exchange, etc. Evaluate the performance of strategies under different market conditions by simulating random processes. Evaluate the risk of a strategy using indicators such as Sharpe ratio, maximum fallback, Value at Risk, and Conditional Value at Risk. Overall, big data technology provides new optimization tools and methods for quantitative investment strategies. Financial institutions and investors need to keep up with the pace of technological development, constantly adjust and improve their strategies to fully utilize the value of big data.

4. Conclusions

The rise of financial big data indicates that the field of financial technology has entered a new stage of development. Big data not only provides participants in financial markets with richer and deeper market information, but also opens up new opportunities for the development of quantitative investment strategies. First of all, the data-driven decision-making model will continue to be favored. The traditional decision-making model based on experience and intuition is inadequate in the face of complex and changeable financial markets. On the contrary, the model based on big data can better capture the subtle changes in the market and provide strong decision support for investors. Secondly, machine learning and deep learning technology will play an increasingly important role in quantitative investment. These advanced algorithms can automatically learn and extract valuable information from a large number of data, which

greatly improves the adaptability and accuracy of investment strategies. The quality, security and privacy protection of data have become the focus of financial institutions and regulators. Ensuring the authenticity and integrity of data is the key to realize the sustained and healthy development of big data in the financial field. In short, financial big data provides a new path and tool for the optimization of quantitative investment strategies. But while giving full play to its potential, we must also pay attention to and deal with the risks and challenges it brings. Looking forward to the future, the integration of financial big data and quantitative investment strategy will be an important development direction in the field of financial technology, which deserves our continuous attention and in-depth study.

References

- [1] Sun Y, Lu Z. Financial Quantitative Big Data Platform Based on High Performance Computing[J]. IEEE, 2019, 42(15):29-35.
- [2] Jin Y, Wang H, Sun C. Offline Big or Small Data-Driven Optimization and Applications[J]. Post-Print, 2021, 8(3):36-41.
- [3] Kannan V. ijarcece financial markets-investment strategy through predictive and statistical analysis using big data and matlab[J]. Journal of Financial Compliance, 2021, 16(5):12-18.
- [4] Abur M, Murphy S. Why financial institutions need a master data strategy as a starting point for a RegTech strategy[J]. Journal of Financial Compliance, 2019, 3(2):31-37.
- [5] Wang S, Zang Y, Ge W, et al. Data-Driven Real-Time Pricing Strategy and Coordinated Optimization of Economic Load Dispatch in Electricity Market[J]. Journal of Physics: Conference Series, 2021, 25(14):13-19.
- [6] Miao B, Liu C, Li H, et al. Big Data Configuration Optimization Strategy Analysis of Thermal-Electric Hybrid Energy System Based on Energy Cascade Utilization[J]. Journal of Physics: Conference Series, 2021, 1852(4):042042-042049.
- [7] Yang, Jidong, Zhu, et al. A practical data-driven optimization strategy for Gaussian beam migration[J]. Geophysics Journal of the Society of Exploration Geophysicists, 2022, 36(14):17-21.
- [8] Guida T. Big Data and Machine Learning in Quantitative Investment[J]. Post-Print, 2019, 25(12):19-24.
- [9] Li J, Yuan Z Y. Research on the Application Strategy of Big Data in University Financial Management[J]. Journal of computational methods in sciences and engineering, 2020, 15(5):10-16.
- [10] Sun G. Quantitative investment prediction analysis for enterprise asset management using machine learning algorithms[J]. Journal of computational methods in sciences and engineering, 2022, 12(5):11-15.