GEORGE BROWN PRESS
JOURNALS, BOOKS & PROCEEDINGS

Economics and Management Innovation

Article

The Application of Big Data Analytics to Stock and Derivatives Trading Strategies

Yilin Fu [1,*]

[1] Brandeis International Business School, Brandeis University, Waltham, Massachusetts, 02453, USA
* Correspondence: Yilin Fu, Brandeis International Business School, Brandeis University, Waltham, Massachusetts, 02453, USA

Abstract: Big data analysis technology has gradually become a key trend in the development of the financial industry in improving trading strategies of stocks and their derivatives. Through the collection and in-depth analysis of a large amount of market information, investors can extract key intelligence, so as to optimize their trading choices and enhance the implementation efficiency of trading strategies. This paper explains the basic theory of price formation and risk management in the stock and derivative markets, and then analyzes the diversified application of big data technology in stock trading strategies, such as market prediction, stock screening, risk early warning model construction and quantitative trading practices. Then it discusses the specific role of big data in the field of derivatives trading, focusing on its application in the optimization of derivatives pricing model, market risk management, arbitrage strategy formulation and market trend prediction.

Keywords: big data analysis; stock trading; derivatives trading; market forecast; risk management

1. Introduction

Financial markets have evolved dramatically over the past few decades. The former manual trading method has been replaced by today's highly intelligent and algorithmic trading mechanism, and technological innovation has become an inexhaustible driving force for market change. In equity investing, the decision-making process has moved beyond a framework based solely on corporate financial reporting and traditional fundamental analysis. Investors can now more accurately predict stock price movements and quickly make investment decisions by deeply analyzing large amounts of market data, social network information, news updates, and more. Traders in the derivatives market also use big data analysis to improve pricing strategies, control risks, and explore arbitrage space. The emergence of data-driven investment models not only improves the execution efficiency of trading strategies, but also significantly reduces the risk of market volatility.

2. Theoretical Basis of Stock and Derivatives Trading

2.1. Pricing and Investment in the Stock Market

The pricing in the stock market is determined through complex market mechanisms and information interactions, which are not only directly affected by market demand and supply, but also influenced by the company's intrinsic value, economic environment, investor sentiment, and other multi-dimensional factors. According to the traditional efficient market hypothesis, the trading price of a stock is regarded as a comprehensive representation of all available information, and those unforeseeable price fluctuations are regarded as random events. However, in real trading, due to the existence of investors' mentality, poor information and other factors, the stock price will often deviate from its real value, which creates arbitrage space for investors [1]. When investing in stocks, investors

Received: 27 April 2025
Revised: 03 May 2025
Accepted: 23 May 2025
Published: 24 May 2025

usually use a combination of fundamental analysis and technical analysis to predict the movement of stock prices. Fundamental analysis mainly conducts in-depth research on the company's financial reports, industry development status, and overall economic situation, with the purpose of assessing the true value of stocks. Technical analysis is to predict the future trend of stock price by studying the dynamic pattern of historical stock price and trading volume [2].

2.2. Pricing and Risk Management of Derivatives Transactions

The trading mode of financial derivatives is a form of trading based on various basic assets, such as stocks, bonds, physical commodities, etc. The pricing of these products is influenced by the fluctuations of the underlying assets and other relevant market factors. The types of derivatives commonly available in the market include futures, options and swaps. When evaluating the value of these derivatives, mathematical models are often used to calculate, among which the most famous is the Black-Scholes formula, which mainly focuses on the pricing of options, and its calculation is based on variables such as the price of the underlying asset, the strike price, the expiration time, the volatility and the risk-free rate of return. In terms of risk management, derivatives trading can not only help investors avoid risks, but also expose them to significant risks. In hedging operations, investors aim to offset potential risks in the spot market through price changes in derivatives. Effective risk management methods include stop loss orders, diversification and hedging techniques. Instruments such as value at risk (VAR) and Greek letters such as Delta and Gamma are often used to measure and manage market, liquidity and credit risk in derivatives transactions [3].

3. Application of Big Data Analysis in Stock Trading Strategies

3.1. Big Data-Driven Market Prediction and Trend Identification

When trading stocks, anticipating market movements and tracking market trends form a key part of the trading strategy. In the past, people relied on historical trading records and chart patterns for technical analysis, but now, relying on the comprehensive analysis technology of big data, we can obtain more accurate market forecasts and faster reactions. The explosive growth of the amount of data makes the behavioral patterns of the stock market more variable and complex, and the traditional analysis methods appear to be inadequate. Therefore, big data analysis technology has emerged at a critical time and has been widely applied to market trend analysis, market prediction, and investment decision-making. This kind of prediction technology based on big data mainly relies on machine learning programs, natural language processing technology, and sentiment analysis. They can conduct real-time analysis of a large amount of news information, social media data, financial media content, and market data to mine key information, and then predict stock price changes and future trends. For example, sentiment analysis tools can quantitatively evaluate the sentiment value of financial news and analysis reports, reveal the market's sentiment changes for specific stocks or industries, and then provide investors with more accurate trading strategy support. The following Table 1 shows the main application areas of big data in market forecasting [4].

Table 1. Application Methods of Big Data in Market Forecasting.

Application field	Concrete method	Effect
Data mining and pattern recognition	Machine learning, deep learning	Identify potential patterns and patterns in the market to improve the accuracy of trend prediction
Sentiment analysis	Natural language processing (NLP), sentiment analysis	Analyze changes in market sentiment and predict short-term price movements

Time series data analysis	Time series analysis, ARIMA model	Identify long-term trends and cyclical fluctuations to provide sound investment decisions
Social media analysis	Social media data capture, emotion monitoring	Use social platform data to predict market sentiment and identify potential market changes in advance

With these technical methods, the deep mining of big data not only enhances the accuracy of market trend prediction, but also allows investors to quickly insight into market opportunities, and thus improve the operational plan of stock investment.

3.2. Application of Big Data in Stock Selection and Screening

In the work of stock selection and screening, advanced big data technology gives investors more comprehensive and precise decision support. Unlike traditional data sources such as financial reports and technical indicators, big data can integrate and process diversified data resources, including real-time stock market transaction information, corporate announcements, media news, and discussions on social networks, thus providing investors with a richer set of information. With big data technology, investors can collect and analyze huge amounts of information in real time to uncover potential investment value. First of all, big data technology will use crawlers, API interfaces and other methods to collect data from multiple channels, such as corporate financial statements, real-time changes in stock market prices, the latest news of the industry and investor sentiment fluctuations [5]. The collected data is often unstructured and must be cleaned, preprocessed, and converted into a standardized data format to facilitate subsequent analysis. Then, through machine learning and natural language analysis techniques, the system is able to extract key features from a wide range of data. Figure 1 below shows the flow of big data in stock screening.

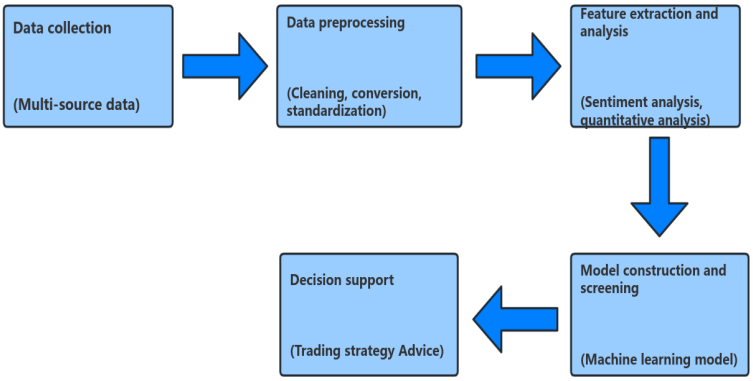

Figure 1. Application Process of Big Data in Stock Screening.

Through this process, investors are not limited to using traditional financial indicators for evaluation, but can also use the real-time analysis function of big data to gain insight into the subtle fluctuations of the market, and then implement more accurate investment choices.

3.3. Application of Big Data Analysis in Risk Prediction Model

In the trading of stocks and derivatives, risk management is the core of strategy. With the rapid progress of big data technology, the old risk management methods have gradually been replaced by data-based risk prediction algorithms. These algorithms help traders accurately identify potential risks in the market and respond quickly through comprehensive analysis of diversified information such as historical trading data, market sentiment fluctuations, and macroeconomic indicators. The application of big data in risk prediction

shows its unique advantages, such as the use of machine learning algorithms (decision trees, support vector machines, deep neural networks) to independently analyze a large number of historical market data and dig out the risk characteristics of the market. These algorithms can effectively predict the market's future volatility trends, price movements, and potential risk exposure. By integrating multiple factors such as market price dynamics, trading volume changes, and macroeconomic indicators, machine learning algorithms can provide investors with accurate risk prediction services in a volatile market environment. Table 2 below shows several common risk prediction models and their characteristics.

Table 2. Common Risk Prediction Models and Their Characteristics.

Model type	Peculiarity	Application field	Advantage
Machine learning model	Automatically identify market patterns and predict future price fluctuations	Stock, futures, options trading	Handles large amounts of data and optimizes itself
Emotion analysis model	Predict market sentiment and risk through text data	Stock market, foreign exchange market	Reflect market sentiment changes in real time
Multifactor regression model	Multifactor regression analysis of asset price risk exposure	Securities portfolio, risk assessment	Precise quantification of risk exposure for a wide range of applications

Relying on the risk prediction model driven by big data, investors can find potential risks in advance in the rapidly changing market, and then implement appropriate risk management methods according to the forecast data, so as to effectively manage investment risks and improve trading decisions.

3.4. Application of Big Data in Quantitative Trading and Strategy

Quantitative trading relies on advanced computer programs and mathematical models to automate trading decisions, aiming to improve the efficiency of trading by mining market information. Under this framework, the use of large-scale data is crucial, which provides sufficient data resources for improving trading strategies. In the field of quantitative trading, the main function of large-scale data is concentrated on the collection and analysis of information. The financial market is filled with a large amount of financial information, such as historical transaction prices, turnover, macroeconomic data, corporate financial reports, etc. Traditional manual analysis is difficult to digest this information efficiently, while large-scale data processing technology can quickly capture and analyze these complex data, laying a solid foundation for trading strategies. Digging deep into these data, quantitative strategies can uncover deeper patterns and movements in the market, thereby enhancing the accuracy of forecasts. In addition, with the help of machine learning and deep learning algorithms, extracting potential signals from data becomes a reality.

4. Application of Big Data Analytics in Derivatives Trading Strategies

4.1. Optimization of Derivative Pricing Model

How derivatives are priced is a central issue in the world of financial engineering. Conventional pricing methods are often based on classical mathematical theories, such as Black-Scholes models, but these traditional models are unable to cope with the non-linear characteristics of the market and volatility changes. With the rise of big data technology, using its analytical ability to improve the pricing mechanism of derivatives has gradually become an innovative method in the financial industry. Through the in-depth mining of past market data, investors can more accurately grasp the law of derivative price fluctuations, and then improve pricing strategies. Thanks to big data technology, we are able to

take into account more influential factors and characteristics, such as macroeconomic indicators, industry development trends, market sentiment, etc., to build more refined derivative pricing models. Black-Scholes option pricing model formula:

$$C = S_0 \cdot N(d_1) - Ke^{-rT} \cdot N(d_2) \quad (1)$$

Among them, C is the option price, S_0 Current price of the underlying asset, K is the option exercise price, r is the risk-free rate, T is the expiration time of the option, N(·) is the cumulative distribution function of the standard normal distribution, d_1 and d_2. They are:

$$d_1 = \frac{ln(S_0/K)+(r+\sigma^2/2)T}{\sigma\sqrt{T}},\ d_2 = d_1 - \sigma\sqrt{T} \quad (2)$$

In the traditional model, volatility is usually treated as a fixed value. However, using big data analysis, dynamic adjustments to volatility can be made with real-time market data and historical price data for more accurate pricing. This dynamic adjustment method not only improves the accuracy of the pricing model, but also captures the sudden changes in the market, so that the pricing of derivatives is more in line with the actual market conditions.

4.2. Derivatives Market Risk Management

Risk management in the derivatives market plays an important role in the trading strategy. With the help of advanced big data analysis technology, financial institutions can gain a deeper insight into market risks and implement effective countermeasures. By collecting and analyzing vast amounts of market transaction information, investor behavior, and macroeconomic indicators, financial institutions can assess risk more accurately and in real time. Big data analysis technology strengthens the ability to monitor and predict market volatility. By processing data such as historical price, volume, and money flow, machine learning algorithms can reveal patterns of price movements, abnormal fluctuations, and potential risks. This has led to improvements in derivatives pricing models that allow them to reflect market dynamics more flexibly, thereby improving the accuracy of risk warnings and the timeliness of adjustments. By exploring the interrelationships between different derivatives, investors can find the most appropriate hedging strategy in asset portfolio management to reduce market risk. In the options and futures markets, investors can use big data analysis to determine the appropriate hedging ratio to mitigate the impact of market fluctuations on the portfolio and ensure the robustness of the portfolio. In the field of quantitative risk management, big data analysis provides strong data support for the construction and improvement of risk control models.

4.3. Arbitrage Strategy Based on Big Data

An arbitrage strategy is a strategy of low-risk trading by taking advantage of market price differences. In the field of financial derivatives, arbitrage techniques based on big data have been widely used, especially in futures, options and other markets. With big data technology, traders can track and analyze key indicators such as price dynamics, trade size and market depth across markets in real time to explore arbitrage potential. Specifically, arbitrage methods based on big data can be roughly divided into several forms such as cross-market arbitrage, cross-variety arbitrage and time series arbitrage. Big data analytics can quickly process information from different exchanges and catch anomalies in price movements in real time, thus seizing fleeting arbitrage opportunities. With the advancement of machine learning and artificial intelligence technology, arbitrage methods based on big data can more accurately mine complex connections and potential irrational factors in the market, thereby improving the success rate of arbitrage strategies.

4.4. Market Forecasting in Derivatives Transactions

When trading derivatives, market prediction is a core part of building an efficient trading strategy. The use of massive data analysis, including the collection of past market data, real-time trading dynamics, and various peripheral information (such as social platforms and news), provides sufficient data support for prediction algorithms, assisting investors in grasping market dynamics more accurately. Common techniques for market prediction include machine learning programs and time series analysis. In machine learning, investors can train regression models based on historical data to predict possible future changes in the price of derivatives. The linear regression model can be expressed as:

$$P_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_n X_{t-n} + \in t \quad (3)$$

Among them, P_t representative time t Moment derivatives prices, X_{t-1}, X_{t-2}, …, X_{t-n} is the historical data that affects price changes, β_0, β_1, …, β_n is the regression coefficient, $\in t$ is the error term. In addition, time series analysis methods such as the autoregressive integrated moving average (ARIMA) model predict future prices by capturing temporal patterns and autocorrelations in price changes. The basic ARIMA model is expressed as:

$$Y_t = c + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \theta_1 \in_{t-1} + \theta_2 \in_{t-2} + \cdots + \theta_q \in_{t-q} + \in_t \quad (4)$$

Among them, Y_t is time$t$The price of the moment, ϵ_t is white noise, $\emptyset_1$, …, $\emptyset_p$. It's an autoregressive coefficient, θ_1, …, θ_q. It's the moving average coefficient. Through these predictive models, traders can effectively capture the trend changes of the market and make more scientific trading decisions, thus improving the success rate of derivatives trading.

5. Conclusion

In the field of equities and derivatives, big data technology has played a significant role in promoting the innovation and improvement of trading tactics. Through in-depth mining of a large amount of market information, investors can more accurately implement market prediction, asset allocation and risk control, and then enhance the rationality and efficiency of trading choices. In stock trading, big data technology effectively promotes trend identification and stock screening, and enhances the implementation of quantitative trading strategies. In derivatives trading, big data technology provides a new perspective and solution for the improvement of price models, market risk monitoring and the formulation of arbitrage strategies. With the continuous development of artificial intelligence and machine learning technology, big data will be more widely used in the financial field in the future, helping investors better adapt to the changing and complex financial market conditions.

References

1. R. Marciano, V. Lemieux, M. Hedges, M. Esteva, W. Underwood, M. Kurtz, et al., "Archival records and training in the age of big data," in *Re-Envisioning the MLS: Perspectives on the Future of Library and Information Science Education*, vol. 44, Emerald Publishing Limited, 2018, pp. 179–199. ISBN: 9781787548848.
2. N. M. Reddy, A. K. Budati, S. Islam, and G. Ramesh, "Enhanced elliptic curve-Diffie Hellman technique with big data analytics for satellite image security enhancement in Internet of Things systems," *Earth Sci. Inform.*, vol. 17, no. 1, pp. 711–723, 2024, doi: 10.1007/s12145-023-01194-2.
3. Y. Jiang, T. Feng, and Y. Huang, "Antecedent configurations toward supply chain resilience: the joint impact of supply chain integration and big data analytics capability," *J. Oper. Manag.*, vol. 70, no. 2, pp. 257–284, 2024, doi: 10.1002/joom.1282.
4. Y. Liu, Y. Bao, Y. Huang, M. Zhang, X. Li, et al., "How the National Health Insurance Coverage policy changed the use of lenvatinib for adult patients with advanced hepatocellular carcinoma: a retrospective cohort analysis with real world big data," *Int. J. Equity Health*, vol. 22, no. 1, p. 256, 2023, doi: 10.1186/s12939-023-02052-9.

5. P. Wangchuk, F. Jie, and K. Wangdi, "Big data analytics in the procurement process: organizational alignment and the behavior of procurement professionals in Bhutanese procuring agencies," *J. Data Inf. Manag.*, vol. 6, no. 1, pp. 15–27, 2024, doi: 10.1007/s42488-023-00109-6.