Optimizing Stock Trend Prediction in the Chinese Market: A Comparative Study of Machine Learning Models with Bayesian Hyperparameter Tuning

Yiming Xiang

Financial Engineering, The Chinese University of Hong Kong, Shenzhen, 518172, China
122020222@link.cuhk.edu.cn

Abstract. Predicting stock market trends has become increasingly vital in today's volatile global economy. This paper addresses the challenge of accurate stock market prediction, particularly within the Chinese stock market, which has been subject to unique regulatory and economic shifts. This paper explores the application of five machine learning models to predict stock trends, comparing their performances and enhancing accuracy through hyperparameter tuning. This paper utilized a dataset from the CSI 300 Index, applying various feature engineering techniques and standardizing inputs. After training and backtesting the models, Bayesian optimization is employed to fine-tune the top-performing models: LightGBM, XGBoost, and GRU. This optimization process focuses on improving the models' annualized returns, Sharpe ratios, and drawdowns. Each model's predictions were tested on a backtest dataset from 2024, and key performance metrics were recorded. The initial results indicated that LightGBM, XGBoost, and GRU outperformed the other models, with LightGBM achieving an annualized return of 16.44%, XGBoost 11.44%, and GRU 16.11%. After Bayesian optimization, LightGBM improved to 20.97%, XGBoost to 17.20%, and GRU reached 27.22%, though with a higher drawdown. These results demonstrate that GRU's risk-adjusted performance can offer significant returns, while LightGBM strikes a balance between risk and return. Future work could extend the dataset beyond the CSI 300 Index for better generalizability. While Bayesian optimization improved gradient boosting models, further tuning of neural networks and exploring advanced architectures like Transformers may enhance performance. Integrating real-world factors, such as transaction costs and macroeconomic indicators may strengthen their practical applications.

Keywords: Machine Learning, Stock Trends Prediction, Bayesian Optimization.

1 Introduction

In the current context of global economic uncertainty and challenges, predicting stock market trends has become an indispensable tool for investors and policymakers. Amid international economic volatility, marked by rising global inflation, geopolitical

J. Lu (ed.), *Proceedings of the International Workshop on Navigating the Digital Business Frontier for Sustainable Financial Innovation (ICDEBA 2024)*, Advances in Economics, Business and Management Research 315,
https://doi.org/10.2991/978-94-6463-652-9_63

tensions, and increased market fluctuations, the importance of accurately forecasting stock market trends has reached an unprecedented level. Within this challenging landscape, the Chinese stock market presents unique characteristics and challenges.

Recent regulatory changes have significantly impacted China's Initial Public Offering (IPO) market. In the first half of 2024, the total funds raised via IPOs on China's major exchanges fell sharply to $4.01 billion from $29.80 billion during the same period in 2023, marking the lowest half-year result since 2020 [1]. This decline reflects a tightening regulatory environment, where authorities such as the China Securities Regulatory Commission (CSRC) and the State Council have implemented new rules aimed at enhancing the quality of new listings and strengthening market stability. However, this shift towards prioritizing "quality over quantity" has also led to a cooling of the IPO market, underscoring the evolving dynamics within the Chinese financial landscape.

In such a complex environment, the capacity to accurately predict stock trends is particularly critical. Given the significant influence of external economic forces on China's market behavior, there is a pressing need for advanced analytical tools capable of navigating these complexities. Recent advancements in machine learning, especially models like LightGBM and XGBoost, have shown considerable promise in enhancing the accuracy of stock market predictions by efficiently handling large, complex datasets [2]. As noted in recent studies, these models have been successfully applied to various markets such as the US and Morocco, demonstrating their ability to improve the identification and interpretation of stock price fluctuations across different economic contexts [3]. However, despite their growing recognition and effectiveness in global applications, the application of LightGBM and XGBoost within the specific context of the Chinese stock market remains underexplored. In addition to these models, neural network architectures — such as Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and Convolutional Neural Networks (CNN) — have also been extensively used for time series forecasting in financial markets due to their ability to capture complex temporal patterns and dependencies in financial data [3]. However, a comprehensive evaluation that compares gradient boosting models like LightGBM and XGBoost with neural network models in the Chinese stock market has not yet been fully explored in the existing literature.

This study aims to fill this gap by conducting a comparative empirical analysis of LightGBM, XGBoost, and three neural network models (LSTM, GRU, and CNN) to evaluate their effectiveness in predicting stock market trends within the Chinese context. By doing so, this research seeks to contribute to the broader understanding of these advanced predictive models and offer valuable insights for investors, policymakers, and researchers engaged in financial market analysis. Furthermore, this study not only aims to enhance stock market prediction methodologies but also to explore the potential synergies of combining different machine learning and neural network models to optimize forecasting accuracy across diverse market contexts.

The structure of the remainder of this paper is organized as follows: Section 2 discusses the methodology, and the dataset utilized in this study. In Section 3, this study analyzed the experimental results and explored potential explanations for the observed outcomes. Finally, Section 4 concludes the paper by summarizing the findings and outlining directions for future research.

2 Method

This section will describe how the data was prepared for model training and testing, including dataset details, feature engineering, and preprocessing steps.

2.1 Dataset

In this study, the dataset comprises stocks from the CSI 300 Index, representing the largest and most liquid companies listed on the Shanghai and Shenzhen Stock Exchanges. To ensure data quality, stocks with more than 30% missing values were excluded from the analysis. After this filtering process, a total of 80 stocks were selected to form the stock pool for this study. The data was collected from the ifind platform, covering the period from January 1, 2021, to July 30, 2024. In total, the dataset includes 70,319 data points, capturing stock prices and company characteristic indicators across this timeframe. The dataset provides comprehensive coverage of the Chinese stock market during a period marked by significant economic and regulatory changes. The target variable in this study is a binary classification indicating the daily stock price trend. Specifically, the variable takes on a value of 1 if the stock's closing price increases on a given day compared to the previous day (upward trend) and 0 if the closing price decreases or remains the same (downward trend). This binary setup allows for a clear distinction between positive and negative movements in stock prices, forming the basis for predicting market direction.

2.2 Factor Engineering

Factors in this study are categorized into two types: fundamental factors and technical factors. A preliminary factor library of 22 factors was selected, consisting of 8 fundamental factors and 14 technical factors, all chosen based on their relevance to stock market behavior and predictive power.

Technical Factors. Technical factors shown in Table 1 focus on price trends, liquidity, and volatility, which are key drivers of short-term stock price movements [4]:

Table 1. Table for Technical Factors

Type	Indicator Name	Description
Trend Indicators	ten_day_MA	10-day's Moving Average
	twenty_day_MA	20-day's Moving Average
Trend-following Indicators	Signal_line	9-day Exponential Moving Average of MACD
	MACD	Moving Average Convergence Divergence
Volume Indicators	volume	Trading Volume
	turn_over_rate	Turnover Rate
	VSTD10	10-day trading volume standard deviation
Momentum Indicators	momentum	Changing rate of stock prices in the past 30 days
Strength Indicators	RSI_7_Day	Relative Strength Index
Volatility Indicators	ten_day_volatility	10-day volatility standard deviation
	twenty_day_volatility	20-day volatility standard deviation
Bollinger Bands Indicators	Bollinger_Upper_20_day	Upper Bollinger Band for a 20-day period
	Bollinger_Lower_20_day	Lower Bollinger Band for a 20-day period
Average True Range Indicator	ATR_14_Day	14-day moving average of stock price volatility

Fundamental Factors. Fundamental factors shown in Table 2 describe the operating conditions of companies through financial data, focusing on size, profitability, and valuation [5]:

Table 2. Table for Fundamental Factors

Type	Indicator Name	Description
Size	market_cap	Market Capitalization
Profitability	PE	Price Earnings Ratio
	ROE	Rate of Return on Common Stockholder's Equity
	ROA	Return on Assets
	EPS	Earnings Per Share
Valuation	PS	Price-to-Sales ratio
	DP	Dividend Payout ratio
	PB	Price-to-Book ratio

Factor Screening and Selection Process. To ensure the relevance and effectiveness of these factors, a rigorous screening process was employed:

- Stationarity Testing: Each factor was tested for stationarity to ensure stability over time.
- Correlation Testing: A correlation matrix was used to identify and eliminate factors with strong correlations to reduce multicollinearity. After this process, 10 factors were retained for further analysis.
- Multiple Linear Regression: The 10 retained factors were then subjected to multiple linear regression against logarithmic returns. Factors with insignificant regression coefficients were eliminated, leaving 8 final factors.

The 8 selected factors shown in Table 3 that demonstrate strong explanatory power and weak self-correlation were then used in the subsequent machine learning models：

Table 3. Table for Chosen Factors

Type	Indicator Name	Description
Profitability	PE	Price Earnings Ratio
Volume Indicators	volume	Trading Volume
Size	market_cap	Market Capitalization
Momentum Indicators	momentum	Changing rate of stock prices in the past 30 days
Trend Indicators	twenty_day_MA	20-day's Moving Average
Strength Indicators	RSI_7_Day	Relative Strength Index
Trend-following Indicators	MACD	Moving Average Convergence Divergence
Average True Range Indicator	ATR_14_Day	14-day moving average of stock price volatility

2.3 Preprocessing

Handling Missing Values. Missing values in the dataset were handled using a forward-filling approach to maintain the continuity of the time series data. This method ensures that any gaps in the data are filled with the most recent available value, preserving the temporal structure required for model training and prediction.

Feature Scaling. In this study, the feature set was standardized using the StandardScaler [6], which transforms the data such that each feature has a mean of 0 and a standard deviation of 1. This step is crucial for ensuring that all features are on a consistent scale, preventing features with larger magnitudes from disproportionately influencing the model's predictions.

Data Splitting. The dataset was divided into two distinct sets: 1) Training Set: The training set comprises data from January 2021 to December 2023, which was used to train the predictive models. The training set was divided into a training and test set in a 4:1 ratio. This split was used to train the predictive models and evaluate their performance. 2) Backtesting Set: The backtesting set includes data from January 2024 to July 2024, and was used to simulate real-world conditions for validating the model's performance. This period was chosen to test the robustness of the model and its ability to generalize about new, unseen market conditions.

Time-Series Data. For time-series prediction, the dataset was transformed into sequences of past trading data, enabling the models to capture temporal dependencies essential for stock trend prediction. The sequence length used was 30 days for all models (LSTM, GRU, and CNN). The time-series data was then structured such that each input sample contained data from the previous 30 trading days. For LSTM and GRU, the input data was converted using TimeseriesGenerator, while for CNN, the data was reshaped into 3D arrays to match the expected input shape.

2.4 Machine Learning-Based Prediction

LightGBM. LightGBM is a gradient-boosting framework known for its ability to efficiently handle large datasets and capture complex relationships between features [7]. It builds decision trees in a leaf-wise manner. In this study, key hyperparameters such as num_leaves, learning_rate, feature_fraction, and bagging_fraction were carefully set to optimize the model ' s performance. These hyperparameters were chosen based on their impact on model complexity, learning speed, and feature sampling, with the goal of striking a balance between accuracy and generalization.

XGBoost. XGBoost is another popular gradient-boosting algorithm that uses regularization techniques to prevent overfitting, making it well-suited for stock trend prediction [8]. In this study, key hyperparameters such as max_depth, learning_rate, subsample, and colsample_bytree were carefully set to optimize the model's ability to handle complex datasets while preventing overfitting. These parameters were chosen to balance model complexity and generalization, ensuring the model could accurately predict stock trends.

LSTM. Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) designed to handle long-term dependencies in sequential data. They are particularly suited for time-series forecasting tasks like stock trend prediction, where capturing temporal patterns is critical [9]. The LSTM model includes two 50-unit LSTM layers and a sigmoid-activated Dense output layer for binary classification. It was compiled with the Adam optimizer and binary cross-entropy loss. The model was trained for 20 epochs with a batch size of 32, using a 10% validation split to monitor performance.

GRU. Gated Recurrent Unit (GRU) networks are a type of recurrent neural network (RNN) that, like LSTMs, are designed to handle long-term dependencies in sequential data. GRUs are computationally more efficient as they use fewer parameters than LSTMs, making them suitable for time-series prediction tasks like stock trend forecasting [10]. The GRU model includes two 50-unit GRU layers with 0.2 dropout, followed by a sigmoid-activated Dense output layer. It was compiled with the Adam optimizer and binary cross-entropy loss. The model was trained for 20 epochs with a batch size of 32, using 30-day sequences and a validation split for performance monitoring.

CNN. Convolutional Neural Networks (CNNs) are widely used in image recognition tasks but can also be applied to time-series data. CNNs capture local patterns within sequences and are particularly effective at recognizing short-term trends in stock prices [11]. The CNN model includes a 1D convolutional layer with 64 filters and a kernel size of 5, followed by MaxPooling1D, Flatten, and a 50-unit Dense layer with a sigmoid output for binary classification. The model was trained for 10 epochs with a batch size of 32, using 30-day sequences and a validation split for performance monitoring.

2.5 Prediction and Backtesting

Prediction accuracy for each model (LightGBM, XGBoost, LSTM, GRU, CNN) was calculated using the 2021 to 2023 training and test sets. The top 30 stocks with the highest prediction accuracy were selected for each model. The weights of the selected stocks in the portfolio were assigned based on their respective prediction accuracy.

During the backtesting phase, for each day in 2024, the factors for each stock were input into the models to generate predictions. Stocks with predicted values greater than or equal to 0.5 were assigned a long position, while those with predicted values less than 0.5 were assigned a short position. The portfolio, constructed based on these daily predictions, was then backtested using the 2024 dataset. The key performance metrics, including annualized returns, maximum drawdown, and the Sharpe ratio, were calculated to evaluate each model's effectiveness.

3 Results and Discussion

3.1 The Performance of LightGBM

The LightGBM model demonstrated strong performance in predicting stock trends within the Chinese market, as reflected in several key metrics: 1) Annualized Return: 16.44% 2) Maximum Drawdown: -4.12 % 3) Sharpe Ratio: 1.55.

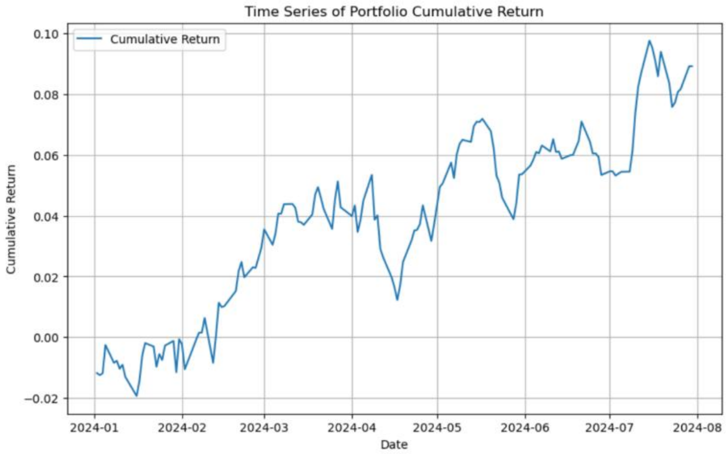

Fig. 1. Cumulative Return for LightGBM Model (Photo/Picture credit: Original).

As shown in the Fig. 1, the portfolio's cumulative returns show a steady increase from January 2024 to July 2024. The performance was relatively stable, with notable upward momentum starting around mid-February, peaking around May 2024 before a slight dip, and then resuming an upward trend. The portfolio's cumulative return peaked at approximately 8.5% in July 2024. This steady increase indicates the effectiveness of the LightGBM model in identifying stock price trends and making appropriate long or short positions over time. The observed dip in May 2024 aligns with general market corrections or volatility during that period, but the model managed to recover and continue generating positive returns.

3.2 The Performance of XGBoost

The XGBoost model displayed slightly lower performance compared to LightGBM, particularly in terms of returns and Sharpe ratio. However, it still showed promising results for stock trend prediction. Key performance metrics for XGBoost are: 1) Annualized Return: 11.44% 2) Maximum Drawdown: -3.39% 3) Sharpe Ratio: 1.21.

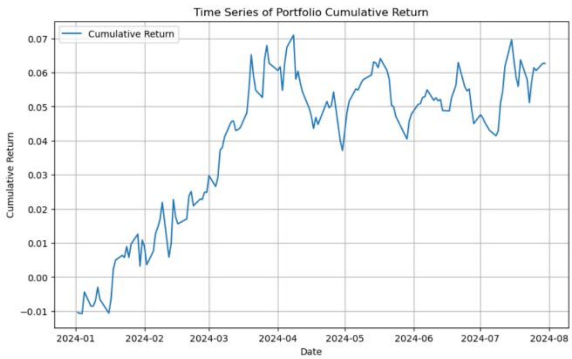

Fig. 2. Cumulative Return for XGBoost Model (Photo/Picture credit: Original).

As illustrated in the Fig. 2, from the beginning of 2024 to July 2024, the portfolio exhibited strong growth early on, particularly during February and March, peaking around May with cumulative returns reaching approximately 7%. Afterward, some fluctuations and volatility occurred, but the portfolio managed to stabilize, ending the period with cumulative returns around 6%. The model's ability to maintain a steady return curve, with recoveries after dips, indicates that XGBoost is effective at navigating through market volatility, especially considering the external economic uncertainties in China during this period.

3.3 The Performance of LSTM

The Long Short-Term Memory (LSTM) model, designed to capture long-term dependencies in time series data, produced the following key performance metrics during the backtesting period: 1) Annualized Return: 6.65% 2) Maximum Drawdown: -4.17% 3) Sharpe Ratio: 0.74.

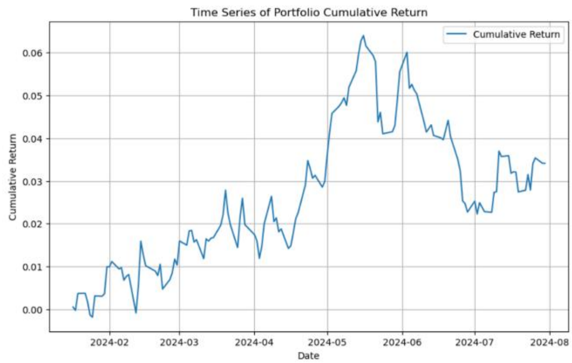

Fig. 3. Cumulative Return for LSTM Model (Photo/Picture credit: Original).

As shown in the Fig. 3, the LSTM model's cumulative returns were generally positive but exhibited higher volatility than expected, particularly in May and June 2024. Initially, the portfolio's cumulative returns steadily increased, reaching approximately 6% by early May. However, from mid-May onwards, the portfolio experienced a significant drawdown, with returns declining sharply before partially recovering in July. This pattern suggests that while the LSTM model was able to identify some upward stock trends early in the year, it struggled to adapt to sudden market shifts during periods of increased volatility. The model's performance demonstrates its ability to capture temporal dependencies in stock price movements, but it may be prone to higher risk in environments with rapid changes.

3.4 The Performance of GRU

The Gated Recurrent Unit (GRU) model, which is another variant of recurrent neural networks (RNNs), produced the following key performance metrics during the backtesting period: 1) Annualized Return: 16.11% 2) Maximum Drawdown:-5.28% 3) Sharpe Ratio: 1.46.

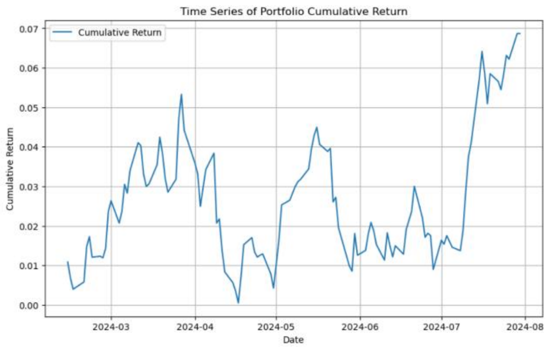

Fig. 4. Cumulative Return for GRU Model (Photo/Picture credit: Original).

The Fig. 4 for the GRU model shows a strong increase in portfolio returns, particularly in the later stages of the backtesting period. The model exhibited significant volatility in the early months of 2024, with multiple peaks and troughs. The GRU model showed rapid gains, pushing cumulative returns to nearly 7% by the end of the backtesting period. This late surge suggests that the GRU model was able to capitalize on favorable market conditions toward the end of the period.

3.5 The Performance of CNN

The Convolutional Neural Network (CNN) model, which was applied to capture short-term patterns in stock prices, exhibited the following key performance metrics during the backtesting period: 1) Annualized Return: 2.49% 2) Maximum Drawdown:-1.31% 3) Sharpe Ratio: 0.12.

The Fig. 5 for the CNN model shows a lower return trend overall compared to the other models. Starting from January 2024, the portfolio fluctuated between positive and negative returns, showing a lack of consistent upward movement. Despite some sharp upward movements in March and late April, the portfolio was unable to sustain gains and reverted to near-zero returns by early May. After May, the model continued to display significant volatility, and while it maintained a positive return by the end of the period, cumulative returns remained low at around 1% by July 2024. The CNN model demonstrated difficulty in identifying stable trends, leading to suboptimal returns.

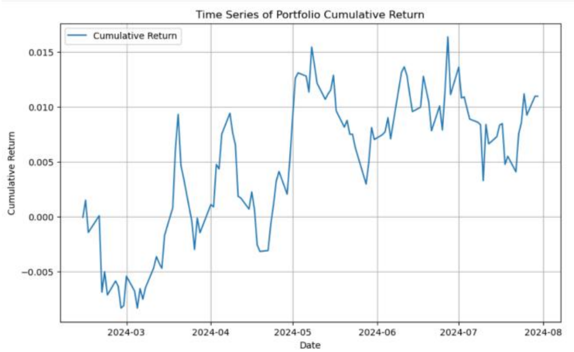

Fig. 5. Cumulative Return for CNN Model (Photo/Picture credit: Original).

3.6 Comparative Summary

Table 4. Table for key performance metrics

Model	Annualized Return (%)	Maximum Drawdown (%)	Sharpe Ratio
LightGBM	16.44	-4.12	1.55
XGBoost	11.44	-3.39	1.21
LSTM	6.65	-4.17	0.74
GRU	16.11	-5.28	1.46
CNN	2.49	-1.31	0.12

During the backtesting period, LightGBM and GRU emerged as the most successful models in terms of annualized returns, generating 16.44% and 16.11% shown in Table 4, respectively. XGBoost performed moderately well with an annualized return of 11.44%, but it fell short compared to the top two models. LSTM and CNN, on the other hand, showed significantly lower returns, with 6.65% and 2.49%, respectively, indicating their limitations in producing high returns. In terms of maximum drawdown, CNN had the smallest drawdown at -1.31%, reflecting its strong risk management capabilities, though it sacrificed potential gains in the process. XGBoost had a relatively modest drawdown of -3.39%, positioning it as a balanced model for risk and return. LightGBM and LSTM experienced similar drawdowns, hovering around -4.12% and -4.17%, while GRU had the highest drawdown at -5.28%, highlighting its riskier nature compared to the other models.

When analyzing the Sharpe ratio, LightGBM again led the pack with a ratio of 1.55, making it the best model for delivering risk-adjusted returns. GRU followed closely with a Sharpe ratio of 1.46, demonstrating that even though it had the highest drawdown, it still managed to strike a good balance between risk and reward. XGBoost achieved a moderate Sharpe ratio of 1.21, showing solid performance, albeit not as optimal as LightGBM and GRU. LSTM and CNN had much lower Sharpe ratios, at 0.74 and 0.12, respectively, suggesting their struggles in generating meaningful returns relative to their risk levels. LightGBM and GRU clearly stand out as the top performers in terms of both return generation and risk-adjusted performance, though GRU's higher volatility may pose a concern. XGBoost offers a balanced approach with moderate returns and well-managed risk. LSTM and CNN lagged behind, with CNN particularly underperforming in returns, despite its excellent risk management capabilities.

3.7 Optimized LightGBM, XGBoost, and GRU Results

Bayesian optimization was applied to the three best-performing models from the previous experiments: LightGBM, XGBoost, and GRU. This approach efficiently tunes hyperparameters by narrowing the search space and focusing on the most promising areas, significantly improving model performance [12]. Below are the updated performance metrics for each model after optimization, followed by the cumulative return charts and individual model analysis.

Optimized LightGBM. Key performance metrics: 1) Annualized Return: 20.97% 2) Maximum Drawdown: -5.11% 3) Sharpe Ratio: 1.68. The cumulative return for LightGBM, as shown in the Fig. 6, indicates a steady upward trend, particularly between March and May 2024, where the portfolio experienced significant growth. There was a brief period of volatility in mid-May, but the model quickly recovered, culminating in a strong positive return by the end of the backtesting period.

Optimized XGBoost. Key performance metrics: 1) Annualized Return: 17.20% 2) Maximum Drawdown: -5.51% 3) Sharpe Ratio: 1.39. XGBoost shows a more volatile performance than LightGBM, as visible in the Fig. 7. While the model generated strong returns early in the backtesting period, particularly in February and March, it experienced periods of drawdown that flattened overall returns by mid-year. However, it managed to regain momentum in the last two months, finishing with respectable cumulative returns.

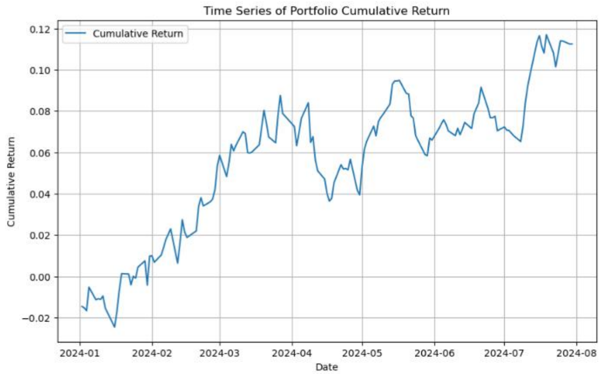

Fig. 6. Optimized Cumulative Return for LightGBM Model (Photo/Picture credit: Original).

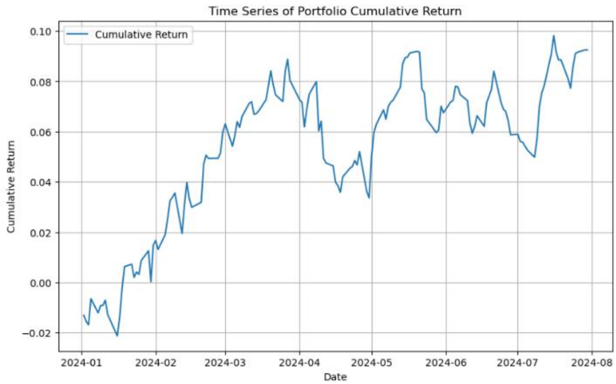

Fig. 7. Optimized Cumulative Return for XGBoost Model (Photo/Picture credit: Original).

Optimized GRU. Key performance metrics: 1) Annualized Return: 27.22% 2) Maximum Drawdown: -5.27% 3) Sharpe Ratio: 2.23. The GRU model, shown in the Fig. 8, demonstrated the highest cumulative return among the three models. It displayed considerable volatility early in the year, with several sharp drawdowns. However, starting from June, the model surged dramatically, ending with the highest cumulative return of the group. This performance highlights the model's ability to capitalize on market opportunities, though it carried higher risk earlier in the year.

After the Bayesian optimization, all three models demonstrated improved performance in both annualized returns and Sharpe ratios. LightGBM continued to provide a balanced return-risk tradeoff, while GRU stood out as the best performer in terms of absolute returns but with higher early volatility. XGBoost offered a solid middle ground, performing well but not exceeding LightGBM in risk-adjusted returns.

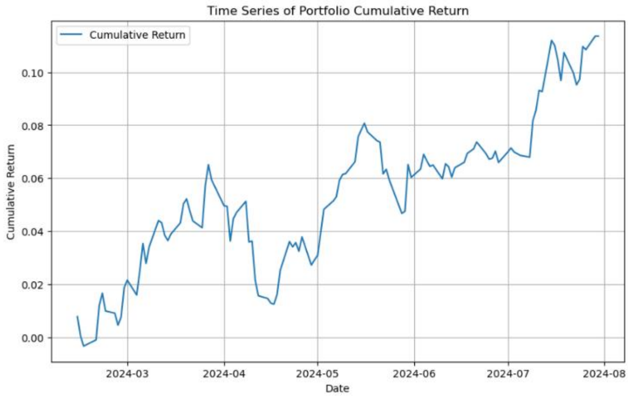

Fig. 8. Optimized Cumulative Return for GRU Model (Photo/Picture credit: Original).

4 Conclusion

This study compared five machine learning models—LightGBM, XGBoost, LSTM, GRU, and CNN—for predicting stock market trends in China. Initially, LightGBM, XGBoost, and GRU performed best based on annualized return, maximum drawdown, and Sharpe ratio. LightGBM had an annualized return of 16.44%, GRU 16.11%, and XGBoost 11.44%, while CNN struggled despite a low maximum drawdown. Bayesian optimization was then applied to fine-tune LightGBM, XGBoost, and GRU, leading to significant improvements. LightGBM's return increased to 20.97% with a Sharpe ratio of 1.68, XGBoost improved to 17.20%, and GRU achieved the highest return of 27.22%, though with greater risk exposure.

Despite the promising results, this study has some limitations. The dataset was restricted to the CSI 300 Index over a specific time period, which may limit the generalizability of the findings to other markets. Additionally, while Bayesian optimization improved the performance of gradient boosting models, the neural network models could benefit from further hyperparameter tuning and more advanced architectures like Transformers or hybrid models. Future research should explore applying these models to broader datasets, improving neural network techniques, and incorporating ensemble methods. Integrating real-world factors such as transaction costs and macroeconomic indicators would also enhance the practical applicability of these models in dynamic financial markets.

References

1. S&P Global Market Intelligence: Chinese IPOs plummet amid tough policy environment as focus shifts offshore. S&P Global (2024, July 9). https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/chinese-ipos-plummet-amid-tough-policy-environment-as-focus-shifts-offshore-82456751
2. Yang, Y., Wu, Y., Wang, P., Xu, J.: Stock price prediction based on XGBoost and LightGBM. In: E3S Web of Conferences, vol. 275, 01040 (2021). https://doi.org/10.1051/e3sconf/202127501040
3. Guennioui, O., Chiadmi, D., Amghar, M.: Improving global stock market prediction with XGBoost and LightGBM machine learning models. Review of Economics and Finance 21, 2603-2610 (2023).
4. Nti, I. K., Adekoya, A. F., Weyori, B. A.: A systematic review of fundamental and technical analysis of stock market predictions. Artificial Intelligence Review 53(4), 3007-3057 (2020).
5. Hatta, A. J., Dwiyanto, B. S.: The Company Fundamental Factors and Systematic Risk in Increasing Stock Price. Journal of Economics, Business, and Accountancy Ventura 15(2), 245-256 (2012).
6. Raju, V. G., Lakshmi, K. P., Jain, V. M., Kalidindi, A., Padma, V.: Study the influence of normalization/transformation process on the accuracy of supervised classification. In: 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), pp. 729-735. IEEE, (2020).
7. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., et al.: LightGBM: A highly efficient gradient boosting decision tree. In: Advances in Neural Information Processing Systems, vol. 30 (2017).
8. Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785-794. ACM, (2016).
9. Yu, Y., Si, X., Hu, C., Zhang, J.: A review of recurrent neural networks: LSTM cells and network architectures. Neural Computation 31(7), 1235–1270 (2019).
10. Dey, R., Salem, F. M.: Gate-variants of Gated Recurrent Unit (GRU) neural networks. In: 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 1597-1600. IEEE, (2017).
11. Alzubaidi, L., Zhang, J., Humaidi, A. J., et al.: Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data 8(53), (2021).
12. Frazier, P. I.: A tutorial on Bayesian optimization. arXiv preprint arXiv:1807.02811, (2018).