Forecasting Reliance Industries Stock Market Movements through Hybrid Feature Selection and Modelling

Smruti Rekha Das[1], Divya Midhun[2], Rajesh Dey[3]

[1,2]Lincoln University College, Malaysia

[3]Gopal Narayan Singh University, India

[1]pdf.smruti@lincoln.edu.my, [2]divya@lincoln.edu.my, [3]rajesh.dey@gnsu.ac.in

Abstract

Stock market prediction is still one of the most difficult tasks because of the dynamic, nonlinear nature of stock markets dependent on macroeconomic, environmental, social, and market sentiment factors. Conventional models tend not to apprehend the interactions between diverse variables, and hence, the predictions are less accurate. To tackle this challenge, we propose a hybrid prediction approach that merges statistical feature selection methods and optimization-based feature selection methods with progressive predictive models. This method uses financial data sets along with technical indicators and statistical parameters. Technical indicators and statistical parameters are calculated with a window size of 12. Key influencing parameters are highlighted before applying them to Long Short-Term Memory (LSTM) and Gradient Boosted Decision Trees (GBDT). The features are optimized using filter, Particle Swarm Optimization (PSO), and Differential Evolution (DE) methods. This research improves decision-making strategies for investors, financial analysts, and policymakers by offering an effective hybrid forecasting model.

Keywords: Stock Market Forecasting; Feature Selection; LSTM; GBDT; Optimization Algorithms

Introduction

The stock market is highly sensitive to macroeconomic trends, environmental and social governance practices (ESG), and investor sentiment. Accurately predicting changes is a historical problem in finance because the underlying factors are nonlinear and complicated. The conventional statistical models are highly dependent on historical price changes and tend to ignore broader market forces. This brings to light the requirement for a model that combines various disciplines, emphasizes significant factors, chooses the optimal features, and applies advanced machine learning models for better forecasting.

Related Work

Past exploration & research in stock market prediction has used a variety of methods, including statistical regression models, machine learning, and deep learning models. For example, earlier studies focused on autoregressive integrated moving average (ARIMA) models [1]. In contrast, more recent research has utilized ensemble models and deep neural networks [2]. Some approaches include

sentiment analysis to capture investor sentiment [3]. But these are lacking in the strength of feature selection or cannot effectively integrate statistical and metaheuristic optimization approaches. Table 1 illustrates the comparison of related works and the proposed approach.

Table 1. Comparison with related work

Reference	Statistical Feature Selection	Optimization Feature Selection	Deep Learning Integration	Sentiment Analysis
[1]	Yes	No	No	No
[2]	Yes	Yes	Partial	No
[3]	No	No	Yes	Yes
This Work	Yes	Yes	Yes	Yes

Key Contribution

The key contributions of this work are as follows:

1. A hybrid approach that integrates both optimization-based and statistical feature selection techniques.
2. Synthesis of technical indicators and statistical measures for comprehensive forecasting.
3. Comparative evaluation of LSTM and GBDT for short-term prediction (1 to 7 days ahead).
4. Demonstration of improved robustness and accuracy over existing models.

Dataset Description, Method, Experiments and Results

Data description

This study uses ten years (Jan 1, 2015 – Jan 1, 2025) of daily stock data for Reliance Industries Limited from the NSE, obtained via the yahoo finance Python library. Each record includes OHLCV attributes (Open, High, Low, Close, Volume). Out of the total sample (2467 rows) the dataset was split 70:30 into training (1713 samples) and testing (754 samples). Preprocessing and a 12-period rolling window were applied, generating 33 input features comprising statistical measures and technical indicators (e.g., moving averages, volatility, oscillators, momentum). For multi-horizon forecasting, the target variable (Open price) was shifted to 1, 3, 5, and 7 day horizons, creating multiple prediction tasks with a consistent structure. The final dataset integrates raw values, engineered indicators, and shifted outputs, providing a strong foundation for model evaluation. The statical indicators are Rolling Mean, Variance, Standard Deviation, Minimum, Maximum, Median, Skewness, Kurtosis and Correlation,

whereas the Technical Indicators are Simple Moving Average (SMA), Exponential Moving Average (EMA), Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD) with signal and histogram, Bollinger Bands, Stochastic Oscillator, Average True Range (ATR), On-Balance Volume (OBV), Rate of Change (ROC), Commodity Channel Index (CCI), Money Flow Index (MFI). Moreover, the data set for Reliance Industries is generated by incorporating technical indicators and statistical measures. Technical indicators are computed considering 12 as window size. The dataset undergoes two streams of feature selection:

1. **Statistical-based feature selection** – using Filter method.
2. **Optimization-based feature selection** – using metaheuristic algorithms such as PSO, DE.

The selected features are used to train predictive models such as **LSTM and GBDT**. Forecasting horizons are set between 1 and 7 days, and the outcomes are compared to assess predictive accuracy. The workflow diagram of the proposed hybrid framework is represented in Figure. 1.

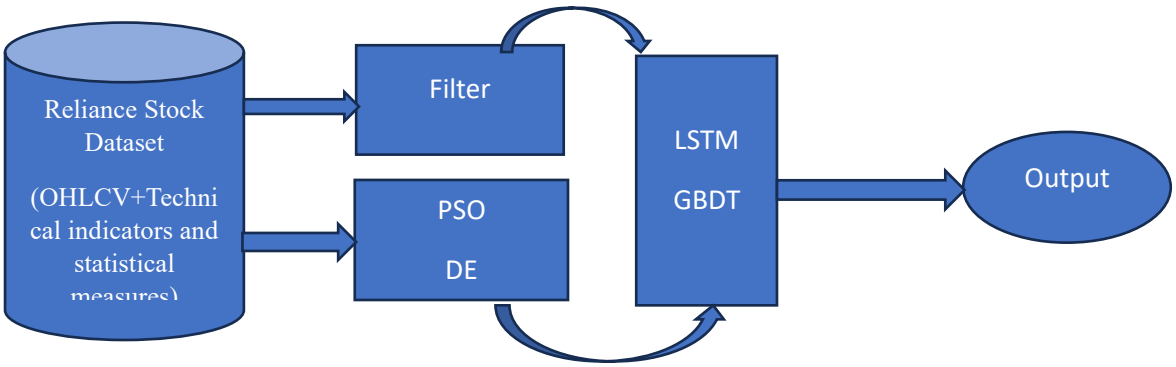

Figure 1. Proposed hybrid framework for stock market prediction

The experimental result given in Table 2. indicate that the hybrid approach significantly outperforms standalone feature selection techniques. For both LSTM and GBDT prediction model, filter method is showing better outcome for 1 day ahead prediction compared to other days ahead. PSO and DE used as feature selection along with optimization for LSTM and GBDT prediction model showed superior performance in case of one ahead prediction. Here both the prediction models effectively handled non-linear interactions.

Table-2 MSE result during Training for Reliance Industries Data Set

Days ahead to be predicted	LSTM_Filter method	GBDT_Filter method	LSTM_PSO method	GBDT_PSO method	LSTM_DE method	GBDT_DE method
1 day	**0.003878**	**0.067544**	0.908766	**0.088865**	**1.476538**	2.474636

3 days	0.007855	0.198038	**0.076533**	1.746542	4.579876	**1.454444**
5 days	0.013322	0.089776	1.879987	1.875654	4.983654	9.434777
7 days	0.233232	2.897654	3.849765	4.329767	9.47394	9.473479

Discussions

The findings highlight the importance of incorporating both **quantitative indicators** and **qualitative models** in market forecasting. Statistical methods provided baseline feature relevance, while optimization techniques refined the feature space to achieve higher accuracy. The synergy between LSTM and GBDT models showed the highest forecasting precision, particularly for horizons within week.

Conclusions

1. **Problem Statement Addressed**: Reliance Stock market forecasting suffers from high complexity due to diverse influencing factors.
2. **Method Used**: LSTM and GBDT prediction models reducing features by using Filter, PSO and DE methods.
3. **Key Findings**: Integration of technical indicators and statistical measures with available attributes of Reliance stock data deploying on hybrid feature selection significantly improves accuracy.
4. **Limitations and Future Work**: Future research should explore real-time data integration, adaptive learning for sudden market shocks, and broader global datasets.

References

1. X. Zhang, X. Zhang and L. Han, "An energy efficient Internet of Things network using restart artificial bee colony and wireless power transfer", *IEEE Access*, vol. 7, pp. 12686-12695, 2019. https://doi.org/10.1109/ACCESS.2019.2892798
2. X. Zhong, L. Zhang and Y. Wei, "Dynamic load-balancing vertical control for a large-scale software-defined Internet of Things", *IEEE Access*, vol. 7, pp. 140769-140780, 2019. https://doi.org/10.1109/ACCESS.2019.2943173
3. M. Malik, M. Dutta and J. Granjal, "A survey of key bootstrapping protocols based on public key cryptography in the Internet of Things", *IEEE Access*, vol. 7, pp. 27443-27464, 2019. https://doi.org/10.1109/ACCESS.2019.2900957