A Hybrid LSTM–CNN Framework for Robust Time-Series Forecasting and Trend Prediction

Danish Jamil[1]*, Aseef Ahmed[2]
[1]Department of Software Engineering, Sir Syed University of Engineering and Technology 75300, Karachi, Pakistan.
[2]Department of Data Science, NED University of Engineering and Technology 75300, Karachi, Pakistan.
*E-mail: djamil@ssuet.edu.pk

ABSTRACT

INTRODUCTION

Stock price forecasting remains one of the most challenging problems in financial analytics due to market volatility, nonlinear behavior, structural breaks, and long-range temporal dependencies. Traditional statistical and linear regression models often fail to capture these complexities, especially when applied across multiple stock tickers and changing market regimes. With the increasing availability of historical financial data and computational resources, machine learning (ML) and deep learning (DL) techniques have emerged as powerful alternatives for modeling complex time-series patterns. Despite this progress, many existing approaches focus primarily on numerical price prediction and neglect other critical aspects such as directional movement, anomaly detection, and real-world validation strategies. Furthermore, studies frequently concentrate on single stocks or short-term predictions, limiting the practical applicability of these methods in diverse and evolving markets. This study addresses these gaps by proposing an advanced, integrated machine learning framework for multi-stock price forecasting and trend direction prediction, aiming to improve prediction accuracy, robustness, and operational relevance.

OBJECTIVE

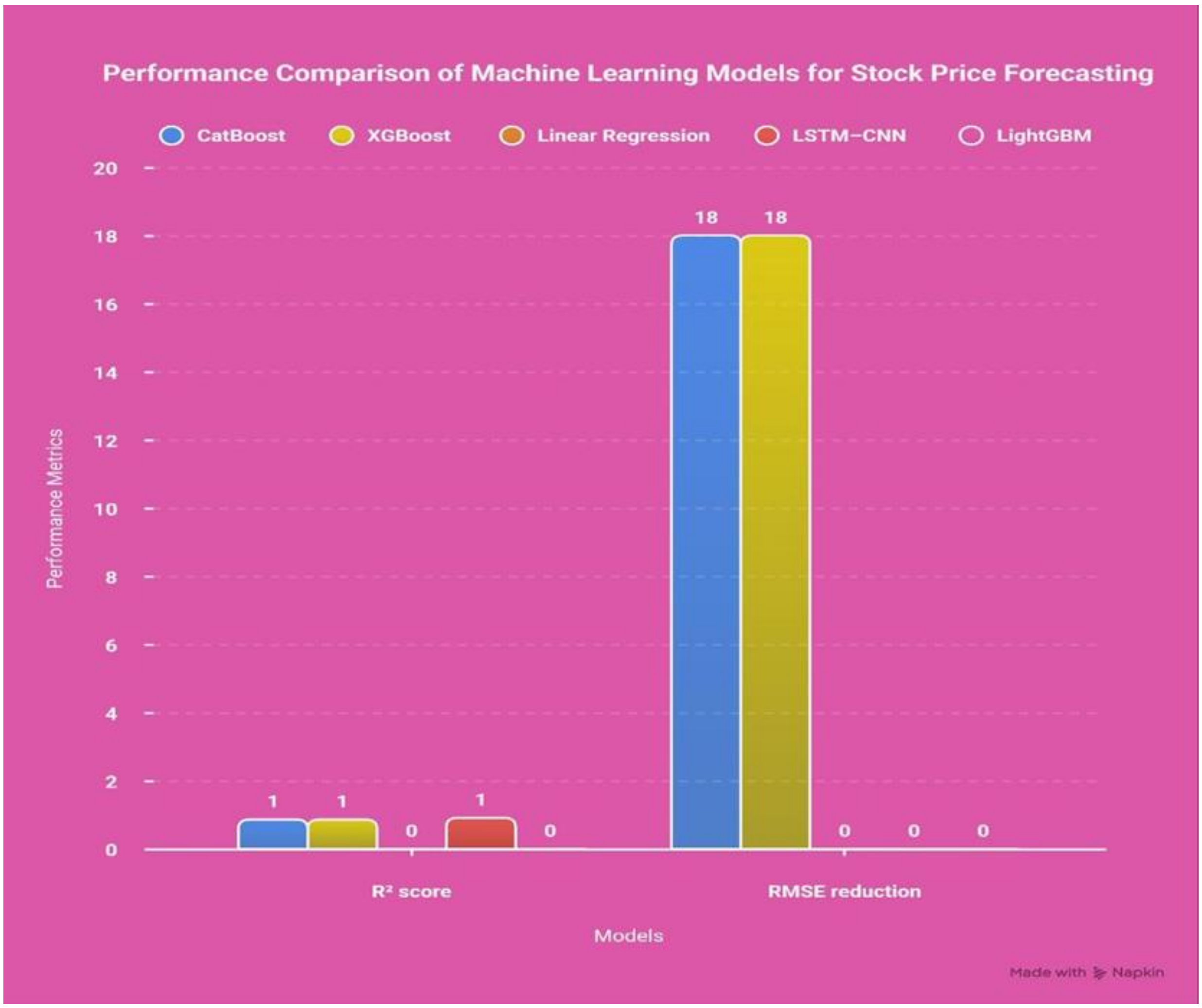
The primary objectives of this study are: (i) to develop a scalable and robust forecasting framework applicable to multiple stock tickers across different market regimes, (ii) to systematically compare the performance of classical machine learning models, boosting algorithms, and deep learning architectures, (iii) to predict both future stock prices and directional market movement (Up/Down), (iv) to enhance model reliability through anomaly detection techniques, and (v) to evaluate real-world performance using walk-forward validation rather than a conventional train–test split. By addressing these objectives, the study seeks to provide a practical decision-support framework for traders, portfolio managers, and financial analysts.

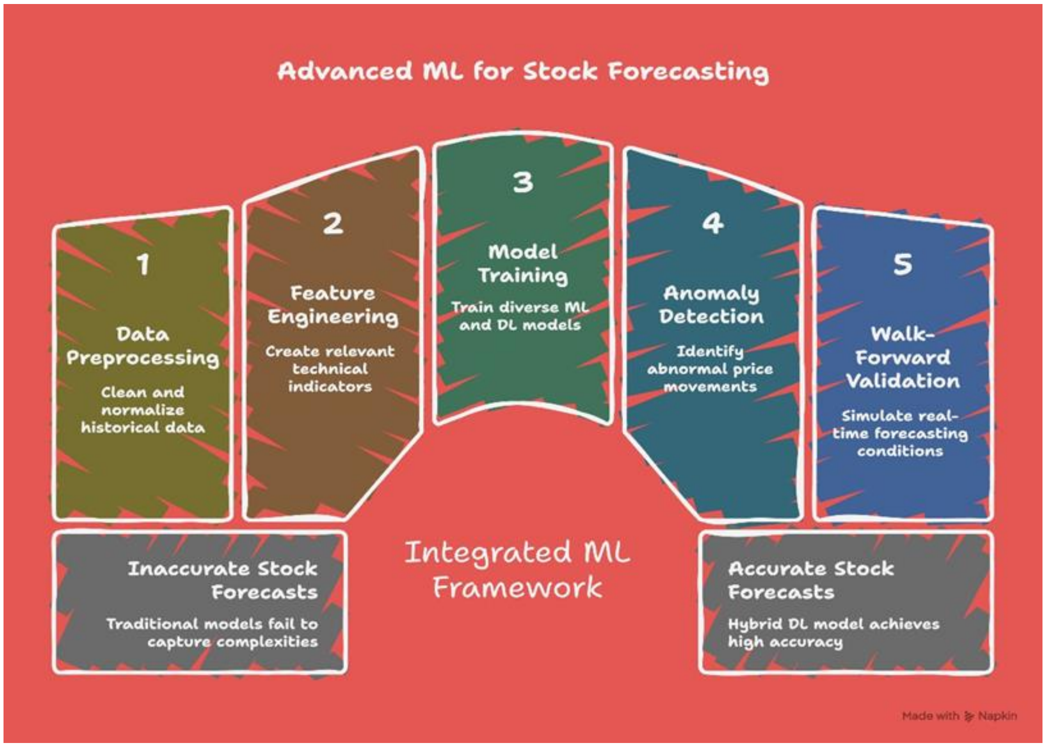
METHODOLOGY

Daily historical OHLCV (Open, High, Low, Close, Volume) data for multiple stock tickers were collected using the Python-based yfinance library. A comprehensive preprocessing pipeline was implemented, including date parsing, missing value imputation, normalization using MinMax scaling, and creation of lagged features to capture temporal dependencies. Feature engineering incorporated widely used technical indicators such as moving averages (MA5, MA10, MA20), Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), Bollinger Band width, lagged closing prices, and volume-based metrics. Several classical and ensemble models were trained and evaluated, including Ridge Regression, Random Forest, XGBoost, LightGBM, and CatBoost. In addition, a hybrid deep learning architecture combining Long Short-Term Memory (LSTM) networks and one-dimensional Convolutional Neural Networks (CNN) was implemented to capture both long-term dependencies and short-term temporal patterns. Model performance was evaluated using standard metrics including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R^2 score for regression tasks, along with directional accuracy and confusion matrices for classification tasks. An Isolation Forest algorithm was employed to detect anomalous price movements, while walk-forward validation simulated real-time forecasting conditions, ensuring that model performance could be generalized to evolving market scenarios.

KEY FINDINGS

Among classical machine learning models, boosting-based methods demonstrated superior performance. CatBoost and XGBoost achieved an average R^2 of approximately 0.86, with an RMSE reduction of nearly 18% compared to linear regression models. The hybrid LSTM–CNN model consistently outperformed all other approaches, achieving an R^2 of 0.91 and the lowest prediction error across multiple stock tickers. Directional movement prediction achieved 72–75% accuracy, with LightGBM exhibiting the strongest classification performance. Anomaly detection successfully identified abnormal price fluctuations during periods of high market volatility, while walk-forward validation confirmed stable performance under evolving market conditions. These findings highlight the importance of combining ensemble methods with deep learning architectures to improve predictive robustness and reliability.





CONCLUSION

The results demonstrate that hybrid deep learning architectures can significantly enhance short-term stock price forecasting, while boosting-based models provide efficient and reliable trend direction prediction. Careful feature engineering and the adoption of walk-forward validation are critical in improving model robustness and real-world applicability. The proposed framework offers a practical and extensible solution for multi-stock forecasting, decision-support systems, and algorithmic trading. Future work may incorporate additional data sources such as macroeconomic indicators, news sentiment analysis, and alternative data streams to further improve predictive performance and enable proactive risk management strategies.

Keywords: Stock Price Forecasting, Machine Learning, LSTM–CNN, Time Series Analysis, Trend Prediction.