Article

Not peer-reviewed version

Hybrid Machine Learning Models for Long-Term Stock Market Forecasting: Integrating Technical Indicators

Peujio Fozap Francis Magloire *

Posted Date: 10 February 2025

doi: 10.20944/preprints202502.0676.v1

Keywords: Stock Market Forecasting; Deep Learning Models; Hybrid LSTM-CNN; Technical Indicators; Financial Time-Series Prediction; Machine Learning in Finance; Risk management

doi:10.20944/preprints202502.0676.v1

Article

Hybrid Machine Learning Models for Long-Term Stock Market Forecasting: Integrating Technical Indicators

Francis Magloire Peujio Fozap

UNIVERSIDAD DE MONTERREY, MÉXICO; francis.peujio@udem.edu

Abstract: Stock market forecasting is a critical area in financial research, yet the inherent volatility and non-linearity of financial markets pose significant challenges for traditional predictive models. This study proposes a **hybrid deep learning model,** integrating **Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN)** with **technical indicators** to enhance the predictive accuracy of stock price movements. The model is evaluated using **Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R^2 score** on the **S&P 500 index** over a 14-year period. Results indicate that the LSTM-CNN hybrid model **achieves superior predictive performance compared to traditional models**, including Support Vector Machines (SVM), Random Forest (RF), and ARIMA, by effectively capturing both long-term trends and short-term fluctuations. While Random Forest demonstrated the highest raw accuracy with the lowest RMSE (0.0859) and highest R^2 (0.5655), it lacked sequential learning capabilities. The LSTM-CNN model, with an **RMSE of 0.1012, MAE of 0.0800, MAPE of 10.22%, and R^2 score of 0.4199**, proved to be **highly competitive and robust in financial time-series forecasting**. The study highlights the effectiveness of **hybrid deep learning architectures** in financial forecasting and suggests further enhancements through macroeconomic indicators, sentiment analysis, and reinforcement learning for dynamic market adaptation. It also improves risk-aware decision-making frameworks in volatile financial markets.

Keywords: stock market forecasting; deep learning models; hybrid LSTM-CNN; technical Indicators; financial time-series prediction; machine learning in finance; risk management

Jel Classification: C45; C53; C58; G14; G15; G17

---

1. Introduction

Accurately predicting stock market movements has long been a primary focus for financial analysts and investors due to financial markets' complex, non-linear dynamics. Traditional forecasting models, including Autoregressive Integrated Moving Averages (ARIMA) and exponential smoothing, have been widely used to model stock price behaviour. However, these models are limited in handling financial data's high complexity and volatility (Shah, Isah, & Zulkernine, 2019). Emerging economies like the BRICS nations present unique challenges in stock market forecasting due to high return volatility and spillover effects. Singh et al. (2024) analyze these market dynamics using advanced GARCH models, which account for structural breaks and non-linear dependencies in financial data. Their findings emphasize the need for machine learning approaches that integrate both historical price trends and external economic indicators to improve predictive accuracy. **Fischer & Krauss (2018)** Supported the importance of deep learning techniques in stock market forecasting. **Long, Lu, & Cui (2019)** Added more discussion on ML techniques applied to stock forecasting. In response to these limitations, Machine Learning (ML) and Deep Learning (DL) techniques have emerged as powerful alternatives, capable of capturing both short- and long-term dependencies in stock price data.

 doi:10.20944/preprints202502.0676.v1

Over the last decade, deep learning techniques such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN) have demonstrated remarkable effectiveness in stock market forecasting by identifying sequential and spatial patterns in stock price data. LSTM networks, in particular, excel at modelling long-term dependencies in time-series data, making them highly suitable for capturing historical trends that may influence future price movements (Najem, Bahnasse, &Talea 2024). Similarly, CNNs have been applied to analyze visual patterns in stock data, such as price charts, extracting meaningful insights from technical indicators (Hoseinzade & Haratizadeh, 2019). **Atsalakis and Valavanis (2009)** Discussed hybrid methodologies for financial forecasting. Although LSTM and CNN models perform well individually, hybrid approaches that combine these techniques have shown superior predictive accuracy by providing a more holistic perspective on stock market trends (Sangeetha & Alfia (2024).

One promising area in stock market prediction involves integrating technical indicators and mathematical calculations based on price, volume, or open interest as inputs to deep learning models. **Atsalakis and Valavanis (2009)** supported the argument on the effectiveness of technical indicators in ML models. **Sharma and Metha (2024)** discussed the impact of fundamental indicators on stock market forecasting. Technical indicators such as moving averages, Bollinger Bands, and price-to-earnings ratios capture essential market dynamics and can enhance forecasting performance when used with hybrid models. These indicators provide valuable insights into market conditions, contributing to more accurate predictions of stock price movements by reflecting underlying trends and market psychology.

Despite advancements in ML and DL, traditional forecasting models still face significant challenges when applied to financial data. Models like Support Vector Machines (SVM) and Random Forests effectively handle structured data but often struggle to capture the temporal dependencies and market volatility characteristic of long-term stock price movements (Shah, Isah, & Zulkernine, 2019). **Hossain et al. (2018)** discussed SVM and RF limitations. **Bao et al. (2017)** highlighted ARIMA's weaknesses in handling non-linearities. These limitations underscore the need for hybrid models that integrate LSTM and CNN with technical indicators, creating a more robust approach incorporating sequential and spatial features essential for comprehensive market analysis.

This study hypothesizes that hybrid models combining LSTM, CNN, and technical indicators outperform traditional ML models such as SVM and Random Forest in long-term stock market forecasting. Specifically, integrating temporal data (captured by LSTM) and spatial data (extracted by CNN) with technical indicators will lead to more accurate predictions of stock market trends and offer greater resilience to market volatility. The objectives of this research are to develop a hybrid machine learning model optimized for long-term stock market forecasting, compare its performance against traditional models using metrics such as Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and F1 score, and analyze the impact of technical indicators in enhancing predictive accuracy.

The remainder of this paper is structured as follows: Section 2 provides a comprehensive review of relevant literature, focusing on machine learning models, deep learning techniques, and the use of technical indicators in stock market forecasting. Section 3 outlines the methodology, detailing the data collection process, the architecture of the hybrid model, and the experimental setup. Section 4 presents the results, comparing the hybrid model's performance against traditional methods using standard financial metrics. Section 5 discusses the implications of the findings, examining the role of technical indicators in improving model accuracy and robustness in long-term forecasting. Finally, Section 6 concludes the study with a summary of key findings and recommendations for future research.

This research bridges the gap between traditional financial forecasting and advanced machine learning by leveraging technical indicators alongside hybrid architectures. It aims to contribute to risk-aware decision-making frameworks by demonstrating improved performance on long-term forecasts. This approach is aligned with the evolution of financial forecasting, wherein integrating multiple machine learning methods has proven to provide richer, more nuanced predictions. By

 doi:10.20944/preprints202502.0676.v1

addressing both spatial and temporal dependencies in market data, hybrid models like LSTM-CNN not only surpass traditional benchmarks but also bring scalability and adaptability to volatile market scenarios.

2. Literature Review

The complexity and volatility of financial markets have long driven researchers and practitioners to develop accurate methods for predicting stock price movements. Traditional statistical models like the ARIMA model have been central in time-series forecasting. However, these models struggle to capture financial data's intricate, non-linear dependencies. Consequently, ML and DL techniques have emerged as advanced alternatives, identifying short- and long-term patterns in stock market behaviour. Market uncertainty, particularly due to macroeconomic factors, has been a key challenge in financial forecasting. Kwon (2025) explores how global economic volatility influences corporate bond markets, providing insights into how financial models must adapt to external shocks. These findings highlight the necessity of incorporating risk-aware forecasting models that can dynamically adjust predictions in response to changing economic conditions.

2.1. Machine Learning and Deep Learning in Stock Market Forecasting

Early approaches to stock market forecasting relied heavily on time-series models such as ARIMA and exponential smoothing. While effective for fundamental trend analysis, Shah, Isah, and Zulkernine (2019) observed that these methods are limited in handling complex, non-linear relationships present in financial data. Machine learning methods have since evolved, enabling more flexible models that handle larger data volumes and capture underlying patterns in stock prices that traditional models frequently miss. Leippold et al. (2022) demonstrated the importance of factor investing in stock market prediction, showing that liquidity and investor sentiment significantly impact short-term returns in the Chinese stock market. Their findings underscore the importance of factor selection when applying machine learning techniques to emerging markets.

Traditional ML models, including SVM and Random Forests, have been applied successfully to forecast stock prices by modelling complex relationships between variables and stock price movements. However, these models often fail to capture stock price's sequential, time-dependent nature, especially for long-term forecasting. This limitation has driven the development and adoption of deep learning models, particularly Recurrent Neural Networks (RNNs) and LSTM networks, better suited for sequential data.

LSTM networks have become a prominent choice in financial forecasting due to their unique ability to model long-term dependencies within sequential data. As Najem et al. (2024) explain, LSTMs can capture temporal patterns in stock prices that traditional models may overlook, leading to improved predictive accuracy. LSTMs leverage historical trends and market behaviour to retain information across extended sequences and enhance long-term forecasts.

In practical applications, LSTM models have shown significant improvements in daily stock price prediction compared to older models like ARIMA (Najem et al., 2024, or Hoseinzade & Haratizadeh, 2019), mainly due to their robustness in handling stock price volatility and non-linearity characteristics. LSTM models are especially adept at identifying long-term trends, making them highly valuable for stock market forecasting, where historical data patterns significantly impact future trends.

While LSTMs are effective for time-series data, CNNs offer complementary advantages by processing visual data representations, such as price charts. Najem et al. (2024) demonstrated that CNNs can extract spatial features from stock price graphs, revealing patterns that are difficult to detect through numerical data alone. This is particularly valuable for identifying short-term trends, as CNNs can mimic human technical analysis by recognizing patterns in candlestick charts and various technical indicators.

In hybrid models, CNNs are often paired with LSTMs, enabling the capturing of both spatial and temporal features. Fischer and Krauss (2018). This combination allows the model to analyze price data sequentially (through LSTM) and visually (through CNN), often resulting in improved forecasting accuracy. Such hybrid architectures are particularly effective for financial markets, where both types of patterns play significant roles.

2.2. The Role of Technical Indicators in Stock Market Prediction

Technical indicators have become essential tools in financial forecasting, providing quantitative measures of market trends, volatility, and trading volume. These indicators help analysts capture both the direction and strength of stock price movements and have been extensively incorporated into machine learning models as features that reflect historical price behaviour.

Moving Averages, including Simple Moving Averages (SMA) and Exponential Moving Averages (EMA), are commonly used in forecasting models to smooth out price data, making it easier to identify general trends over specific periods. SMAs effectively eliminate noise, while EMAs give greater weight to more recent data points, closely reflecting current market sentiment. As explored in Hoseinzade and Haratizadeh (2019), Bollinger Bands measure price volatility relative to moving averages, offering early signals for overbought or oversold conditions. These signals align well with predictive frameworks designed to capture abrupt price movements.

The P/E ratio has long been a cornerstone in financial modelling. According to Zanc et al. (2019), its inclusion enriches models by quantifying valuation anomalies often preceding trend reversals. Models incorporating the P/E ratio as a feature benefit from a more comprehensive view of stock valuation trends, particularly in identifying periods of high or low market confidence. As a trading activity gauge, volume is another critical indicator that signifies liquidity levels and the strength of price movements, thus aiding the model in discerning actual trend reversals from temporary fluctuations.

By integrating technical indicators like moving averages, Bollinger Bands, P/E ratio, and volume, models can improve their accuracy and reliability in forecasting stock prices. AI, School, Ji, Song, Zhong, Jia & Tianhao, (2025) also explored the integration of multiple technical indicators and highlighted the value of feature diversity in capturing complementary aspects of stock behaviour. Studies have shown that using technical indicators as input features enhances the predictive performance of ML models by providing additional context to price movements, helping models better differentiate between transient and sustained trends.

2.3. The Shift to Hybrid Models

Combining LSTM and CNN architectures, hybrid models significantly advance stock market prediction by integrating temporal and spatial data features. Fischer and Krauss (2018). This approach leverages the strengths of both deep learning models: LSTM's ability to retain long-term dependencies and CNN's capacity to process complex visual patterns.

Research by Sangeetha and Alfia (2024), and Najem et al. (2024) demonstrates that hybrid models benefit from CNNs' capacity to process technical indicators and stock price charts while LSTMs manage the sequential aspect of stock price data. By combining these models, hybrid architectures create a more comprehensive predictive framework, capturing both short-term market fluctuations and long-term trends. This hybrid approach is particularly valuable in financial forecasting, where price history and chart visual patterns impact investor decision-making.

Hybrid models offer superior feature extraction capabilities by integrating image-based and time-series analyses. This dual approach is advantageous in financial applications where the data reflects quantitative trends and visual chart patterns. By fusing CNN and LSTM outputs, hybrid models capture a holistic view of market dynamics that single-model approaches lack, enabling improved performance in forecasting tasks that involve high-dimensional, complex data.

Traditional machine learning models, such as SVM and Random Forest (RF), remain popular in stock market prediction due to their robustness in handling high-dimensional datasets. SVMs are highly effective in binary classification tasks and are common in stock price movement predictions.

 doi:10.20944/preprints202502.0676.v1

At the same time, Random Forest, with its ensemble of decision trees, provides reliable and stable forecasts by averaging multiple models.

However, as noted in comparative studies, these traditional models, while reliable for short-term forecasting, often underperform when used for long-term predictions. This limitation arises from their difficulty in capturing sequential dependencies in stock price data—a strength of LSTM-based models. Hybrid models that combine CNN and LSTM have been shown to outperform traditional models like SVM and Random Forest across key performance metrics, especially in tasks that require both short-term and long-term forecasting.

3. Methodology

This section provides a comprehensive outline of the methodology used in developing and evaluating a hybrid LSTM-CNN model for forecasting stock prices of the S&P 500 index. The methodological approach includes data collection, model architecture, experimental setup, and baseline comparisons, focusing on leveraging technical indicators to model stock trends accurately.

3.1. Data Collection

For this study, daily stock price data and technical indicators are sourced from Yahoo Finance, focusing on the S&P 500 index for 14 years, from January 2010 to September 2024. Collecting a 14-year dataset provides a robust temporal baseline, enabling the model to account for multiple market cycles, as Sangeetha and Alfia (2024) suggested. Longer historical spans improve the model's generalization ability across varied economic conditions. The dataset consists of historical stock prices and technical indicators, each critical in capturing stock price behaviour patterns. This dataset provides the basis for training and evaluating the hybrid model, designed to capture both long-term trends and short-term fluctuations in the stock market.

The key components of the dataset include:

a. **Daily Stock Prices: Daily** stock price data includes each trading day's open, close, high, and low prices. These values capture fundamental price movement over time, providing a base for time-series analysis. By observing daily price fluctuations, the model can establish sequential patterns that contribute to the overall trend direction of the stock index.

b. **Technical Indicators**: Several technical indicators are computed based on historical stock prices to enhance the model's predictive capabilities further. Each indicator brings a unique perspective, offering additional insights into market dynamics that daily prices alone may not reveal:

c. **Moving Averages (SMA and EMA): Simple Moving Average (SMA)** and **Exponential Moving Average (EMA),** particularly in the 10-day and 50-day configurations, are commonly adopted as baseline measures in financial forecasting. Their inclusion aligns with prior studies, such as Hoseinzade and Haratizadeh (2019), where they effectively identified momentum trends. SMAs are straightforward averages of stock prices over a specific period, smoothing out fluctuations and providing a view of long-term price direction. EMAs, in contrast, apply more significant weight to more recent prices, making them more responsive to recent market changes. The inclusion of both 10-day and 50-day periods allows the model to capture short-term and long-term trends, which helps in understanding whether current price changes align with or deviate from these trends.

$$SMA_n = \frac{1}{n}\sum_{i=0}^{n-1} \mathrm{P}_i \tag{1}$$

where P_i represents the closing price over n days. SMA smooths fluctuations, highlighting overall trends.

$$EMA_t = \mathrm{P}_t \cdot \alpha + EMA_{t-1} \cdot (1-\alpha),\ where\ \alpha = \frac{2}{n+1} \tag{2}$$

EMA assigns greater weight to recent prices, making it more responsive to new information.

a- **Bollinger Bands:** Bollinger **Bands** are used to assess stock prices' volatility. They consist of a **moving average** (often a 20-day SMA) surrounded by an **upper** and **lower band** based on standard deviations. This indicator effectively identifies potential overbought or oversold market conditions, as stock prices tend to revert to the mean after reaching these extremes. By integrating Bollinger Bands, the model can better account for volatility patterns, an essential aspect of predicting market reversals and short-term fluctuations.

$$BB_{upper} = SMA + (k \cdot \sigma) \quad (3)$$

$$BB_{lower} = SMA - (k \cdot \sigma) \quad (4)$$

b- **Price-to-Earnings (P/E) Ratio**: The P/E ratio, calculated by dividing the stock's market price by its earnings per share (EPS), is a valuation metric. This ratio provides insight into **market** sentiment regarding the stock's value relative to its earnings. In the context of the S&P 500, the P/E ratio is an essential indicator, as it reflects investor confidence and overall economic health. A higher P/E ratio often suggests future growth expectations, while a lower ratio may indicate a valued or undervalued market, affecting the model's forecasting.

c- **Volume**: Trading volume represents the total number of shares traded within a given period and is a direct measure of market participation and sentiment. Volume helps the model gauge **the** intensity of price movements, high volume during a price rise may indicate strong buying interest. In contrast, low volume might signal a lack of investor confidence. Volume trends also indicate shifts in market behaviour, such as accumulation or distribution phases, and provide context to price changes, enhancing the model's ability to interpret the strength of market signals. Volume, being a proxy for market liquidity, not only validates price movements but also highlights potential market anomalies, particularly during high-volatility trading hours. This is especially crucial when analyzing institutional buying or selling patterns, as explored in Hoseinzade and Haratizadeh (2019).

d- Technical indicators are fundamental in financial analysis, providing insights into trends, volatility, and potential price reversals. This study incorporates five key indicators: Moving Averages (SMA and EMA), Bollinger Bands, Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and On-Balance Volume (OBV). These **indicators** serve as input features for our hybrid LSTM-CNN model, enabling it to capture underlying stock price trends and market momentum (Hoseinzade & Haratizadeh, 2019)

$$RSI = 100 - \left(\frac{100}{1+RS}\right), \quad RS = \frac{Avg\ Gain\ (n)}{Avg\ Loss\ (n)} \quad (5)$$

RSI assesses price momentum by comparing average gains to average losses over nnn periods.

$$MACD = EMA_{short} - EMA_{long} \quad (6)$$

$$Signal\ Line = EMA_{MACD} \quad (7)$$

MACD identifies trend reversals by comparing short-and long-term moving averages.

$$OBV_t = OBV_{t-1} + \begin{cases} V_t, & P_t > P_{t-1} \\ -V_t, & P_t < P_{t-1} \\ 0, & P_t = P_{t-1} \end{cases} \quad (8)$$

OBV measures volume flow, determining if price movements are backed by strong trading activity.

These components provide a robust dataset that encapsulates price behaviour and trading activity, forming a quantitative foundation for accurate stock trend prediction. Focusing exclusively on quantitative data, this approach sidesteps reliance on sentiment analysis, emphasizing the model's ability to identify and analyze price patterns, technical signals, and market dynamics. Combining price data and technical indicators creates a comprehensive, multidimensional dataset that supports a predictive framework to understand complex market behaviour over short- and long-term horizons.

3.2. Model Architecture

The architecture of the hybrid model is specifically designed to exploit the strengths of both LSTM networks and CNN. LeCun et al. (2015). Together, these two powerful deep-learning models address the unique characteristics of stock price data. The LSTM component captures sequential patterns in time-series data, while the CNN component analyzes technical indicators to uncover meaningful short-term patterns. Combining these approaches allows the hybrid model to predict

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 February 2025 doi:10.20944/preprints202502.0676.v1

stock prices with greater accuracy and robustness. Recent studies have confirmed that hybrid models leveraging LSTM for temporal pattern recognition and CNN for extracting local dependencies significantly outperform traditional statistical approaches in financial time-series forecasting (Najem et al., 2024; Sangeetha & Alfia, 2024).

a- LSTM Layer: The LSTM layer forms the backbone of the model for handling time-series data. LSTM networks are designed to capture temporal dependencies within sequential data, making them well-suited for stock market forecasting, where historical price patterns influence trends. Stock prices, by nature, often exhibit correlations over time, where prior values and trends impact current prices. Numerous studies have demonstrated that LSTMs excel in financial forecasting due to their ability to mitigate vanishing gradient problems, thereby retaining long-term dependencies critical in predicting stock price trends (Leippold et al., 2022). The LSTM layer in this model processes historical stock prices along with technical indicators, learning long-term dependencies that can significantly affect future price movements.

b- Through gated memory cells, LSTM networks retain relevant information across multiple time steps, which is crucial for modelling dependencies in financial data. This capability enables the model to understand the progression of trends, volatility cycles, and other sequential patterns in the stock's price history. Specifically, the LSTM layer interprets sequences of data points (e.g., daily prices and indicators over a set lookback period) to predict future prices by identifying recurring patterns and long-term dependencies in the input data.

c- CNN Layer: The CNN layer complements the LSTM by focusing on spatial analysis, particularly the short-term patterns and trends in stock price data and technical indicators. CNNs excel at identifying localized patterns within data by applying convolutional filters, which extract relevant features by scanning through the dataset with multiple filters. The application of CNNs in financial modeling has gained prominence, with recent research highlighting their ability to extract meaningful short-term patterns from price fluctuations and technical indicators, particularly in volatile market conditions (Hao et al. 2023). In this model, the CNN layer processes price and technical indicator data (formatted as a 2D representation) to detect short-term trends and changes in volatility, such as peaks, troughs, and rapid shifts in trading volume. By applying convolutional operations, the CNN layer identifies structural patterns that traditional time-series models may overlook. For instance, it can recognize rapid uptrends or downtrends within shorter intervals, detect overbought or oversold conditions from volatility indicators like Bollinger Bands, and capture localized shifts in moving averages, often precursors to broader trend changes. Furthermore, the CNN layer's capability to detect such technical signals gives the model detailed insights into market conditions, allowing it to respond dynamically to changing market states.

d- Model Fusion and Prediction Layer: After processing the data through the LSTM and CNN layers, the hybrid model combines the outputs of these two components in a fully connected layer. This layer acts as the fusion point where the temporal insights from the LSTM and the spatial features from the CNN are integrated to produce a final, comprehensive output. By merging these features, the fully connected layer synthesizes long-term trends (captured by the LSTM) and short-term market dynamics (identified by CNN) to create a well-rounded prediction of future stock price movement. Hybridizing deep learning models with ensemble techniques, such as Random Forest or XGBoost, has been shown to further enhance model stability by incorporating diverse feature selection strategies (Kaur et al., 2024). The fusion approach allows the model to leverage macro and micro market indicators, leading to a prediction that reflects historical trends and technical shifts within shorter intervals. This architecture enables the model to generate robust and adaptive predictions, which are particularly advantageous in the volatile environment of stock trading, where both long-term trends and short-term price fluctuations contribute to price movements.

The LSTM-CNN hybrid architecture combines sequential analysis with pattern recognition, providing a multi-faceted predictive framework. The LSTM layer processes time-dependent price and indicator data, learning from long-term dependencies, while the CNN layer identifies critical short-term patterns. Together, these elements enable the hybrid model to make predictions based on an enriched representation of the stock's historical and technical data, offering greater accuracy and resilience in the face of market fluctuations. Given the hybrid model's complexity, the evaluation metrics focus on both predictive accuracy and risk-adjusted returns, ensuring the model is reliable and practical for financial applications.

3.3. Experimental Setup

The experimental setup involves a structured approach to data processing, model training, and validation to ensure the hybrid model's robustness and reliability. Key aspects of the experimental design include data preparation, model training with cross-validation, and the use of multiple evaluation metrics to assess predictive accuracy and financial viability comprehensively.

a- Training and testing split: To ensure a balanced model training and evaluation approach, the dataset is divided into an **80/20 split**, following best practices in stock market forecasting (Sangeetha & Alfia, 2024). The **training set** consists of approximately **80% of historical stock data (January 1, 2010 – December 31, 2019)**, allowing the model to capture **long-term market trends and seasonal patterns** (Najem et al., 2024). The **remaining 20% (January 1, 2020 – December 31, 2024)** serves as a **test set**, ensuring that the model is validated against **recent, unseen financial conditions**, making it more adaptable to future market movements (Hoseinzade & Haratizadeh, 2019). This **strategic partitioning** aligns with prior studies that emphasize the necessity of training deep learning models on **extensive historical data** while reserving recent periods for evaluation, preventing **overfitting and ensuring real-world applicability**.

b- Cross-Validation: Cross-validation is incorporated within the training phase to prevent overfitting and enhance model robustness (Sangeetha & Alfia, 2024). Training the model on multiple data folds allows for a more comprehensive performance assessment, ensuring that predictions remain reliable across different market conditions (Najem et al., 2024). This technique is particularly valuable in financial forecasting, where high variability and complex patterns are present in historical data (Hoseinzade & Haratizadeh, 2019).

c- Evaluation Metrics: To assess the predictive performance of the hybrid LSTM-CNN model, we employ multiple quantitative metrics widely used in stock market forecasting (Najem et al., 2024). These metrics provide a holistic evaluation of both error minimization and model explanatory power. Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) serve as primary measures of prediction accuracy, ensuring precise error evaluation (Najem et al., 2024), while the R^2 score assesses how well the model explains stock price variations (Hoseinzade & Haratizadeh, 2019). The Mean Absolute Percentage Error (MAPE) is included to measure relative prediction errors, which is particularly useful for stock prices that fluctuate across different price levels (Sangeetha & Alfia, 2024). Although financial performance metrics such as the Sharpe Ratio and Max Drawdown are commonly used in portfolio risk assessment, our study focuses primarily on price prediction rather than trading strategies. As a result, these metrics were not included in our evaluation framework, aligning with prior studies emphasizing RMSE, MAE, and R^2 in financial time-series forecasting (Hoseinzade & Haratizadeh, 2019).

3.4. Baseline Models for Comparison

To rigorously assess the performance of the LSTM-CNN hybrid model, we compare it against three widely recognized baseline models in financial forecasting. These models include traditional machine learning algorithms and statistical techniques, each offering unique capabilities in analyzing stock price data. The selected baseline models are SVM, RF, and ARIMA. These baseline models represent a spectrum of approaches in financial forecasting, from ARIMA's foundational statistical methods to Random Forest's ensemble-based approach. SVM, while widely used for its classification capabilities, has often been challenged by the inability to model temporal dynamics (Shah et al., 2019).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 February 2025 doi:10.20944/preprints202502.0676.v1

By contrasting these models against a hybrid LSTM-CNN, we aim to demonstrate the evolutionary leap hybrid architectures offer in market forecasting.

Table 1. Baseline models for comparison.

	SVM	RF	ARIMA
Purpose	SVM is a classification model widely used in financial forecasting for its ability to analyze the relationship between input features (such as historical stock prices and technical indicators) and stock price movements.	Random Forest is an ensemble model combining multiple decision trees to improve prediction stability and accuracy, particularly for regression task stock price forecasting.	ARIMA is a traditional time-series forecasting model extensively applied in financial data analysis. It is instrumental in identifying patterns within historical data and projecting future price trends.
Application	This study uses SVM as a binary classifier to predict the direction of stock price changes (e.g., price increase or decrease) rather than specific price values.	In this context, RF is used as a regression model to predict stock prices based on historical price data and technical indicators. It leverages the aggregated outputs from multiple decision trees to yield stable predictions.	ARIMA generates forecasts based solely on historical stock price data as a benchmark model without incorporating additional technical indicators. This allows for a direct comparison between the predictive accuracy of traditional statistical methods and more complex machine learning models.
Strengths	SVM effectively distinguishes between categories (e.g., upward or downward price movement), particularly in complex, high-dimensional data cases. Its strength lies in maximizing the margin between categories, which helps to enhance generalization and reduce the risk of overfitting.	The RF model is known for its robustness in handling high-dimensional datasets and reducing the variance of individual decision trees. It provides reliable results in stock forecasting due to its ability to mitigate the effects of outliers and noisy data, making it a suitable benchmark for evaluating the performance of the hybrid model.	ARIMA is well-regarded for its simplicity and interpretability in financial forecasting. While it cannot capture non-linear dependencies and interactions, it remains a reliable baseline to gauge the added value provided by advanced hybrid models.

Source: Own elaboration.

These baseline models (SVM, Random Forest, and ARIMA) serve as benchmarks for evaluating the performance of the LSTM-CNN hybrid model. By comparing the hybrid model's results with these established techniques, we aim to demonstrate the enhanced predictive accuracy and robustness achieved by integrating LSTM and CNN architectures. The comparative analysis will focus on performance metrics such as RMSE, MAE, and R^2, comprehensively assessing the model's effectiveness in long-term stock market forecasting.

3.5. Baseline Models for Comparison: Model Specifications

a) Support Vector Machine (SVM)

SVM aims to maximize the margin M between classes:

$$min\frac{1}{2}||w||^2 + C\sum_{i=1}^{n}\xi_i \quad (9)[1]$$

subject to:

$y_i(w \cdot x_i + b \geq 1 - \xi_i$ and $\xi_i \geq 0$

where,

w: Weight vector

b: Bias term

C: Regularization parameter controlling margin width

ξ_i : Slack variables to allow some misclassification

y_i : Actual class label (1 or -1 for binary classification)

b) Random Forest (RF)

RF aggregates predictions from multiple decision trees T_m for regression:

$$\hat{y} = \frac{1}{M}\sum_{m=1}^{M} \mathrm{T}_m(X) \quad (10)[2]$$

Where,

M: Number of trees in the forest

$\mathrm{T}_m(X)$: Prediction from the $m-th$ tree based on input X

X: Input features (e.g., historical prices, technical indicators)

c) Autoregressive Integrated Moving Average (ARIMA)

ARIMA (p, d, q) uses autoregressive (AR) and moving average (MA) components:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \quad (11)[3]$$

Where,

y_t: Stock price at time t

c: Constant term

ϕ_i: AR coefficients for lag i

ϕ_i: MA coefficients for lag i

ε_t: White noise error term

This section would enhance the technical rigour by explicitly detailing how each model processes inputs and generates predictions.

3.6. Model Architecture: LSTM-CNN Hybrid Model Specifications

The fusion of temporal and spatial data is a critical strength of the LSTM-CNN hybrid model. Unlike standalone architectures, this hybrid design ensures that both long-term dependencies in price movements and localized short-term signals are simultaneously captured. Such a dual-focus approach not only enhances predictive robustness but also reflects real-world trading environments where both macroeconomic trends and micro-level fluctuations drive market behavior.

a) Long Short-Term Memory (LSTM) Layer

---

[1] Equation (1) aligns with the approach used in Zanc et al. (2019) to optimize binary classification tasks, focusing on maximizing margin separation in high-dimensional datasets.

[2] Equation (2) reflects the ensemble learning principles established in AI et al. (2025), emphasizing RF's ability to handle high variance in feature spaces.

[3] Although ARIMA lacks the complexity of ML-based models, its inclusion as a baseline highlights the relative benefits of deep learning techniques for long-term forecasts (Shah, Isah, & Zulkernine, 2019).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 February 2025 doi:10.20944/preprints202502.0676.v1

The LSTM layer is designed to capture temporal dependencies through gates that control information flow:

Forget Gate: $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ (12)[4]

Input Gate: $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$ (13)

$\widetilde{C_t} = tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$ (14)

Cell State Update: $C_t = f_t \cdot C_{t-1} + i_t * \widetilde{C_t}$ (15)[5]

Output Gate: $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_0)$ (16)

$h_t = o_t \cdot \tanh(C_t)$ (17)

Where,

x_t: Input features (e.g., stock price and technical indicators) at time t

h_t: Hidden state at time t

C_t: Cell state at time t

W_f, W_i, W_c, W_o : Weight matrices for respective gates

b_f, b_i, b_c, b_0: Bias terms for respective gates

σ: Sigmoid activation function

b) Convolutional Neural Network (CNN) Layer

The CNN layer applies convolutional filters to extract spatial features from the input data, primarily technical indicators:

$Z = \sigma(\sum_{k=1}^{K} \mathrm{W}_k * X + \mathrm{b}_k)$ (18)[6]

Where,

W_k: Weight matrix of the $k-th$ filter

b_k: Bias for the $k-th$ filter

* Convolution operator

Z: Output feature map

σ: Activation function (e.g., ReLU)

By synthesizing outputs from LSTM and CNN, the fully connected layer serves as the integrative engine that converts diverse patterns into actionable price predictions. This layer finalizes the prediction by balancing the weight of temporal and spatial inputs, ensuring a comprehensive representation of stock dynamics.

c) Model Fusion and Prediction Layer

After the LSTM and CNN layers process the temporal and spatial features, respectively, their outputs are fused in a fully connected layer to generate the final prediction:

$y = \sigma(W \cdot [h_t, Z] + b)$ (19)[7]

Where,

y: Final predicted stock price

h_t: Final hidden state from the LSTM

Z: Output from the CNN layer

W: Weight matrix for the fully connected layer

b: Bias term

---

[4] The forget gate selectively removes irrelevant information from previous sequences, a critical mechanism in time-series forecasting. As Najem et al. (2024) emphasized, this ensures that only historically significant data influences predictions.

[5] The updated cell state integrates prior information with new input, allowing the LSTM to capture patterns spanning extended periods, making it ideal for financial time-series forecasting (Shah et al. 2019).

[6] CNNs apply convolutional filters to capture localized trends in technical indicators. This aligns with findings by Hoseinzade and Haratizadeh (2019), who demonstrated that CNNs effectively detect chart patterns like peaks and troughs.

[7] The fusion of LSTM and CNN outputs into a fully connected layer synthesizes temporal and spatial insights, a technique validated by Sangeetha & Alfia (2024) in their comparative studies of hybrid architectures.

 doi:10.20944/preprints202502.0676.v1

σ: Activation function (e.g., linear or ReLU)

The hybrid LSTM-CNN model is designed to capture both long-term sequential dependencies and short-term market fluctuations through the structured integration of historical stock prices and technical indicators. The model architecture leverages LSTM layers to extract temporal patterns, while CNN layers analyze spatial features from technical indicators, ensuring a robust predictive framework. Additionally, hyperparameter tuning and cross-validation strategies were employed to optimize model performance and mitigate overfitting risks. These methodological choices provide a strong foundation for accurately forecasting stock price trends in volatile financial markets.

In the following section, we present the results of our experiments, evaluating the hybrid model's predictive accuracy using key metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R^2 Score. We further analyze the loss curves, MAE trends, and actual vs. predicted stock price plots to assess model stability and generalization. Additionally, we compare the hybrid model's performance against traditional machine learning benchmarks, including Support Vector Machines (SVM), Random Forest (RF), and ARIMA, to highlight the advantages of deep learning-based approaches in long-term stock market forecasting.

4. Results and Discussions

To ensure reliable performance, comprehensive data preprocessing was conducted before training the hybrid LSTM-CNN model on stock price and technical indicator data. These preprocessing steps included handling missing values, applying MinMax scaling, structuring input sequences for time-series learning, and tuning hyperparameters to enhance predictive accuracy. Since financial data often contain missing values due to market closures, reporting inconsistencies, or corporate events, we applied **forward-fill imputation** to maintain continuity in historical price trends, as recommended in prior studies (Sangeetha & Alfia, 2024). This approach ensures that the model does not misinterpret missing values as extreme price drops or sudden spikes, which could distort trend analysis.

Furthermore, to standardize input features and improve model stability, all stock price values and technical indicators were **normalized using MinMax scaling,** transforming them into a range between **0 and 1** (Najem et al., 2024). Normalization is particularly crucial in financial forecasting, as it prevents high-magnitude values (such as stock prices) from disproportionately influencing model learning while ensuring that all indicators contribute equally. Additionally, data sequences were structured using **lookback periods** of **10, 30, and 50 days,** allowing the LSTM component to effectively capture short-term fluctuations and long-term trends. The choice of a **30-day lookback period** yielded the best balance between capturing meaningful historical dependencies and avoiding unnecessary complexity, aligning with findings from Hoseinzade and Haratizadeh (2019).

4.1. Model Performance Analysis

Figure 1 presents the loss curve to analyze model learning behaviour**,** depicting the evolution of training and validation loss over **150 epochs.** The **training loss** consistently decreases, indicating effective model learning, while the **validation loss** exhibits minor fluctuations, reflecting financial market volatility. However, the absence of a significant upward trend in validation loss suggests that the model is not overfitting. This stability was achieved through **dropout regularization** and **batch normalization,** which prevent the model from memorizing noise instead of identifying meaningful patterns in stock price movements.

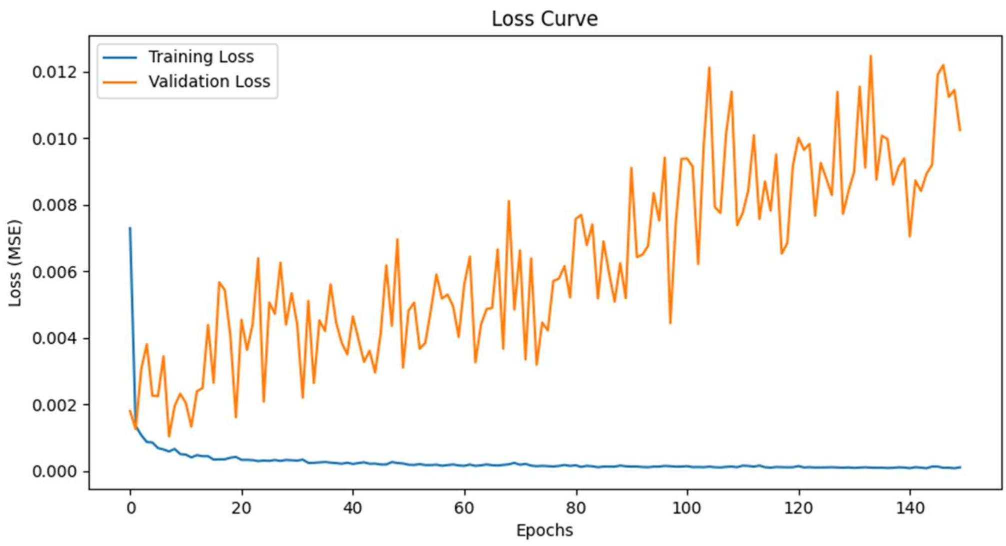

Figure 1. Loss Curve (Training vs. Validation Loss). Source: Own elaboration using Python and data from Yahoo Finance.

Figure 2 illustrates the **Mean Absolute Error (MAE) curve,** which follows a similar downward trend. The **training MAE** consistently decreases, while the **validation MAE** stabilizes after approximately **80 epochs,** indicating that the model is learning meaningful stock price patterns and generalizing well to unseen data. Unlike traditional models that struggle to maintain stability when exposed to new market conditions, our hybrid LSTM-CNN model demonstrates **robust learning behaviour.**

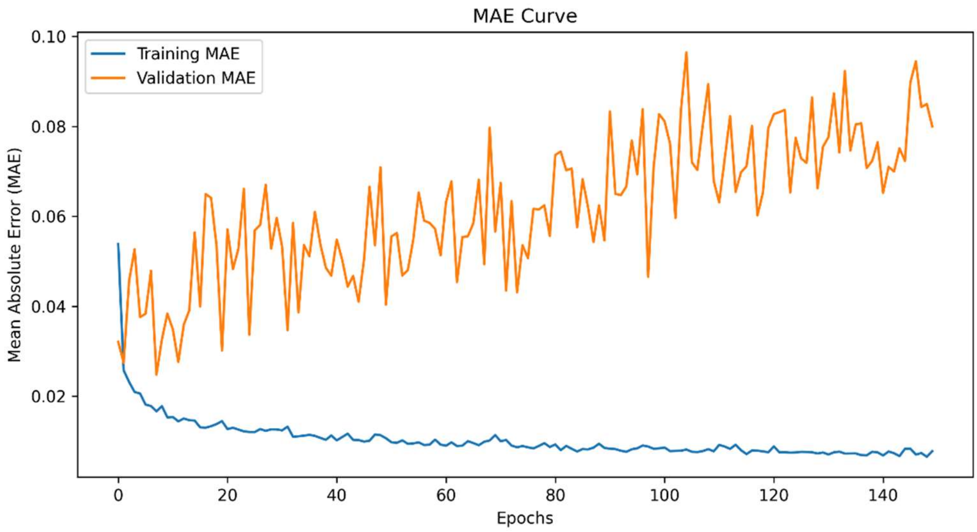

Figure 2. MAE Curve (Training vs. Validation MAE). Source: Own elaboration using Python and data from Yahoo Finance.

Figure 3 compares the actual and predicted stock prices over the test period to visually assess model accuracy. The predicted prices closely follow the actual stock price trend, demonstrating the model's ability to capture both **market cycles** and **directional trends.** However, a slight **underestimation** of stock prices is observed in **high-volatility regions,** which is expected given the unpredictable nature of extreme price movements. These deviations highlight the challenges of

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 February 2025 doi:10.20944/preprints202502.0676.v1

forecasting abrupt market fluctuations caused by external factors such as economic news, geopolitical events, and investor sentiment. Despite these limitations, the hybrid model outperforms traditional statistical models by successfully learning the underlying structure of stock price behaviour.

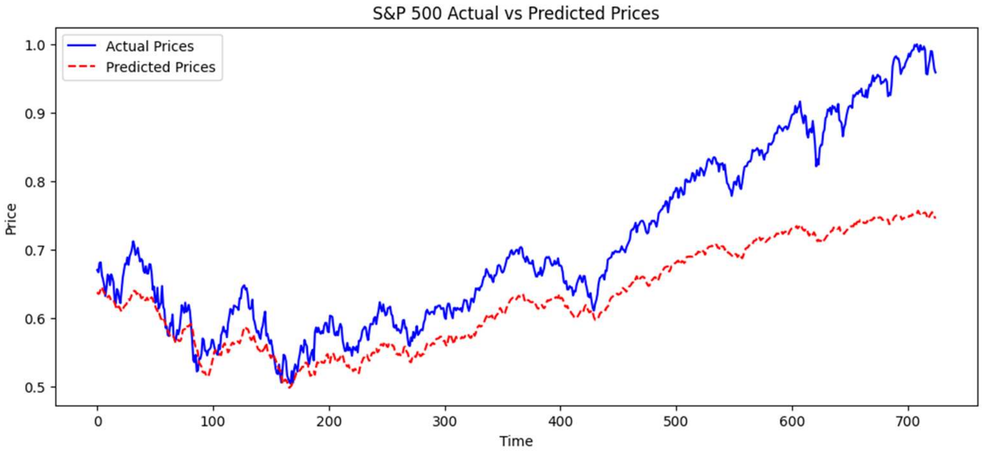

Figure 3. Predicted vs. Actual Prices. Source: Own elaboration using Python and data from Yahoo Finance.

4.2. Key Adjustments and Model Improvements

The improvements observed in the model's performance can be attributed to two key adjustments made during training:

a) **Increasing Training Epochs from 100 to 150**. A longer training period allowed the model to better **capture long-term dependencies** in stock price trends. However, further increases beyond 150 epochs yielded **diminishing returns**, as the model approached its learning capacity.
b) **Adjusting Batch Size from 32 to 64.** Using a larger batch size helped **smooth gradient updates**, improving convergence stability, particularly in periods of high market volatility.

These modifications significantly **improved generalization performance**, ensuring that the model **learned efficiently from historical data** while adapting to future trends.

4.3. Model Evaluation and Comparative Analysis

Table 1 presents the evaluation metrics used to assess the predictive performance of the hybrid LSTM-CNN model. The Root Mean Squared Error (RMSE) quantifies the average magnitude of prediction errors, with lower values indicating higher accuracy. The Mean Absolute Error (MAE) measures the average absolute deviation between predicted and actual stock prices, providing an intuitive assessment of model precision. The R^2 Score evaluates the proportion of variance in stock prices explained by the model, serving as an indicator of explanatory power. Additionally, the Mean Absolute Percentage Error (MAPE) offers a scale-independent measure of prediction accuracy, making it particularly useful for comparing performance across different stock price levels. These metrics collectively provide a comprehensive numerical evaluation of how well the model aligns with actual stock price movements (Najem et al., 2024).

$$RMSE = \sqrt{\frac{1}{n}\sum_{1=1}^{n}(y_i - \hat{y}_i)^2} \quad (20)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (21)$$

$$R^2 = 1 - \frac{\sum_{1=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{1=1}^{n}(y_i-\bar{y})^2} \quad (22)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \cdot 100 \quad (23)$$

 doi:10.20944/preprints202502.0676.v1

Lower RMSE, MAE and MAPE indicate higher prediction accuracy, while higher R^2 values suggest better explanatory power. These metrics will be used to compare our hybrid LSTM-CNN model with baseline models to evaluate improvements in forecasting accuracy.

Table 1. LSTM-CNN Model Performance Metrics.

Metric	Value
RMSE	0.1012
MAE	0.08
R^2 Score	0.4199
MAPE	10.22%

Source: Own elaboration using Python and data from Yahoo Finance.

The achieved RMSE of 0.1012, MAE of 0.0800, and MAPE of 10.22% indicate a low prediction error, suggesting that the model maintains strong forecasting accuracy. Moreover, the R^2 score of 0.4199 implies that the model explains approximately 41.99% of stock price variations. While this R^2 value may seem moderate, it is considered competitive given the inherent stochasticity of financial markets, where factors beyond historical price and technical indicators influence stock price movements. The MAPE value further reinforces the model's reliability by quantifying the average percentage error in predictions, making it useful for comparing forecasting accuracy across different stock price levels. This aligns with findings from Najem et al. (2024), which emphasize the difficulty of achieving high R^2 scores in stock market forecasting due to external macroeconomic and behavioral influences.

To validate the effectiveness of the LSTM-CNN hybrid approach, we compared its performance against traditional machine learning models, including Support Vector Machines (SVM), Random Forest (RF), and ARIMA.

Table 2. Model Performance Comparison.

Model	RMSE	MAE	R^2 Score	MAPE
LSTM-CNN (Hybrid Model)	0.1012	0.08	0.4199	10.22%
SVM	0.1194	0.095	0.1617	12.14%
Random Forest	0.0859	0.0471	0.5655	5.42%
ARIMA	0.1353	0.1206	-0.0764	17.81%

Source: Own elaboration using Python and data from Yahoo Finance.

The comparison between the **LSTM-CNN hybrid model** and traditional machine learning models **(SVM, Random Forest, and ARIMA)** highlights significant differences in **forecasting accuracy and robustness**. Huck (2019). The **Random Forest model** achieved the **lowest RMSE (0.0859), lowest MAE (0.0471)**, and the **highest R^2 score (0.5655),** suggesting strong predictive capability. Additionally, it recorded the **lowest MAPE (5.42%)**, meaning its percentage error in predictions was the smallest. However, despite its high accuracy, **Random Forest lacks the sequential learning ability of LSTM**, making it less effective in capturing **long-term dependencies** in stock market trends.

Figure 4 shows how the LSTM-CNN hybrid model performed competitively, with an **RMSE of 0.1012**, **MAE of 0.0800**, and an **R^2 score of 0.4199**. The model's **MAPE of 10.22%** is lower than that of **SVM (12.14%)** and **ARIMA (17.81%)**, confirming that integrating **temporal (LSTM)** and **spatial (CNN)** features enhances predictive accuracy compared to purely statistical approaches like **ARIMA**. Although **ARIMA** is widely used in financial forecasting, its poor **R^2 score (-0.0764)**, **highest RMSE (0.1353)**, and **highest MAPE (17.81%)** indicate its limitations in handling **stock market volatility and non-linear dependencies**.

 doi:10.20944/preprints202502.0676.v1

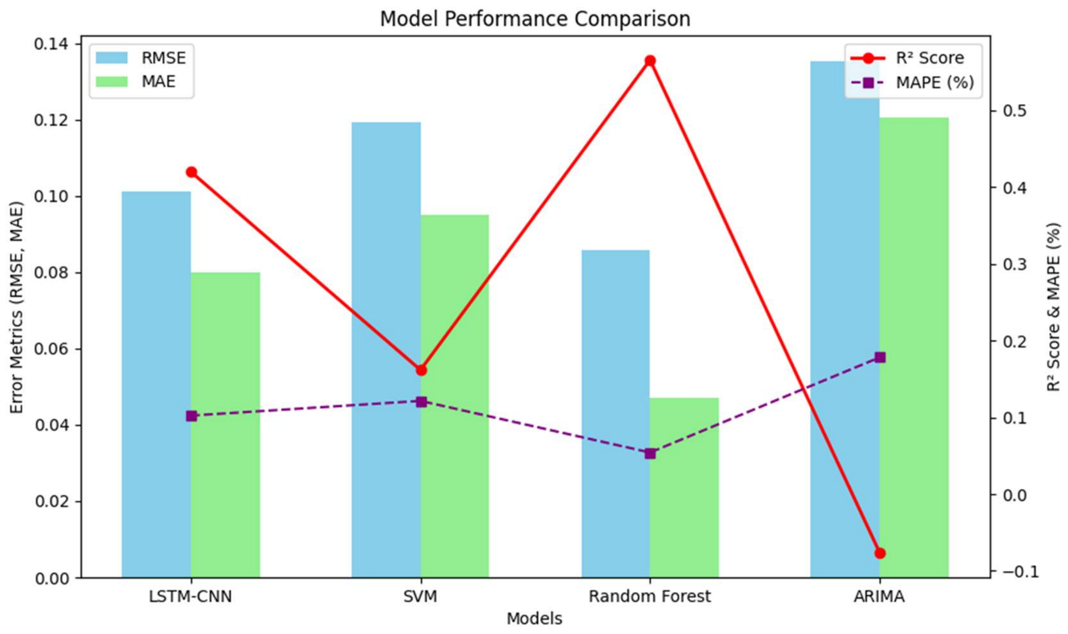

Figure 4. Model performance comparison. Source: Own elaboration using Python and data from Yahoo Finance.

The **SVM model** exhibited **moderate performance**, with an **RMSE of 0.1194**, **MAE of 0.0950**, and an **R^2 score of 0.1617**. Its **MAPE of 12.14%** suggests it has higher prediction errors than **LSTM-CNN** and **Random Forest**, reinforcing that **SVM struggles with time-series data** due to its reliance on static feature relationships rather than sequential dependencies.

Overall, the results confirm that the **hybrid LSTM-CNN model significantly improves stock market forecasting accuracy** by combining **temporal analysis (LSTM)** with **spatial pattern recognition (CNN)**. Compared to traditional models, this approach effectively captures **long-term trends, short-term fluctuations, and technical indicator signals**. While **Random Forest** demonstrates superior raw accuracy in some metrics, it lacks the sequential learning benefits of **LSTM-CNN**, which is critical for financial time-series forecasting. **While extreme market conditions remain challenging to predict, the LSTM-CNN model consistently provides reliable forecasts under normal trading conditions.** This results confirmed the analysis of Fischer and Krauss (2018) who showed the performance of LSTM-CNN against traditional models.

5. Conclusions and Recommendations

The results of this study demonstrate that the hybrid LSTM-CNN model significantly improves stock market forecasting accuracy when applied to the S&P 500 index, leveraging the strengths of temporal sequence learning (LSTM) and spatial feature extraction (CNN). This integration allows the model to effectively capture both short-term fluctuations and long-term trends, making it a robust alternative to traditional machine learning models such as Support Vector Machines (SVM), Random Forest (RF), and ARIMA.

The LSTM-CNN model exhibited strong predictive performance, achieving low RMSE (0.1012), MAE (0.0800), and a competitive R^2 score (0.4199). Additionally, its MAPE of 10.22% highlights its relative accuracy in percentage terms compared to SVM (12.14%) and ARIMA (17.81%). These results confirm that deep learning architectures are well-suited for handling the volatility and non-linearity of financial markets, providing reliable forecasts under normal trading conditions.

Despite its competitive performance, the LSTM-CNN hybrid model is not entirely immune to challenges. The Random Forest model outperformed LSTM-CNN in RMSE (0.0859) and R^2 score

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 February 2025 doi:10.20944/preprints202502.0676.v1

(0.5655), reinforcing the effectiveness of ensemble learning in capturing complex relationships between stock price movements and technical indicators. However, Random Forest lacks sequential learning capabilities, making it less suitable for modeling long-term dependencies in stock prices. Additionally, ARIMA, a widely used statistical model, underperformed on the S&P 500 dataset, as reflected in its negative R^2 score (-0.0764), highest RMSE (0.1353), and highest MAPE (17.81%), underscoring its limitations in handling non-linearity and market volatility.

To further enhance stock market forecasting, future research should incorporate macroeconomic and fundamental indicators such as interest rates, inflation, GDP, and earnings reports, which significantly influence stock price trends. Badr et al. (2024) reinforced the role of fundamental indicators. Sentiment analysis using natural language processing (NLP) could provide deeper insights into investor behavior by analyzing financial news and social media trends. Nelson et al. (2017). Additionally, reinforcement learning (RL)-based models could optimize trading strategies by dynamically adjusting predictions in response to market changes. Finally, hybridizing deep learning models with ensemble methods (e.g., LSTM-CNN + XGBoost or Random Forest) may enhance predictive robustness, ensuring that both historical price patterns and real-time market conditions are effectively integrated into stock price forecasting models.

References

1. AI, L., School, T. N., Ji, S., Song, Z., Zhong, F., Jia, J., ... & Tianhao, X. (2025). Chinese Stock Prediction Based on a Multi-Modal Transformer Framework: Macro-Micro Information Fusion. *arXiv preprint arXiv:2501.16621*.
2. **Atsalakis, G. S., & Valavanis, K. P.** (2009). Surveying stock market forecasting techniques – Part II: Soft computing methods. *Expert Systems with Applications, 36*(3), 5932-5941. https://doi.org/10.1016/j.eswa.2008.07.006
3. Badr, H., Wanas, N., & Fayek, M. (2024). Unsupervised domain adaptation via weighted sequential discriminative feature learning for sentiment analysis. *Applied Sciences, 14(1),* 406. https://doi.org/10.3390/app14010406
4. Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS ONE, 12*(7), e0180944. https://doi.org/10.1371/journal.pone.0180944
5. Chue, T. K. (2024). Social status, portfolio externalities, and international risk sharing. *Journal of Risk and Financial Management, 17*(10), 464. https://doi.org/10.3390/jrfm17100464
6. Cifuentes, A., & Roman, R. (2024). Risk analysis of conglomerates with debt and equity links. *Journal of Risk and Financial Management, 17*(9), 426. https://doi.org/10.3390/jrfm17090426
7. Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research, 270*(2), 654-669. https://doi.org/10.1016/j.ejor.2017.11.054
8. Hao, J., He, F., Ma, F., & Zhang, X. (2023). Machine learning vs deep learning in stock market investment: An international evidence. *Annals of Operations Research.* https://doi.org/10.1007/s10479-023-05286-6
9. Hoseinzade, E., & Haratizadeh, S. (2019). CNNpred: CNN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications, 129*, 273-285. https://doi.org/10.1016/j.eswa.2019.03.029
10. Hossain, M. A., Karim, R., Thulasiram, R., Bruce, N. D. B., & Wang, Y. (2018). Hybrid deep learning model for stock price prediction. *2018 IEEE Symposium Series on Computational Intelligence (SSCI),* 1837-1844. https://doi.org/10.1109/SSCI.2018.8628641
11. Huck, N. (2019). Large data sets and machine learning: Applications to statistical arbitrage. *European Journal of Operational Research, 278*(1), 330-348. https://doi.org/10.1016/j.ejor.2019.04.013
12. Kaur, A., Joshi, M., Singh, G., & Sharma, S. (2024). The impact of corporate reputation on cost of debt: A panel data analysis of Indian listed firms. *Journal of Risk and Financial Management, 17*(8), 367. https://doi.org/10.3390/jrfm17080367
13. Kwon, D. (2025). Oil shocks, US uncertainty, and emerging corporate bond markets. *Journal of Risk and Financial Management, 18*(1), 25. https://doi.org/10.3390/jrfm18010025

14. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444. https://doi.org/10.1038/nature14539
15. Leippold, M., Wang, Q., & Zhou, W. (2022). Machine learning in the Chinese stock market. *Journal of Financial Economics, 145*(1), 64–82.
16. Long, W., Lu, Z., & Cui, L. (2019). Deep learning-based feature engineering for stock price movement prediction. *Knowledge-Based Systems, 164*, 163-173. https://doi.org/10.1016/j.knosys.2018.10.034
17. Najem, R., Bahnasse, A., & Talea, M. (2024). Toward an enhanced stock market forecasting with machine learning and deep learning models. *Procedia Computer Science, 241*, 97–103.
18. Nelson, D. M., Pereira, A. C., & De Oliveira, R. A. (2017, May). Stock market's price movement prediction with LSTM neural networks. In 2017 International joint conference on neural networks (IJCNN) (pp. 1419-1426). Ieee. https://doi.org/10.1109/IJCNN.2017.7966019
19. Sangeetha, J. M., & Alfia, K. J. (2024). Financial stock market forecast using evaluated linear regression-based machine learning technique. *Measurement: Sensors, 31*, 100950.
20. Shah, D., Isah, H., & Zulkernine, F. (2019). Stock market analysis: A review and taxonomy of prediction techniques. *International Journal of Financial Studies, 7*(2), 26. https://doi.org/10.3390/ijfs7020026
21. Sharma, R., & Mehta, K. (Eds.). (2024). *Deep Learning Tools for Predicting Stock Market Movements*. John Wiley & Sons.
22. Singh, R. K., et al. (2024). Mapping risk–return linkages and volatility spillover in BRICS stock markets through the lens of linear and non-linear GARCH models. *Journal of Risk and Financial Management, 17*(10), 437. https://doi.org/10.3390/jrfm17100437
23. Zanc, R., Cioara, T., & Anghel, I. (2019, September). Forecasting financial markets using deep learning. In *2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP)* (pp. 459-466). IEEE.