Journal of Computer Science

Research Article

Deep Learning Models for Predicting Stock Closing Prices in the Saudi Stock Market

Areej Alhumaidi[1] and Hoda A. Abdelhafez[1,2]

[1]*Information Technology Department, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia*
[2]*Faculty of Computers and Informatics, Suez Canal University, Ismailia, Egypt*

Article history
Received: 21-09-2025
Revised: 30-12-2025
Accepted: 20-01-2026

Corresponding Author:
Hoda A. Abdelhafez
Information Technology Department, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia
Email: hodaabdelhafez@gmail.com

Abstract: Financial time series data are inherently volatile and nonlinear, presenting considerable challenges in accurately forecasting stock market trends. This study aims to predict the closing prices of companies listed on the Saudi stock market by analyzing historical data spanning from 2013 to 2023. To this end, four deep learning models were employed-Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Gated Recurrent Units (GRU), and Convolutional Neural Networks (CNN). In addition, two hybrid models, GRU-LSTM and CNN-LSTM, were developed to enhance predictive performance. The models were evaluated using five key regression metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). The findings indicate that the GRU and GRU-LSTM models demonstrated superior performance relative to the other models. The GRU model attained the lowest MSE and RMSE, demonstrating superior overall accuracy and robustness, while the GRU-LSTM model achieved the lowest MAE and MAPE, reflecting more precise pointwise and relative error estimates. Although the LSTM and CNN-LSTM models exhibited reasonable performance, the Bi-LSTM and CNN models were comparatively less effective. This study underscores the efficacy of deep learning, particularly GRU-based and hybrid architectures, in forecasting stock prices with high accuracy. The results offer valuable insights for investors and financial institutions operating within the Saudi stock market, demonstrating the practical implications of applying advanced deep learning methodologies. Moreover, the study contributes to the broader body of literature on financial forecasting and supports the development of more informed, data-driven investment strategies.

Keywords: Deep Learning, Stock Market, Hybrid Models, Stock Market Prediction

Introduction

The stock market represents a vital component of the global financial and economic system, functioning as a central hub for daily transactions involving billions of dollars (Kumar et al., 2020). Its significance has attracted widespread attention from researchers across both financial and technical disciplines. Accurate prediction of stock price movements holds substantial value, given its potential impact on investment strategies and economic planning. However, forecasting stock trends remains inherently challenging due to the influence of various dynamic factors, including macroeconomic indicators, political events, corporate performance, and investor sentiment. Moreover, the semi-strong form of market efficiency limits the ability to generate predictive insights from publicly available information. High volatility and noise further complicate the modelling of stock market time series data. Despite these challenges, deep learning models designed to emulate the neural processes of the human brain have shown considerable promise in identifying hidden patterns and delivering meaningful forecasts for investors and financial analysts (Göçken et al., 2016).

Investors, traders, and analysts have long grappled with the challenge of predicting stock market patterns.

Traditionally, stock market analysis has relied on technical or fundamental analysis. Fundamental analysis draws on financials, management insights, and industry trends, while technical analysis focuses on interpreting price patterns and market data to identify trading opportunities (Ilyas et al., 2022). The Saudi stock market, known as Tadawul, stands as a major exchange in the MENA region and is the largest in the area. It plays a crucial role in both Saudi and regional financial markets by offering investment opportunities to domestic and international investors. The Tadawul All Share Index (TASI), serving as the primary market index, encompasses a wide variety of companies, each marked by distinct characteristics, tendencies, and projected paths. The stability and vitality of this market reflect Saudi Arabia's economic strength and progress, highlighting the importance of accurate market price predictions. Predictive modelling in finance, a complex and nuanced field, aims to forecast future financial patterns and outcomes using historical and current market data. Traditional linear models frequently prove inadequate in capturing the complexities of the stock market due to the intricate nature of financial time series data and their inherent nonlinearity (Gandhmal and Kumar, 2019). Recent studies have demonstrated that advancements in artificial intelligence and deep learning have substantially improved stock market prediction by enabling the extraction of complex, non-linear patterns from large-scale financial time-series data. Neural network–based deep learning models, such as LSTM and GRU, have been reported to outperform traditional statistical and machine-learning approaches in forecasting stock prices and market movements due to their strong feature-learning and temporal modelling capabilities (Fischer and Krauss, 2018; Sezer et al., 2020).

As stock prices are shaped by a myriad of factors within complex and dynamic systems, developing models that can reliably predict stock prices is a formidable task. Successfully addressing this challenge could empower investors to make more informed decisions, potentially increasing their returns and reducing the risk of market downturns (Obthong et al., 2020). Additionally, Stock market data are often noisy, incomplete, and inconsistent, which complicates modelling and prediction. Effectively Addressing these data quality challenges is essential for achieving accurate short-term stock price forecasts, which can enhance investment decisions and business opportunities. Moreover, it contributes to a deeper understanding of deep learning techniques in predicting stock prices, offering insights that may be transferable to other financial markets or analogous problems in various domains. This highlights the importance of this research study in advancing predictive modelling to enhance the accuracy and reliability of financial forecasting.

This study aims to explore the potential of deep learning for stock market price prediction through three primary objectives: Evaluating the accuracy of deep learning models in forecasting stock prices, comparing different architectures to determine the most effective models, and developing hybrid models that integrate multiple architectures to enhance predictive accuracy and reliability. Although deep learning has been applied to stock market prediction, few studies have systematically compared multiple architectures, including hybrid models, on large-scale datasets from under-studied markets such as Saudi stock companies. Moreover, previous research often lacks alignment between model selection and dataset characteristics and does not fully assess the advantages of hybrid architectures. This study addresses these gaps by developing, benchmarking, and evaluating both individual and hybrid deep learning models to improve stock price forecasting and support informed investment decisions.

Because financial time-series data are highly non-linear, volatile, and exhibit both short- and long-term dependencies, which motivates the use of hybrid deep learning architectures rather than single models. The GRU-LSTM model is chosen to balance computational efficiency and long-term dependency learning, where GRU layers capture short-term market fluctuations with fewer parameters, and LSTM layers model long-term trends. The CNN-LSTM model is selected to extract local features from historical data, where CNN layers learn short-term price patterns and LSTM layers capture temporal dependencies across trading periods.

Predictive accuracy is evaluated using five regression metrics: Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) to penalize large prediction errors, Mean Absolute Error (MAE) for average absolute deviations, Mean Absolute Percentage Error (MAPE) for relative error measurement, and the coefficient of determination (R^2) to assess the explained variance in stock closing prices. Lower MSE, RMSE, MAE, and MAPE values, combined with higher R^2 scores, indicate improved predictive performance and provide a comprehensive measure of forecast accuracy and reliability.

While the deep learning models and hybrid models are widely studied in literature, the originality of this work is not lie in proposing new model structures. Instead, its primary contribution is the scale, empirical breadth, and contextual specificity of the analysis. The study conducts a comprehensive evaluation of six deep learning models across 131 publicly listed companies in the Saudi Exchange (Tadawul) over a ten-year period. This represents one of the most extensive comparative studies of deep learning–based stock price prediction within the Saudi financial market. The study offers valuable empirical insights into model behavior, robustness, and generalizability that are not captured in prior, smaller-scale or cross-market studies.

Related Work

AI approaches have proven successful in various domains, including stock market analysis. Numerous studies have applied DL and AI techniques to analyze economic data across various nations and environments. The data-driven decision-making approach has gained popularity due to AI's ability to generalize and model data, providing a new way to understand behavior and uncover patterns. This section provides an overview of relevant literature reviews, discussing its strengths and limitations.

Predicting Stock Market Using Deep Learning Methods

Padh Ghosh et al. (2024), employed the National Stock Exchange (NSE) of India and the New York Stock Exchange (NYSE) historical data to train four robust deep learning models, MLP, RNN, LSTM, and CNN, to forecast stock prices. The primary focus of the analysis was on the day-wise closing prices of each stock, as these were typically the basis for investment decisions. Testing was conducted on a diverse array of NSE and NYSE companies, in addition to training utilizing Tata Motors data obtained from the NSE. The size of the window and duration of the prediction were carefully considered, with 200 window sizes found to be ideal for a 10-day forecast period. The authors used MAPE to assess the accuracy of the predicted output. The CNN model demonstrated outstanding accuracy in predicting NYSE stock prices, despite being trained on NSE data. The research emphasized the effectiveness of neural networks in forecasting stock market trends, showcasing their ability to capture intricate nonlinear patterns that traditional linear models like ARIMA struggle with. The study highlighted the interdependence of many markets and the potential of deep learning structures in properly predicting stock values.

Sako et al. (2022) explored the application of simple RNN, LSTM, and GRU models to forecast the closing prices of six foreign exchange rates and eight stock market. The study employed both univariate and multivariate models, with the univariate models focusing on the closing price and the multivariate models incorporating high, low, open, and closing. Preprocessing steps included normalization using the MinMaxScaler function, scaling the data range between zero and one, and replacing missing values with the median of each column. The study also used the sliding window approach to transform time series data into a mini-batch, which consisted of a sequence, time steps, and observations at a specific time step. The GRU model shown superior performance in univariate out-of-sample forecasting for currency exchange rates and multivariate out-of-sample forecasting for stock market indexes.

Alkhatib et al. (2022) presented a research study exploring the application of deep learning models for adjusted closing price prediction. It introduces a novel feature set comprising six variables (High, Low, Volume, Open, HiLo, and OpSe) for improved forecasting accuracy, contrasting with the traditional four-variable approach. The study assessed the efficacy of six deep learning models (MLP, GRU, LSTM, Bi-LSTM, CNN, and CNN-LSTM). The study evaluated the impact of data size by using datasets of different scales from various sectors, such as Apple, Tesla, Snapchat, and ExxonMobil. The data preprocessing method involved normalization using a min-max scaler. Then, split into 70% for training, 15% for testing, and 15% for validation. This specific distribution was chosen to prevent overfitting of the models and to ensure accurate evaluation. The study's results demonstrated that LSTM-based models showed enhanced performance when utilizing the novel feature approach. However, all models provided comparable results, with no model consistently surpassing the others. Ultimately, the introduction of new characteristics had a beneficial impact on the performance of the prediction models.

Al-Nefaie and Aldhyani (2022), explored the use of LSTM and MLP models for forecasting stock market prices. It focused on the four sectors of the Saudi Stock Exchange (Tadawul): communication, energy, financial, and industrial, from 2018 to 2020, with twelve features are included in these records: The lowest and highest price, the opening and closing price, the stock market's change, the quantity exchanged, and the number of trades. The closing price was regarded as a reliable indicator for forecasting the future values of the stock market over a 60-day timeframe. and assessed their performance using four metrics: MSE, RMSE, NRMSE and R-squared. The study effectively demonstrated the high efficacy of LSTM and MLP models in predicting stock market trends, with both models achieving R2 values greater than 99%, indicating excellent predictive accuracy and robustness in their results.

Faraz et al. (2020) proposed a novel stock market prediction strategy using autoencoder long short-term memory (AE-LSTM) networks. This approach integrated technical analysis with deep learning methods by adding technical indicators and oscillators as features to the dataset for predicting the closing price for the S&P 500 index. The dataset for this study comprised historical records of the S&P 500 index over a span of twenty years, from March 1, 2000, to April 11, 2019, obtained from Yahoo Finance. and features such as open, high, low, closing, and adjusted closing prices, along with trading volume. The model was assessed using several evaluation metrics, including MAE, RMSE, and average return (AR). The results showed that AE-LSTM outperforms Generative Adversarial Networks (GAN) in predicting the S&P 500's daily adjusted closing price, proving the effectiveness of this prediction method in stock market forecasting (Song and Choi, 2023). Moghar and Hamiche (2020), investigated the application of RNN-based LSTM models for predict stock market values. The paper's primary aim was to assess the precision of machine learning algorithms, particularly LSTM, in stock market prediction and to explore how the number of training epochs influences the model's performance. The dataset comprised the daily opening prices of two stocks, Google (GOOGL) and Nike (NKE), from the New York Stock Exchange. The data covered different time periods for each stock, their approach involved training the model with various epochs (12, 25, 50, and 100), was evaluated using MSE. The finding was that increasing the number of epochs led to a decrease in loss for both stocks and noted the significant impact of data length on the result, indicating enhanced predictive accuracy of the model with more extensive training. This outcome underscored the effectiveness of LSTM models and the importance of sufficient training in achieving high accuracy.

Muhammad et al. (2024) introduced a deep learning model integrated with explainable AI (XAI) techniques to predict multi-class stock market trends. It assessed the model's performance on four key indices, S&P500, DAX30, FTSE100, and Nikkei225, using historical data from Yahoo! Finance spanning 1990 to 2022. This dataset encompassed various market cycles, economic crises, and major global events. To improve prediction accuracy, the study employed feature engineering, incorporating factors such as returns, moving averages, volatility, Relative Strength Index (RSI), and momentum alongside closing prices. The findings revealed that the deep learning model significantly outperforms conventional machine learning approaches, including Random Forest, Support Vector Machine, and Logistic Regression, achieving a 94.9% accuracy and an F1-score of 94.85%, demonstrating its robustness across diverse market conditions.

Barua et al. (2024) conducted a comparative analysis of deep learning models for predicting stock prices in the Indian stock market. It examined five models: Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), Gated Recurrent Units (GRU), and Attention LSTM. Stock price data were sourced from the National Stock Exchange of India, focusing on five major entities, HDFC, TCS, ICICI, Reliance, and Nifty 50, spanning the years 2016 to 2021. Model performance was assessed using MAE, MSE, RMSE, and R^2. The findings indicated that CNN and GRU excel in predicting stable stocks, while Attention LSTM was more effective for volatile stocks. RNNs performed the poorest due to their limitations in capturing long-term dependencies.

Predicting Stock Market Using Hybrid Deep Learning Methods

Song and Choi (2023) proposed novel hybrid models were developed to forecast both the one-time-step and multi-time-step closing prices. by using CNN-LSTM, GRU-CNN, and ensemble models. The dataset included three major stock market indices: DAX, DOW, and S&P500.These were analysed over different periods, including long periods from 2000 to 2019, and short periods before and after the COVID-19 pandemic. The authors introduced a unique feature, termed "medium", representing the average of the high and low price. Along with daily open, high, low, and medium prices, trading volume, and change (OHLMVC), the data was segmented using the sliding window technique. This technique involved selecting a predetermined window size of time-series data as the input and generating outputs for different look-back periods of 5, 21, and 42 days. The ensemble models exhibited significant effectiveness in forecasting stock market indices, outperforming benchmarks in various periods. From 2000 to 2019, the models surpassed benchmarks in 77.8% of cases for both MSE and MAE. Post-pandemic the models achieved 100% outperformance. However, the study also noted the increased challenge in long-term predictions, where errors grew with longer prediction steps, yet the ensemble models still outperformed in these longer-term forecasts.

Zaheer et al. (2023) proposed a hybrid deep-learning forecasting model based on CNN, RNN, LSTM, CNN-RNN, and CNN-LSTM. The models predict two stock parameters: The closing price and the high price for the next day. The dataset comprised daily trading data for the Shanghai Composite Index from July 3, 1997, to January 24, 2022, encompassing a total of 5951 trading days with six factors: Opening price, highest price, lowest price, closing price, adj-close, and trading volume. Data standardization was used Z-Score standardization to transform a set of data into a similar scale. The obtained outcome indicates that CNN had the poorest performance, while LSTM surpassed CNN-LSTM. Additionally, CNN-RNN outperforms CNN-LSTM and LSTM. Furthermore, the suggested single-layer RNN model outperforms all other models because its R2 value was the closest to 1. The experimental results confirm the efficacy of the suggested approach, which will aid investors in maximizing their profits through informed decision-making.

Karim et al. (2023), introduced a novel hybrid model, combining Bi-LSTM and GRU, for predicting the opening prices of stock. The model was tested using the NIFTY-50 stocks market dataset, which included daily data on opening, closing, highest, lowest, and volume-weighted average prices from January 2000 to March 2021. The dataset was split into various training and testing ratios for three different window sizes (30, 60, and 90) in fifteen different combinations to optimize performance and reduce bias. Before feeding into the model, the input raw data undergoes normalization using the MinMaxScaler function, followed by a 1D CNN layer for data pattern extraction. The hybrid model demonstrated superior accuracy and reduced error rates in comparison to the separate models. The hybrid model performed best with a window size of 60 and a training-testing split ratio of 65, resulting in the lowest error values.

Aldhyani and Alzahrani (2022) developed a model to predict for stock market values with financial time-series data as inputs. The model incorporated LSTM and a hybrid of CNN with LSTM to predict the closing prices of Apple and Tesla stocks, where 70% of the data was used for training, and the remaining 30% was used for testing. Using for evaluation metrics MSE, RMSE, NRMSE and Pearson's correlation R. Apple's stock, the CNN-LSTM model was more accurate and reliable, with R percentages of 99.83% during training and 99.48% during testing, as opposed to LSTM's 99.68 and 98.66%. This reinforces the conclusion that the CNN-LSTM model slightly outperformed LSTM, demonstrating its promising potential in stock market predictions. The study emphasized the importance of AI in financial forecasting and its ability to handle complex time series data.

Patra and Mohanty (2022) presented a hybrid LSTM-GRU for predicting the adjusted closing price of the S&P500 index. The study addressed the challenge of stock price prediction, which is complicated due to the nonlinearity and volatility of market data. The data spanned the period from January 3, 2000, to October 30, 2017, covering 4486 recorded days. The dataset contained six features: Open, high, low, close, volume, and adjusted close. However, for enhanced accuracy, the number of features was expanded to 25 by incorporating several technical indicators. The models were evaluated using several metrics such as: R2, MSE, return ratio, optimism ratio, and pessimism ratio. The proposed model showed superior performance compared to the standalone LSTM, GRU, and MLP models. It achieved a training loss of 0.00024, a return ratio of -3.3103, an MSE of 0.005, and an R2 score of 0.939. The results suggested that the hybrid model could more accurately and reliably predict stock prices by leveraging the strengths of both the LSTM and GRU architectures.

Jarrah and Salim (2019) proposed a hybrid model RNN combined with a Discrete Wavelet Transform (DWT) for predicting stock prices, utilizing these techniques to remove noise from the data and predict future stock prices. The study employed the Saudi stock market, which consisted of 146 stocks that were lanuched between January 1, 2011, and March 31, 2016. These records included five features, such as open, close, low, high, and volume daily. The study compared its approach with traditional prediction algorithms like ARIMA, showing the effectiveness of the DWT and RNN combination in improving prediction accuracy. The effectiveness of this approach was evaluated using standard criteria like MSE, MAE, and RMSE which were common metrics for assessing the accuracy of predictive models in finance.

Previous studies on stock market prediction have demonstrated the effectiveness of deep learning models such as LSTM, GRU, and CNN; however, several limitations remain unaddressed. First many studies focus on a single market, index, or a limited number of sectors, which restricts the generalizability of their findings. Second several studies report exceptionally high predictive accuracy without sufficiently addressing overfitting risks or robust validation strategies. Additionally, few studies provide a systematic comparison between machine learning models, and multiple deep learning architectures within the same experimental framework.

The novelty of this study lies in its large-scale evaluation of 131 companies in the Saudi stock market (Tadawul), an underexplored domain in stock prediction research. It combines multiple deep learning models (CNN, LSTM, Bi-LSTM, GRU) with hybrid architectures (CNN-LSTM, GRU-LSTM) to capture both temporal and complex nonlinear patterns in stock data. By integrating extensive market features, performing rigorous preprocessing, and evaluating models using multiple metrics (MSE, MAE, RMSE, MAPE, R^2), the research provides a robust, market-specific predictive framework that outperforms prior approaches and offers practical insights for investors and financial institutions.

Materials and Methods

This section consists of six phases. The first phase involved selecting and gathering the dataset from Saudi Exchange (Tadawul). The second phase focused on exploring and visualizing the data. The third phase was dedicated to pre-processing the dataset. In the fourth phase, deep learning models were applied. The fifth phase entailed evaluating the performance of these prediction models using five regression metrics. The final phase involved using the model to forecast the closing prices for the next five days and comparing the results with the actual values as shown in Figure 1.

Data Collection

This study utilized historical data related to stocks from the Saudi stock exchange, known as Tadawul. The downloaded data covered the time frame from January 1, 2013, to December 31, 2023, and included a total of 516,364 observations. The database contained information on stock transactions for 249 companies listed in Saudi Arabia, grouped by sectors related to different industries.

After applying the filtering criteria, the dataset was reduced to 131 companies with a trading history of 11 years, covering 2745 trading days for each, through a filtering process. The listed data for these companies included details like industry group, symbol, company name, date, opening and closing prices, highest and lowest prices of the day, volume of shares traded, number of trades, price change, and the value of trades in Saudi Riyals for each trading day. Table 1 offers a detailed explanation of each data feature.

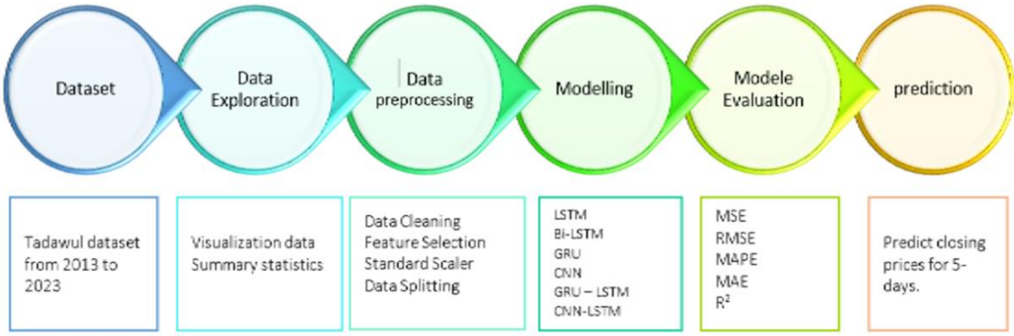

Fig. 1: Methodological phases for applying the proposed system

Table 1: Detailed explanation of all dataset features

Features	Description	Data Type
Industry of group	The category of the industry to which the company belongs.	String
Symbol	A unique four-digit code is used to identify a company.	Integer
Company name	The name of the company to which the stock belongs.	String
Date	Date of day exchange	Date
Closing price	The closing price of the same day	Decimal
Open price	The opening price of the same day	Decimal
Highest price	The highest price of the same day	Decimal
Lowest price	The lowest price of the same day	Decimal
No. of trades	Number of trades placed on a stock of the same day	Decimal
Change	The difference in the closing price of the stock from the closing price of the previous day.	Ratio
Volume traded	The number of shares that changed hands during the trading session.	Decimal
Value traded R. S	The total value of shares traded during the day	Decimal

Data Exploration

Data exploration is a fundamental initial phase in the data analysis process, aimed at gaining a deep understanding of the dataset's structure and uncovering inherent patterns. In this study, various exploration techniques were employed, including data visualization, temporal comparisons, statistical analysis, and time series decomposition into components such as trend, seasonality, and residuals. The study focuses on analyzing the daily stock trends of companies listed on the Saudi stock market that remained consistently active from 2013 to 2023. Notably, upward trends in value and volume traded are key indicators of increasing investor interest, reflecting both market valuation and liquidity, critical factors in investment decisions.

Data Preprocessing

Data preprocessing is an essential and crucial phase in building models for the purpose of predicting stock market prices. Data wrangling encompasses the processes of cleansing, converting, and structuring the data in a manner that optimizes the efficiency and efficacy of the analysis.

Data Cleaning: This process entails cleaning and transforming the data to enhance its quality, ensuring it is better suited for insight extraction or predictive modeling. It includes tasks such as verifying the data types of each feature (or column) in the dataset, identifying and handling missing values, and other preprocessing steps. For example, the presence of missing values can be addressed by either removing them or filling with the previous or the next value (Aseeri, 2023). Another technique involves using the median of each column to fill in missing values. The primary advantage of employing the median is its robustness in dealing with potential outliers (Sako et al., 2022).

Features Selection is the process of choosing a subset of pertinent features to be used in constructing a model. Dimensionality reduction is a technique that enhances model performance by lowering the number of features in the data. This, in turn, improves its interpretability, and can decrease the computation time needed for training and prediction. SelectKBest is a commonly employed feature selection strategy that identifies the top K features with the strongest link to the target variable using statistical testing such as mutual_info_regression by estimating how much information the presence or absence of a feature contributes to making accurate predictions about the target variable. Features that have higher mutual information scores are seen as more significant (Otchere et al., 2022). The correlation matrix is simply a table that displays the correlation coefficients for different variables. It provides information about the relationship

DOI: 10.3844/jcssp.2026.1370.1386

between features and the target feature. A correlation matrix is a visual representation of data that helps detect the degree of relationship between several features and the target feature. The dataset represents each attribute using colors, which convey information to researchers about the association between features (Buyrukoğlu and Akbaş, 2022).

Data Standardization: The z-score normalization is a technique used to scale numerical data prior to modeling in machine learning and statistics. This process transforms the features, so they have a mean (average) of 0 and a standard deviation of 1. The standardization helps ensure that each feature contributes equally to the distance calculations in the model, making it particularly important for algorithms that rely on distance measures (Nayak et al., 2014):

$$Z = x - \mu\sigma \quad (1)$$

Where x is the original value of the feature, μ is the mean of the feature across all data points, and σ is the standard deviation of the feature across all data points (Nayak et al., 2014).

Deep Learning Models

The following are the deep learning techniques that were implemented in this study.

Long Short-Term Memory (LSTM): The fundamental concept underlying LSTMs is the cell state, which functions as a type of continuous pathway that spans across all LSTM units in the sequence. The presence of this cellular state enables the transmission of significant data across the network, facilitating the retention of information over extended sequences. LSTMs possess the ability to discern which data to retain and which to discard, facilitated by three distinct types of gates (Baru, 2019). The forget gate uses a sigmoid function to regulate the retention or removal of information from the previous hidden state based on the current input, producing an output between 0 and 1. The input/ update gate determines the incorporation of new information into the memory through a sigmoid layer that controls information flow and a "tanh" layer that scales the importance of the selected values between -1 and 1. Output Gate initially employs a sigmoid layer to ascertain which components of the LSTM memory should influence the output. Subsequently, it applies a non-linear tanh function to scale these values to a range between -1 and 1. The final step involves multiplying this outcome by the output from another sigmoid layer (Baru, 2019). Benefiting from the three control gates and a memory cell, it efficiently manages long-term information storage and retrieval. The size of the weight matrix allows for customizable output dimensions. This architecture effectively handles long input-feedback delays while preventing gradient issues, thanks to a consistent error flow within its memory cell.

Bidirectional Long Short-Term Memory (Bi-LSTM): LSTM model is designed to train a network by leveraging both past and future sequences of input data. It processes this data through two interconnected layers, enabling the Bi-LSTM to predict or label each element in a sequence by considering the contextual information from both preceding and succeeding elements. This functionality is achieved by operating two LSTM networks in parallel: One processes the sequence in a forward direction (left to right), and the other in a backword direction (right to left). The combined prediction or tagging from these two directional flows is referred to as a composite output. Empirically, this approach has demonstrated significant effectiveness. Utilizing the LSTM architecture twice results in enhanced acquisition of long-term dependencies, ultimately leading to improved accuracy of the model.

Gated Recurrent Unit (GRU) shares similarities with LSTM, however, it has a more streamlined structure. The GRU model employs two gates: An update gate and a reset gate. The GRU model possesses a smaller number of parameters, which thus allows for potentially faster training or the ability to generalize with less input. However, when dealing with a significant amount of data, LSTMs that have a greater capacity for expressing complex patterns and relationships may yield superior outcomes (Medar et al., 2017).

Update gate functions similarly to the forget and input gates in an LSTM, determining which information to discard and which new information to retain. Reset gate determines how much of the previous concealed state to ignore and how much new input to consider for the current state. The gate is implemented as a sigmoid function, with a value of 1 allowing all information through and 0 blocking it (Medar et al., 2017).

Convolutional Neural Network (CNN) model consists of many convolutional and pooling layers, which are utilized to extract features. Conventional convolutional layers employ two-dimensional filters (kernels) and activation functions to analyze picture characteristics. However, in the field of stock prediction, Convolutional Neural Networks (CNNs) are employed to analyze time series data, which consist of one-dimensional features. CNNs for time series employ a one-dimensional filter that moves over the time series using a stride set by the data granularity, in order to account for the variation in data shape (Baru, 2019). The Convolution layer attempts to extract the most useful features from the one-dimensional matrix and performs calculations to get a complex output. The Pooling layer utilizes the convolutions' output as an input. The max pooling function is employed to choose the most significant features with high weights in the pooling layer. The output of the pooling layer is transmitted to the flatten layer. The main purpose of the flatten layer is to transform the data into a unified array format. The Fully connected layer takes the output of the

DOI: 10.3844/jcssp.2026.1370.1386

flatten layer and processes it to obtain the results (Zaheer et al., 2023).

Hybrid Models Techniques

Long Short-Term Memory

Gated Recurrent Unit (LSTM-GRU): The hybrid approach combines different techniques to improve the performance of prediction models. Hybrid algorithms enhance the effectiveness of prediction models by improving their efficiency. By integrating both GRU and LSTM layers, the model attempts to capture the benefits of each: The efficiency and simplicity of GRUs with the powerful long-term dependency modeling capabilities of LSTMs. This can potentially lead to improved performance and efficiency in tasks that involve complex sequential data, as the model can learn and retain information over both short and long sequences more effectively (Moghar and Hamiche, 2020; Hamayel and Owda, 2021).

Convolutional Neural Network – Long Short-Term

Memory (CNN-LSTM): CNNs, known for their ability to focus on prominent features within their field of view, are extensively applied in feature engineering tasks. Conversely, LSTMs excel in processing data in a sequential, time-dependent manner, making them ideal for time series analysis. Leveraging the distinctive advantages of CNNs and LSTMs, a stock forecasting model employing a CNN-LSTM architecture has been developed. This model comprises an input layer, a one-dimensional convolutional layer, a pooling layer, an LSTM hidden layer, and a fully connected layer, combining the strengths of both CNNs and LSTMs to predict stock market effectively (Saud and Shakya, 2020).

Evaluation Metrics

Evaluating models is an essential component of every research project. When assessing the effectiveness of different models. The following evaluation metrics were used in the research study. The evaluation metrics are widely used in regression analysis, forecasting, and machine learning to evaluate the performance of predictive models (Saboor et al., 2023; Medar et al., 2017).

Mean square Error (MSE): Is a metric used to evaluate the performance of a predictor. It is always non-negative, with values closer to zero indicating higher performance. Mean Squared Error (MSE) takes into consideration both the bias and variance of a prediction model, which refers to the extent to which the forecasts differ between different data samples. MSE is calculated as the second moment of error around the origin. The Mean Squared Error (MSE) quantifies the deviation of anticipated values from the mean, similar to how it calculates the distance from the true origin. The formula for MSE is (Saboor et al., 2023).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2 \quad (2)$$

Mean Absolute Error (MAE) computes the mean of the absolute differences between the expected and actual values. A lower MAE value indicates a model that predicts more closely to the actual data points. The formula for MAE is (Jingyi et al., 2022):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \widehat{Y}_i\right| \quad (3)$$

(RMSE) is metric used to assess the accuracy of predictions. It calculates the square root of the average squared differences between predicted values and actual values. RMSE is like Mean Squared Error (MSE), but it includes the square root step to directly relate the errors to the units of the variable being predicted. The formula for RMSE is (Jingyi et al., 2022):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2} \quad (4)$$

Mean Absolute Percentage Error (MAPE) is a metric that computes the average of the absolute percentage errors between the projected values and the actual values. It offers valuable information about the relative magnitude of the errors expressed as percentages. The formula for MAPE is (Jingyi et al., 2022):

$$MAPE = \frac{100\%}{N}\sqrt{\sum_{i=1}^{n}\left|\frac{Y_i - \widehat{Y}_i}{Y_i}\right|} \quad (5)$$

n is the number of observations, Y_i is the actual value, and $\widehat{Y}_i$ is the predicted value.

The coefficient of determination (R^2) is a statistical metric that indicates the proportion of variance in the dependent variable that is explained by the independent variable (s) in a regression model. The value ranges from 0 to 1. In simpler terms, it quantifies how well a model explains and predicts future outcomes, based on the comparison between the model's predictions and the actual data. The formula for R^2 is (Patra and Mohanty, 2022):

$$R^2 = 1 - \frac{SR}{TR} \quad (6)$$

Where *TR* is the total square sum and *SR* is the sum of square residuals.

Implementation

During the implementation phase, SelectKBest was employed for feature selection. After completing data preprocessing, the resulting multivariate time-series data were utilized for model development. Google Colab was chosen as the implementation environment due to its cloud-based architecture, Python support, access to high-performance GPUs, and Pro+ features, which enable efficient training of deep learning models.

DOI: 10.3844/jcssp.2026.1370.1386

For applying deep learning models, the data must be cleansed, consistent, and standardized to the same range. As a result, diverse pre-processing steps must be executed on the data, contingent upon its characteristics. Consequently, the following steps were implemented to adequately prepare the data for subsequent analysis.

Companies Selection: Company selection was based on the trading days threshold. It involved counting the occurrences of each company name in the dataset to determine how many trading days were recorded for each company. Then companies were screened based on their activation for exactly 2,745 trading days, corresponding to the total number of trading days spanning from 2013 to 2023. Through this filtering mechanism, a list was compiled, comprising the names of companies that meet this criterion of consistent trading activity throughout the specified timeframe.

Handling Missing Values: The dataset sourced from Tadawul included instances of companies with zero values. This phenomenon is attributed to punitive measures enforced by Capital Market Authorities or due to financial losses experienced by the companies, leading to a cessation of trading activities. Consequently, any company that underwent a suspension of trading for a duration of one month or more was omitted from further analysis.

There were 303 records remaining missing value for the feature's open, low, and high, with 5 records remaining for the close. To rectify this situation, two methodologies were used. The first method, known as forward filling, involved replacing missing values with the nearest preceding non-missing value for the same attribute. The second method used the median value of each column to fill in the missing values. Both strategies produced identical results when applied to the dataset.

Feature selection: To identify the features with the strongest relationships to the dependent variable, two complementary feature selection methods were employed. The first method, SelectKBest, utilized mutual information regression to quantify the nonlinear dependency between each feature and the closing price. The results, presented in Table 2, indicate that price-related variables such as Low, High, and Open exhibit the highest mutual information scores, reflecting their strong predictive relevance. In contrast, the number of trades variable shows significantly lower mutual information and correlation values, leading to its exclusion from subsequent analysis.

The second method involved a correlation matrix to analyze linear relationships and potential multicollinearity among the features. As illustrated in Figure 2, the closing price is strongly correlated with Open, High, and Low prices, confirming the relevance of these variables. By jointly considering the mutual information scores and the correlation structure, the final feature set was selected to retain highly informative variables while avoiding redundant or weakly contributing features. This combined strategy enhances model robustness and reduces noise, ultimately contributing to improved predictive accuracy of the proposed deep learning models.

Table 2: Feature Importance Based on Mutual Information Regression

Feature	MI score
Low	4.651685
High	4.537840
Open	3.928191
Volume Traded	0.152864
Value Traded (SAR	0.049313
No. of Trades	0.032274

Fig. 2: Correlation coefficients among various features

Data Splitting*:* The variables in the study were categorized into features (X) and the target (Y). The features comprised open, low, high, volume traded, and value traded, while the target consisted of the closing price. The dataset was divided into training and testing sets based on specific dates: The training set covered the period from January 1, 2013, to December 31, 2022, and the testing set covered from January 1, 2023, to December 31, 2023. Given the sequential nature of the data, it was treated as a time series.

Figure 3 displays a visual depiction of the training and testing sets for the closing price. The curved line represents the highest closing price observed in the dataset, while the shaded area beneath the line illustrates the range of all the closing prices over time.

Data Standardization: Standardization is a crucial preprocessing technique that centers data to a mean of zero and scales it to a unit standard deviation, which helps accelerate model convergence and improve performance. After splitting the dataset into training and testing subsets, Z-score normalization was applied. This method subtracts the mean and divides by the standard deviation of the training data, transforming the features to have a mean of 0 and a standard deviation of 1. This process minimizes the impact of outliers while preserving the relative relationships among values within each feature.

After completing the data preprocessing phase, proceed to implementing the models to process multivariate time series data. Google Colab was selected as the development environment because it is a cloud-based platform that supports Python, provides high-performance GPUs, and offers Pro+ capabilities, which facilitate faster training of deep learning models. Six predictive models were utilized to predict the closing price of publicly listed companies in the Saudi stock market. The models incorporated in the analysis consist of LSTM, Bi-LSTM, GRU, and CNN, as well as two hybrid models: CNN-LSTM and GRU-LSTM. The architecture of LSTM, BiLSTM and GRU models comprises various parameters that are crucial for configuring the models and significantly impact their performance. These include the number of neurons, layers, dropout rate, optimizer (AdamW), epochs (100), learning rate (0.001), and batch size (64). Dropout is used to prevent overfitting, with LSTM and BiLSTM using a 0.2 rate while GRU a 0.0 drop rate means no units are dropped, maintaining them all during training. Convolutional Neural Network (Conv1D): Before training the model, it's crucial to preprocess the data. This includes scaling the data and reshaping it from a 2D array format (326976, 6) to a 3D array (326976, 6, 1) to meet the input requirements of CNN models. The architecture consists of input, max pooling, flattening, dense, batch normalization, dropout, and output layers, with training for 15 epochs. Hybrid models combine CNN and LSTM or GRU and LSTM for better time-series predictions. CNN-LSTM starts with a Conv1D layer (256 filters, kernel size 3), followed by max pooling, LSTM (128 units), batch normalization, dropout (0.1), and dense layers (128 units). GRU-LSTM combines a GRU layer (256 units) with an LSTM layer (256 units), using batch normalization, dropout (0.2), and dense layers (128 and 256 units). To optimize the performance of deep learning models throughout the training process, two crucial callbacks were used: The model checkpoint callback and the early halting callback. These callbacks helped in saving the best-performing model and prevent overfitting by stopping training when validation loss does not improve for 10 epochs.

Fig. 3: Training and testing dataset split

Results and Discussion

This section evaluates the performance of the six constructed models using five metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2).

Evaluating the Models for all Companies

Table 3 presents the evaluation metrics for the developed models to predict closing prices for 131 companies.

The GRU model outperformed other models in terms of MSE, RMSE, and R^2, indicating the lowest average squared errors, smallest root-mean-square error, and highest explanatory power for closing price variance. This superior performance can be attributed to the GRU's simpler architecture, which efficiently captures short-term dependencies with fewer parameters. Its streamlined structure reduces the risk of overfitting and allows faster convergence on volatile financial time series, making it particularly effective for the Saudi stock market dataset. The GRU-LSTM hybrid model achieved the lowest MAE and MAPE values (0.4370 and 2.2527, respectively), indicating more accurate predictions in terms of absolute and relative errors. The hybrid's advantage likely arises from the synergy between GRU and LSTM layers: GRU layers model short-term market fluctuations, while LSTM layers retain long-term temporal dependencies. This combination enables the model to balance sensitivity to recent trends with memory of historical patterns, enhancing overall predictive accuracy.

In contrast, the CNN model displayed higher error metrics, suggesting it is less suited for this dataset, possibly because it focuses primarily on local patterns and lacks inherent temporal memory compared to recurrent architectures. These results highlight that model architecture should align with the temporal characteristics of financial time series, and hybrid models can leverage complementary strengths to improve prediction. Figure 3 illustrates the comparative performance of all models.

Evaluation models for selected companies Furthermore, an in-depth examination of the models' performance is conducted, with a focus on selecting the top three companies based on their trading value among various companies.

Saudi Basic Industries Crop. The performance evaluation of the models developed for Saudi Basic Industries is summarized in Table 4. The results indicate that the GRU models outperformed the others, with the Hybrid models also showing strong performance. Additionally, Figure 5 depicts the performance comparisons of these models.

Alinma Bank: The evaluation presented in Table 5 shows that the LSTM model demonstrated the most robust performance among the models tested for Alinma Bank, closely followed by the Bi-LSTM model. This superiority is evident across multiple key metrics. Further analysis and visual comparisons of these results to actual values are detailed in Figure 6.

Al Rajhi Bank: Table 6 provides a detailed comparison of different models' performance metrics for Al Rajhi Bank. Among the models tested, the GRU model stands out with the highest R^2 value of 98.76 and the lowest error metrics, suggesting it is the most accurate model for predicting this dataset. The GRU-LSTM model follows closely, demonstrating similarly strong performance. In contrast, the LSTM and Bi-LSTM models show significantly higher error rates and lower R^2 values. Further analysis and graphical comparisons of these results are depicted in Figure 7.

Prediction: In this study, we assessed the effectiveness of advanced predictive models in forecasting the closing stock prices for the next five days, focusing on the top three value-traded companies. The models were applied to previously unseen data starting from January 1, 2024, allowing for a practical evaluation of their predictive performance beyond the training and test sets.

Table 3: Performance comparison of deep learning models in predicting stock closing prices for 131 companies

Models	MSE	RMSE	MAE	MAPE	R^2
LSTM	0.6184	0.7864	0.5015	2.2875	99.96
Bi-LSTM	0.5898	0.7680	0.5354	3.8161	99.96
GRU	0.4595	0.6779	0.4478	3.6962	99.97
CNN	0.7746	0.8801	0.5904	3.2694	99.95
GRU-LSTM	0.7159	0.8461	0.4370	2.2527	99.95
CNN-LSTM	0.6385	0.7991	0.5168	2.5789	99.96

Table 4: Performance comparison of deep learning models for predicting closing prices of Saudi Basic Industries Corporation

Models	MSE	RMSE	MAE	MAPE	R^2
LSTM	0.7439	0.8625	0.6637	0.790	97.22
Bi-LSTM	1.2777	1.1304	1.0302	1.200	95.23
GRU	0.213	0.4615	0.3608	0.420	99.21
CNN	1.6982	1.3032	1.127	1.270	93.66
GRU-LSTM	0.2937	0.5419	0.4381	0.500	98.9
CNN-LSTM	0.3144	0.5607	0.4428	0.510	98.83

DOI: 10.3844/jcssp.2026.1370.1386

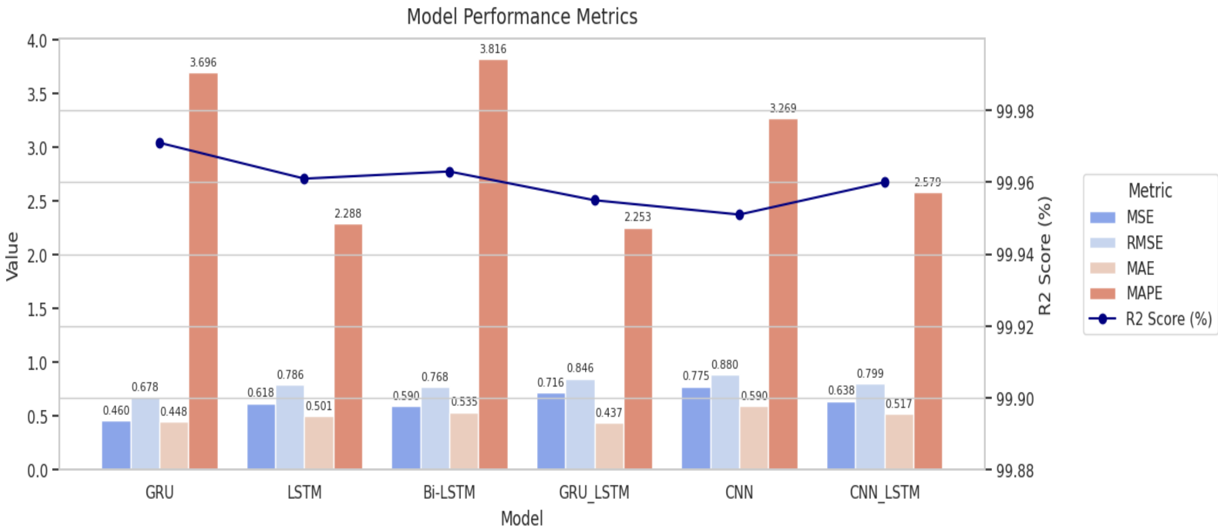

Fig. 4: Performance of models in predicting closing prices of 131 companies

Fig. 5: Performance of models in predicting stock prices of Saudi Basic Industries Corporation

Table 5: Performance comparison of deep learning models for predicting closing prices of Alinma Bank

Models	MSE	RMSE	MAE	MAPE	R^2
LSTM	0.0956	0.3092	0.6637	0.730	98.5
Bi-LSTM	0.1298	0.3603	0.2867	0.880	97.97
GRU	0.2722	0.5217	0.468	1.40	95.73
CNN	0.1979	0.4449	0.3687	1.090	96.9
GRU-LSTM	1.1516	1.0731	0.4381	3.030	81.95
CNN-LSTM	1.2969	1.1388	0.4428	3.180	79.67

Table 6: Performance comparison of deep learning models for predicting closing prices of Al Rajhi Bank

Models	MSE	RMSE	MAE	MAPE	R^2
LSTM	3.224	1.7956	1.7371	2.380	83.26
Bi-LSTM	3.2632	1.8064	1.7218	2.340	83.06
GRU	0.2397	0.4896	0.3831	0.530	98.76
CNN	0.3819	0.618	0.4952	0.670	98.02
GRU-LSTM	0.266	0.5158	0.387	0.520	98.62
CNN-LSTM	0.9088	0.9533	0.7179	0.990	95.28

Fig. 6: Performance of models in predicting stock prices of Alinma Bank

Fig. 7: Performance of models in predicting stock prices of Al Rajhi Bank

This extended analysis provides insight into the models' ability to capture short-term market trends and generate reliable forecasts in a real-world setting. The predicted prices were compared with actual closing prices to assess accuracy, complementing the standard evaluation metrics and demonstrating the robustness and generalizability of the proposed models. Table 7 (a, b, c) presents sample results for three companies where the model was applied to forecast the next five trading days: January 1, 2, 3, 4, and 7, 2024.

DOI: 10.3844/jcssp.2026.1370.1386

Table 8 presents the results of applying various deep learning models to forecast stock prices for the top three companies on previously unseen data. The GRU-LSTM model consistently delivers the most accurate predictions, as indicated by its lower error metrics, while the GRU and LSTM models also show strong performance. Other models exhibit varied results, generally achieving moderate accuracy. This section demonstrates how the models perform beyond the standard training and test sets, providing practical validation of their generalization capability and robustness in real-world prediction scenarios.

Table 7: Predicted closing prices for the next five days for three companies using unseen data

(a) Saudi Basic Industries Corp

Actual Close prices	LSTM	Bi-LSTM	GRU	GRU -LSTM	CNN	CNN -LSTM
83.7	83.32	84.64	83.42	84.30	85.35	84.21
83.3	83.05	84.46	83.24	83.94	85.21	84.29
82	82.19	83.50	82.17	82.78	84.38	83.33
80.6	80.18	81.29	79.87	80.05	81.50	80.93
81.5	81.19	82.19	80.75	81.39	82.72	81.65

(b) Alinma Bank

Actual Close prices	LSTM	Bi-LSTM	GRU	GRU -LSTM	CNN	CNN -LSTM
39.9	39.92	40.47	40.12	40.14	39.27	40.16
39.5	40.12	40.10	40.15	40.06	39.38	40.26
38.5	39.21	39.09	39.28	39.17	38.72	39.38
39.85	39.49	39.76	39.67	39.64	38.93	39.14
41.8	41.25	42.43	41.52	41.97	40.14	41.27

(c) Al Rajhi Bank

Actual Close prices	LSTM	Bi-LSTM	GRU	GRU -LSTM	CNN	CNN -LSTM
86.8	86.54	88.31	87.09	87.04	87.96	87.32
86.9	85.79	87.60	86.42	86.36	87.48	86.71
84.7	85.02	86.83	85.56	85.62	86.66	85.94
87.5	85.60	87.48	86.05	85.90	86.66	86.02
88.8	89.10	90.78	89.78	88.79	89.46	88.93

Table 8: Evaluation of the proposed models applied for each Company

Company	Models	MSE	RMSE	MAE	MAPE
Saudi Basic Industries Corp.	LSTM	0.103	0.321	0.310	0.377
	Bi-LSTM	1.086	1.042	0.996	1.210
	GRU	0.241	0.491	0.398	0.488
	GRU-LSTM	0.339	0.582	0.536	0.651
	CNN	2.867	1.693	1.612	1.956
	CNN-LSTM	0.628	0.793	0.662	0.803
Alinma Bank	LSTM	0.264	0.514	0.452	1.137
	Bi-LSTM	0.288	0.536	0.496	1.243
	GRU	0.238	0.488	0.422	1.069
	GRU-LSTM	0.179	0.423	0.370	0.939
	CNN	0.812	0.901	0.710	1.747
	CNN-LSTM	0.441	0.664	0.628	1.582
Al Rajhi Bank	LSTM	1.020	1.010	0.778	0.893
	Bi-LSTM	2.246	1.499	1.268	1.462
	GRU	0.823	0.907	0.812	0.933
	GRU-LSTM	0.751	0.867	0.662	0.765
	CNN	1.333	1.155	1.040	1.204
	CNN-LSTM	0.810	0.900	0.712	0.824

Comparing With Previous Studies

Table 9 provides a qualitative comparison between the proposed study and selected related works. The compared studies differ significantly in terms of stock markets, data scales, prediction targets, and evaluation metrics. As a result, direct numerical comparison of performance metrics is not meaningful. Instead, this table highlights methodological aspects such as the types of models employed, target markets, and evaluation criteria. This comparison demonstrates that while similar deep learning architectures have been explored in developed and emerging markets, this study extends prior work by conducting a large-scale evaluation of individual and hybrid deep learning models on Saudi stock market.

Table 9: Assessment of the proposed models in relation to prior Studies

Reference	Dataset	Model	Evaluation Metrics
Proposed Model	The Saudi stock exchange, Tadawul (Large-scale evaluation on 131 companies)	LSTM, Bi-LSTM, GRU, CNN, LSTM-GRU, and CNN-LSTM	MSE, RMSE, MAE, MAPE, R^2
Zaheer et al., 2023	The Shanghai Composite Index, 5951 trading days	CNN, LSTM, RNN, CNN-LSTM, and CNN-RNN	MSE, RMSE, MAE, MAPE
Alkhatib et al., 2022	Data from four firms representing different industry sectors-Tesla, Apple, Snapchat, and ExxonMobil were used	MLP,CNN, GRU, LSTM, Bi-LSTM, and CNN-LSTM	MSE, MAPE

Conclusion

Predicting stock prices is challenging due to the complexity and nonlinearity of financial time series data. This study applied deep learning models, including LSTM, Bi-LSTM, GRU, CNN, and hybrid models (GRU-LSTM, CNN-LSTM), to forecast closing prices of 131 Saudi companies from 2013 to 2023. The methodology comprised six stages: Data collection, exploration, preprocessing (handling missing values, feature selection, normalization, train-test split), model implementation, evaluation (MSE, RMSE, MAE, MAPE, R^2), and prediction on selected key companies. The results show that GRU and the hybrid GRU-LSTM models outperformed the other approaches in forecasting stock prices, achieving the lowest error metrics and highest accuracy. LSTM and CNN-LSTM also performed well, while Bi-LSTM and CNN showed comparatively lower performance. This study demonstrates that deep learning and hybrid models can effectively predict stock prices by capturing subtle patterns in complex financial data. The findings advance knowledge of the Saudi stock market and provide practical insights to support informed investment decisions. Despite the strong predictive performance achieved by the proposed deep learning models, this study has limitations. First, stock market prediction inherently faces challenges such as market non-stationarity, sudden shocks, and the influence of external factors such macroeconomic indicators and news sentiment, which were not incorporated into the models. Second, the models were evaluated within a single market context, which may limit the generalizability of the findings to other stock markets with different characteristics. Future work could explore the impact of economic, social, and political factors on stock prices and develop models for real-time, tick-level data to support automated trading. Such studies would provide valuable insights for investors while addressing the inherent unpredictability of the Saudi stock market.

Acknowledgment

The authors would like to thank the Saudi Exchange (Tadawul) for providing access to the stock market data used in this study. No external funding was received for this work.

Funding Information

This research received no external funding.

Author’s Contributions

Areej Alhumaidi: Conceptualization of the study, methodology, data collection, analysis, and drafted of the manuscript.

Hoda Abdelhafez: Contributed to the study design, provided academic guidance, and was involved in the critical review and edited of the manuscript.

Ethics

The authors commit to promptly addressing any ethical issues that may arise after publication in accordance with journal policies.

References

Aldhyani, T. H. H., & Alzahrani, A. (2022). Framework for Predicting and Modeling Stock Market Prices Based on Deep Learning Algorithms. *Electronics*, *11*(19), 1–20. https://doi.org/10.3390/electronics11193149

Alkhatib, K., Khazaleh, H., Alkhazaleh, H. A., Alsoud, A. R., & Abualigah, L. (2022). A New Stock Price Forecasting Method Using Active Deep Learning Approach. *Journal of Open Innovation: Technology, Market, and Complexity*, *8*(2), 1–16. https://doi.org/10.3390/joitmc8020096

Al-Nefaie, A. H., & Aldhyani, T. H. H. (2022). Predicting Close Price in Emerging Saudi Stock Exchange: Time Series Models. *Electronics*, *11*(21), 3443. https://doi.org/10.3390/electronics11213443

Aseeri, A. O. (2023). Effective short-term forecasts of Saudi stock price trends using technical indicators and large-scale multivariate time series. *PeerJ Computer Science*, *9*, e1205. https://doi.org/10.7717/peerj-cs.1205

Baru, C. (2019). Institute of Electrical and Electronics Engineers. *and IEEE Computer Society, 2019 IEEE International Conference on Big Data: Proceedings: Dec 9*, 1–17.

Barua, M., Kumar, T., Raj, K., & Roy, A. K. (2024). Comparative Analysis of Deep Learning Models for Stock Price Prediction in the Indian Market. *FinTech*, *3*(4), 551–568. https://doi.org/10.3390/fintech3040029

Buyrukoğlu, S., & Akbaş, A. (2022). Machine Learning based Early Prediction of Type 2 Diabetes: A New Hybrid Feature Selection Approach using Correlation Matrix with Heatmap and SFS. *Balkan Journal of Electrical and Computer Engineering*, *10*(2), 110–117. https://doi.org/10.17694/bajece.973129

Faraz, M., Khaloozadeh, H., & Abbasi, M. (2020). Stock Market Prediction-by-Prediction Based on Autoencoder Long Short-Term Memory Networks. *2020 28th Iranian Conference on Electrical Engineering (ICEE)*, 1-c5. https://doi.org/10.1109/icee50131.2020.9261055

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

Gandhmal, D. P., & Kumar, K. (2019). Systematic analysis and review of stock market prediction techniques. *Computer Science Review*, *34*, 100190. https://doi.org/10.1016/j.cosrev.2019.08.001

Göçken, M., Özçalıcı, M., Boru, A., & Dosdoğru, A. T. (2016). Integrating metaheuristics and Artificial Neural Networks for improved stock price prediction. *Expert Systems with Applications*, *44*, 320–331. https://doi.org/10.1016/j.eswa.2015.09.029

Hamayel, M. J., & Owda, A. Y. (2021). A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms. *AI*, *2*(4), 477–496. https://doi.org/10.3390/ai2040030

Ilyas, Q. M., Iqbal, K., Ijaz, S., Mehmood, A., & Bhatia, S. (2022). A Hybrid Model to Predict Stock Closing Price Using Novel Features and a Fully Modified Hodrick–Prescott Filter. *Electronics*, *11*(21), 3588. https://doi.org/10.3390/electronics11213588

Jarrah, M., & Salim, N. (2019). A Recurrent Neural Network and a Discrete Wavelet Transform to Predict the Saudi Stock Price Trends. *International Journal of Advanced Computer Science and Applications*, *10*(4), 140–146. https://doi.org/10.14569/ijacsa.2019.0100418

Jingyi, Z., Qinwen, Z., Yanjun, J., & Heng-Tze, C. (2022). Stock Market Prediction via Deep Learning Techniques: A Survey. *ArXiv (Cornell University)*, *7*(1), 1–33. https://doi.org/10.48550/arXiv.2212.12717

Karim, M. E., Foysal, M., & Das, S. (2023). Stock Price Prediction Using Bi-LSTM and GRU-Based Hybrid Deep Learning Approach. *Springer Link / Springer Nature*, *493*, 701–711. https://doi.org/10.1007/978-981-19-3148-2_60

Kumar, D., Sarangi, P. K., & Verma, R. (2022). A systematic review of stock market prediction using machine learning and statistical techniques. *Materials Today: Proceedings*, *49*, 3187–3191. https://doi.org/10.1016/j.matpr.2020.11.399

Medar, R., Rajpurohit, V. S., & Rashmi, B. (2017). Impact of Training and Testing Data Splits on Accuracy of Time Series Forecasting in Machine Learning. *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, 1–6. https://doi.org/10.1109/iccubea.2017.8463779

Moghar, A., & Hamiche, M. (2020). Stock Market Prediction Using LSTM Recurrent Neural Network. *Procedia Computer Science*, *170*, 1168–1173. https://doi.org/10.1016/j.procs.2020.03.049

Muhammad, D., Ahmed, I., Naveed, K., & Bendechache, M. (2024). An explainable deep learning approach for stock market trend prediction. *Heliyon*, *10*(21), e40095. https://doi.org/10.1016/j.heliyon.2024.e40095

Nayak, S. K., Misra, B. B., & Behera, H. S. (2014). Impact of Data Normalization on Stock Index Forecasting. *International Journal of Computer Information Systems and Industrial Management Applications*, *6*, 257–269.

Obthong, M., Tantisantiwong, N., Jeamwatthanachai, W., & Wills, G. (2020). A Survey on Machine Learning for Stock Price Prediction: Algorithms and Techniques. *Proceedings of the 2nd International Conference on Finance, Economics, Management and IT Business*, 63–71. https://doi.org/10.5220/0009340700630071

Otchere, D. A., Ganat, T. O. A., Ojero, J. O., Tackie-Otoo, B. N., & Taki, M. Y. (2022). Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions. *Journal of Petroleum Science and Engineering*, *208*, 109244. https://doi.org/10.1016/j.petrol.2021.109244

Padh Ghosh, B., Shafiquzzaman Bhuiyan, M., Das, D., Ngoc Nguyen, T., Jewel, M., Mia, M. T., & Cao, D. (2024). Deep Learning in Stock Market Forecasting: Comparative Analysis of Neural Network Architectures Across NSE and NYSE. *Journal of Computer Science and Technology Studies*, *6*(1), 68–75. https://doi.org/10.32996/jcsts.2024.6.1.8

Patra, G. R., & Mohanty, M. N. (2022). An LSTM-GRU based hybrid framework for secured stock price prediction. *Journal of Statistics and Management Systems*, *25*(6), 1491–1499. https://doi.org/10.1080/09720510.2022.2092263

Sako, K., Mpinda, B. N., & Rodrigues, P. C. (2022). Neural Networks for Financial Time Series Forecasting. *Entropy*, *24*(5), 657. https://doi.org/10.3390/e24050657

Saud, A. S., & Shakya, S. (2020). Analysis of look back period for stock price prediction with RNN variants: A case study on banking sector of NEPSE. *Procedia Computer Science*, *167*, 788–798. https://doi.org/10.1016/j.procs.2020.03.419

Saboor, A., Hussain, A., Lord Y. Agbley, B., ul Haq, A., Ping Li, J., & Kumar, R. (2023). Stock Market Index Prediction Using Machine Learning and Deep Learning Techniques. *Intelligent Automation & Soft Computing*, *37*(2), 1325–1344. https://doi.org/10.32604/iasc.2023.038849

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning : A systematic literature review: 2005–2019. *Applied Soft Computing*, *90*, 106181. https://doi.org/10.1016/j.asoc.2020.106181

Song, H., & Choi, H. (2023). Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models. *Applied Sciences*, *13*(7), 4644. https://doi.org/10.3390/app13074644

Zaheer, S., Anjum, N., Hussain, S., Algarni, A. D., Iqbal, J., Bourouis, S., & Ullah, S. S. (2023). A Multi Parameter Forecasting for Stock Time Series Data Using LSTM and Deep Learning Model. *Mathematics*, *11*(3), 590. https://doi.org/10.3390/math11030590