

## Article

# Assessing the Predictive Power of Transformers, ARIMA, and LSTM in Forecasting Stock Prices of Moroccan Credit Companies

Karima Lahboub \*  and Mimoun Benali

Laboratory of Research and Studies in Management, Entrepreneurship and Finance (LAREMEF), Nation School of Commerce and Management of Fez, Sidi Mohamed Ben Abdellah University, Fes 30050, Morocco; mimoun.benali@usmba.ac.ma

\* Correspondence: lahoubkarima.encg@gmail.com

**Abstract:** In this paper, we present a data-driven approach to forecasting stock prices in the Moroccan Stock Exchange. Our study tests three predictive models: ARIMA, LSTM, and transformers, applied to the historical stock price data of three prominent credit companies (EQD, LES, and SLF) listed on the Casablanca Stock Exchange. We carefully selected and optimized hyperparameters for each model to achieve optimal performance. Our results showed that the LSTM model achieved high accuracy, with R-squared values exceeding 0.99 for EQD and LES and surpassing 0.95 for SLF. These findings highlighted the effectiveness of LSTM in stock price forecasting. Our study offers practical insights for traders and investors in the Moroccan Stock Exchange, demonstrating how predictive modeling can aid in making informed decisions. This research contributes to advancing stock market forecasting in Morocco, providing valuable tools for navigating the Casablanca Stock Exchange.

**Keywords:** stock market; Moroccan Stock Exchange; machine learning; LSTM; transformer



**Citation:** Lahboub, Karima, and Mimoun Benali. 2024. Assessing the Predictive Power of Transformers, ARIMA, and LSTM in Forecasting Stock Prices of Moroccan Credit Companies. *Journal of Risk and Financial Management* 17: 293. <https://doi.org/10.3390/jrfm17070293>

Academic Editor: Gulasekaran Rajaguru

Received: 22 May 2024

Revised: 29 June 2024

Accepted: 5 July 2024

Published: 9 July 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The stock market operates as a dynamic commercial arena, where buying and selling transactions occur. Unlike conventional markets, what sets it apart is the steadfast nature of its regulatory framework, which remains consistent irrespective of contractual changes. Tradable assets on the stock exchange must possess significant economic value, be storable, and offer public utility, distinguishing them from commodities traded in other markets. Transactions within the stock exchange are exclusively facilitated by licensed and registered brokerage firms and intermediaries, adhering to pertinent laws and regulations (Ahmed and Huo 2021). Among the diverse realms of stock markets, we find the financial market, encompassing staples such as wheat, sugar, and corn, alongside pivotal currencies, such as the US (United States) dollar, Japanese yen, Euro, Swiss franc, Canadian dollar, Australian dollar, and New Zealand dollar. Furthermore, there exists the stock market, the bond market, and the pivotal market for raw materials, including oil, copper, and cotton (Løkken and Aas 2020). Leading the global arena are eminent exchanges, such as the New York Stock Exchange in the United States of America, the venerable London Stock Exchange in England, the dynamic Frankfurt Stock Exchange in Germany, the Tokyo Stock Exchange in Japan, the vibrant Sydney Stock Exchange in Australia, and the bustling Hong Kong Stock Exchange in Hong Kong (Ma et al. 2016; Kuvshinov and Zimmermann 2022).

In Morocco, the Casablanca Stock Exchange serves as a burgeoning financial hub boasting over 75 Moroccan companies spanning various sectors, including energy, food, and pharmaceuticals (Zaimi 2022). The Casablanca Stock Exchange holds a notable global standing, ranking among the top 30 stock exchanges worldwide and securing a position among the three most robust in the Arab world (Azzam 2015). Remarkably, it clinches the title of the foremost exchange in Africa, as affirmed by the esteemed British agency, “ZYN” (Dibiah and Mojekwu 2023). By the end of 2023, the MASI (Moroccan All Shares Index)

concluded at an impressive 12,000 points, reflecting a market value exceeding 626 billion dirhams for the same period (Baali et al. 2023).

To entice investors to engage with the Casablanca Stock Exchange, offering them tailored insights and predictive models that cater to their investment strategies is essential. One effective approach involves analyzing the performance of various sectors within the Moroccan market, leveraging historical data and sophisticated statistical models to provide valuable foresight into the trajectory of listed companies. Furthermore, the development of predictive models to forecast the closing values of key stock market indices, such as the MASI, serves as a crucial tool for informed investment decision-making. By delving into market trends, economic indicators, and geopolitical developments, investors can make well-informed choices regarding the timing of their investment activities (Nti et al. 2020).

In essence, providing investors with comprehensive and adaptable analytical tools not only enhances their confidence in the Casablanca Stock Exchange but also empowers them to navigate the complexities of the market with greater precision and efficiency.

Indeed, the field of data science, as a subset of artificial intelligence, offers a performant algorithm specifically designed to analyze stock market data and generate predictive insights (Nosratabadi et al. 2020). These algorithms use advanced statistical techniques, called machine learning models and deep learning architectures, to extract valuable patterns and trends from vast datasets.

By leveraging these algorithms, investors can gain deeper insights into various aspects of the stock market, including price movements, trading volumes, market sentiment, and volatility (Kompella and Chakravarthy Chilukuri 2020). This enables them to make more informed decisions regarding investment opportunities, whether it involves selecting the best-performing companies or predicting the future performance of key market indices.

Moreover, data science algorithms can also assist in identifying hidden correlations and dependencies within the data, uncovering potential opportunities for arbitrage or risk mitigation strategies. Additionally, they can help in optimizing portfolio allocation and asset allocation strategies based on investors' risk preferences and investment objectives (Bhowmik and Wang 2020).

Overall, the application of data science algorithms in stock market analysis not only enhances the efficiency and accuracy of investment decision-making but also contributes to the development of innovative investment strategies and approaches. As such, it plays a crucial role in empowering investors to navigate the complexities of the stock market with greater confidence and success (Brière et al. 2022).

Machine learning (ML) algorithms play a crucial role in predicting future outcomes, such as in the stock market. One major category of machine learning is supervised learning, where algorithms learn from labeled data to make predictions (Dridi 2021). In the stock market context, supervised learning algorithms have evolved over time. Initially, linear algorithms were prominent, relying on linear functions to model problems. Examples include linear regression, lasso, and support vector machines (SVM). These algorithms provided a foundational understanding but had limitations in capturing complex patterns (Mishra and Padhy 2019).

Later, decision tree algorithms emerged, marking a significant advancement in machine learning. Decision trees (DT) revolutionized the field by offering competitive performance across various sectors (Zhang et al. 2022). Techniques such as bagging and boosting, which combine multiple decision trees, led to the development of ensemble algorithms, such as XGBoost, random forest, and LightGBM (Mohammad 2023). These algorithms excel in handling nonlinear relationships and capturing intricate patterns in data. Furthermore, the advent of deep learning (DL) introduced neural networks capable of learning intricate patterns from data. Recurrent neural networks (RNNs) specialize in sequential data (Chen et al. 2021), making them suitable for time series analysis in the stock market. Recently, transformers have emerged as the latest generation of algorithms, leveraging attention mechanisms to capture long-range dependencies and improve model performance (Khan et al. 2022). Each generation of algorithms builds upon the previous ones, incorporating

advancements in computational power, data availability, and algorithmic techniques. By leveraging these algorithms, investors can better analyze market trends, identify profitable opportunities, and make informed investment decisions.

In exploring the realm of stock price forecasting, the comparative studies discussed exhibit a diverse array of methodologies and datasets, each offering unique insights into the predictive capabilities of various models. While the study of [Prasad et al. \(2022\)](#), delved into the comparison between ARIMA (Autoregressive Integrated Moving Average) and SARIMAX (Seasonal Autoregressive Integrated Moving Average with Exogenous Regressors) models, utilizing data sourced from Yahoo Finance, its emphasis on closing values underscores the significance of accurate forecasting for investors in navigating the complexities of stock markets. Conversely, the study of [Low and Sakk \(2023\)](#) broadened the scope by evaluating ARIMA and LSTM (Long Short-Term Memory) models across ten different stock tickers, demonstrating the versatility of ARIMA in making precise point predictions of closing prices for exchange-traded funds. Moving to the study of [Wahyudi \(2017\)](#) the focus shifted to the Indonesia CSPI (Composite Stock Price Index) dataset, where an in-depth examination of different ARIMA models revealed the effectiveness of ARIMA (0,1,1) in capturing the daily movements of stock prices. Also, the study of [Pulungan et al. \(2018\)](#) extended the discussion to the impact of ARIMA (3,1,1) on the SRI-KEHATI (Sustainable and Responsible Investment (SRI)-KEHATI) Index, shedding light on the intricate relationship between socially responsible investment and market dynamics. Together, these studies underscore the pivotal role of forecasting models in empowering stakeholders with actionable insights, ultimately contributing to informed decision-making within the realm of financial markets.

Furthermore, various studies have assessed the performance of predictive models over different forecast horizons. For example, [Patel et al. \(2015\)](#) compared ANN (artificial neural network), SVM, random forest, and naïve Bayes models for short-term (1 day ahead) and medium-term (7 days ahead) forecasts using Indian Stock Market data. Their findings showed that the random forest model outperformed others, with an average accuracy of 83.59%, highlighting its robustness in capturing stock market trends across different horizons. Similarly, [Ballings et al. \(2015\)](#) benchmarked ensemble methods (random forest, AdaBoost, and kernel factory) against single classifier models (neural networks, logistic regression, SVM, and K-nearest neighbor) using data from 5767 European companies to predict stock price direction one year ahead. Random forest emerged as the top-performing algorithm, with the highest mean AUC (area under the ROC) ranking (1.0) and the lowest interquartile range (0.0061), followed by SVM and kernel factory.

At the deep learning level, [Wu et al. \(2023\)](#) delved into the comparison between SACLSTM (Self-Attentive Convolutional Long Short-Term Memory), SVM, CNN (convolutional neural networks), and ANN models, using data from ten stocks in the American and Taiwan markets to predict the direction of the stock market. Emphasizing historical data, futures, and options as input features, the study evaluated accuracy as its metric, revealing SACLSTM's relatively superior performance compared to other models. Meanwhile, [Wang et al. \(2021\)](#) focused on BiSLSTM (Bidirectional Sequence-to-Sequence Long Short-Term Memory) against MLP, RNN, LSTM, BiLSTM, CNN-LSTM, and CNN-BiLSTM models, using the Shenzhen Component Index data to forecast closing prices. With a comprehensive set of input features and metrics, including MAE (mean absolute error), RMSE (root mean square error), and  $R^2$  (coefficient of determination), CNN-BiSLSTM emerged as the optimal performer with superior values. Transitioning to the study of [Lu et al. \(2021\)](#), which explored CNN-BiLSTM-AM (attention mechanism) among MLP, CNN, RNN, LSTM, BiLSTM, CNN-LSTM, and other variants, the focus shifted to predicting the next day's stock closing price in the Shanghai Composite Index. With a similar set of input features and metrics, CNN-BiLSTM-AM yielded the best results, demonstrating its robust predictive capability.

Additionally, [Bao et al. \(2017\)](#) presented a novel deep learning framework combining wavelet transforms (WT), stacked autoencoders (SAEs), and LSTM for stock price forecasting. Their results demonstrated that the WSAEs-LSTM model significantly outperformed

other models in predictive accuracy and profitability for 1-day- and 5-days-ahead forecasts, achieving a MAPE of 0.019 and Theil U of 0.013 for the CSI 300 index, with an R-value of 0.944. Similarly, Li et al. (2018) introduced an attention-based multi-input LSTM (MI-LSTM) model capable of extracting valuable information from low-correlated factors. Their experimental results on China's Stock Market data showed that the MI-LSTM model achieved superior performance in profit comparison, particularly in 1-day- to 5-days-ahead forecasts, significantly outperforming standard LSTM models and the CSI 300 index. These findings further underscore the effectiveness of advanced LSTM variants in providing accurate short-term stock price predictions.

By examining the results from various studies, it becomes evident that LSTM models consistently outperformed ARIMA and linear models in most cases. For instance, in the study of (Wu et al. 2023), SACLSTM demonstrated relatively superior performance compared to SVM, CNN-cor, CNNpred, and ANN models in predicting the direction of the stock market. In the study of (H. Wang et al. 2021), CNN-BiLSTM achieved optimal values for metrics such as MAE, RMSE, and  $R^2$ , indicating its effectiveness in forecasting closing prices using the Shenzhen Component Index data. Similarly, the study in (Wang 2023) showcased the superior performance of CNN-BiLSTM-AM in predicting next-day stock closing prices in the Shanghai Composite Index dataset, as evidenced by its impressive MAE, RMSE, and  $R^2$  values. The same was seen for CNN-BiLSTM-ECA and BiLSTM-MTRAN-TCN models, which emerged as the superior performers in predicting next-day closing prices across multiple indices, underscoring the robustness of LSTM architectures in stock market prediction tasks. Collectively, these findings suggest that LSTM models offer superior predictive capabilities compared to ARIMA and linear models, making them a preferred choice for stock price forecasting tasks.

In this paper, our focus lies on the Moroccan Stock Exchange dataset, particularly within the consumer credit sector. This sector encompasses four prominent companies integrated into the Casablanca Stock Exchange. Our objective is to predict the closing prices of each company using three distinct methodologies: ARIMA, LSTM, and transformers. While ARIMA offers a traditional approach to time series forecasting, LSTM leverages recurrent neural networks for sequential data prediction. Furthermore, we introduce transformers as a novel concept for predicting sequential data, exploring its potential in the domain of stock price forecasting. Through this comparative analysis, we aim to discern the strengths and limitations of each approach and provide insights into their effectiveness in predicting stock prices within the Moroccan consumer credit sector.

In the subsequent sections of our study, we will detail the materials and methods employed, encompassing data analysis techniques and the approach undertaken. Following this, we will present the obtained results along with their discussion, elucidating any significant findings and their implications. Finally, we will draw conclusions based on our analysis, summarizing the key insights and potential implications for future research and practical applications.

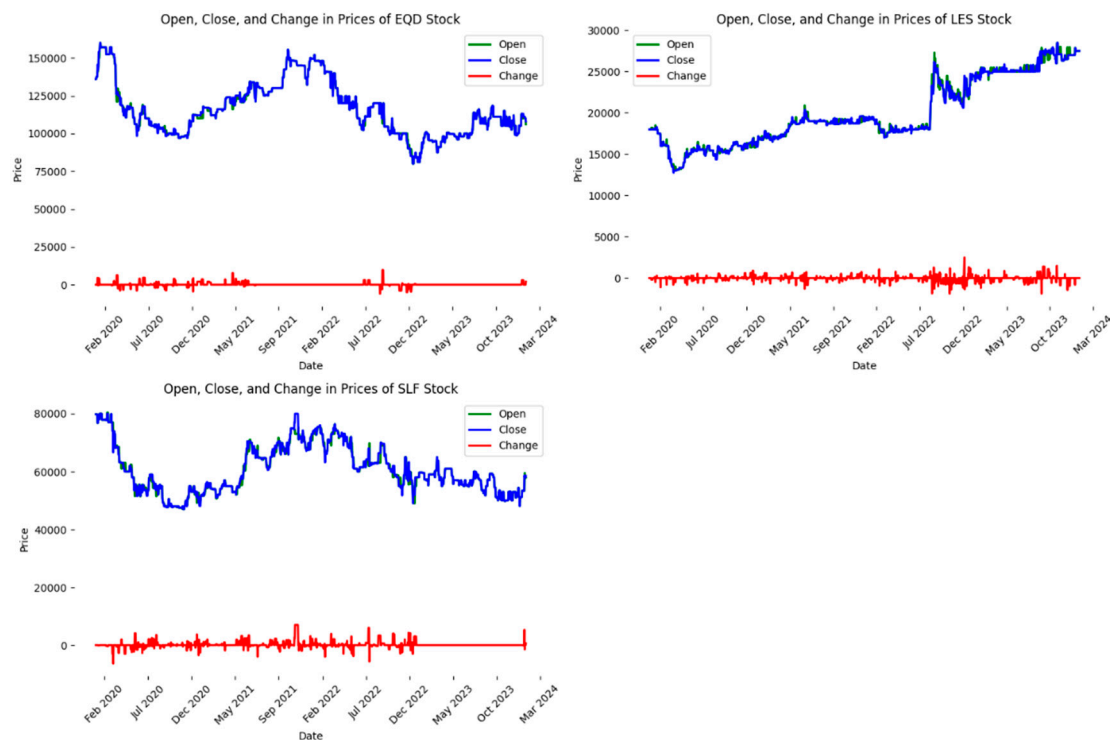
## 2. Materials and Methods

### 2.1. Data

In this study, we analyzed financial data from three prominent credit companies in Morocco: EQD, SLF, and LES. These companies have been integrated into the Casablanca Stock Market since 2020. Our dataset spans from January 2020 to March 2024, encompassing both the opening and closing prices of these companies' stocks over this period.

Figure 1 illustrates the distribution of open and close prices, as well as the change between them, for the three companies throughout the study period. The graphs depict a highly heterogeneous pattern in stock price changes, suggesting significant variability and potential opportunities for analysis. The observed heterogeneity in stock price changes indicates the presence of interesting dynamics that can be leveraged to develop predictive models. This variability provides an opportunity to explore the underlying factors driving stock price movements and to construct robust forecasting models. By using this

dataset, we aim to uncover insights that can inform investment strategies and enhance our understanding of the dynamics within the Moroccan credit market. Overall, the observed heterogeneity in stock price changes underscores the potential for developing predictive models and highlights the significance of our dataset in uncovering valuable insights for investors and researchers alike.



**Figure 1.** Comparison of open and close prices.

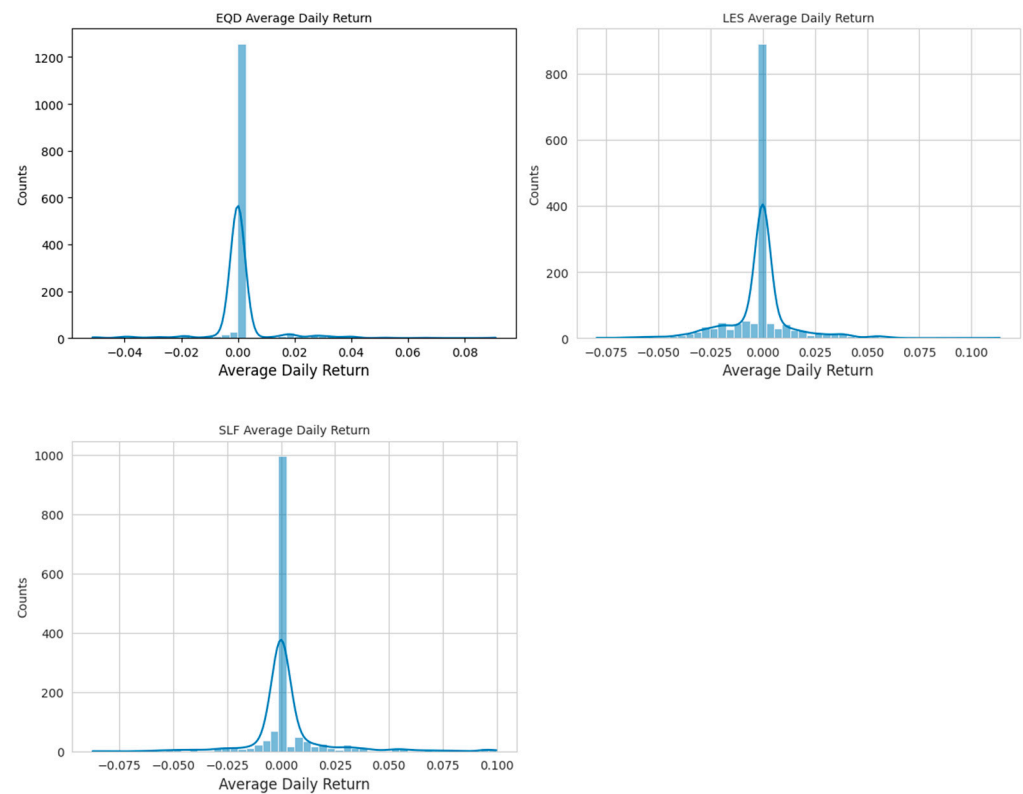
To further analyze the data, Figure 2 displays the average return for each day for the three companies. As depicted in the figure, there were important fluctuations in the daily returns of companies 2 and 3 (LES and SLF) across both positive and negative axes. This indicates substantial volatility in their returns, with notable variations in both upward and downward directions. Additionally, for the first company (EQD), there were discernible changes within certain days, albeit less pronounced compared to LES and SLF.

One common observation across all three companies was that the magnitude of return changes tended to be relatively small, as indicated by the proximity of the returns to 0. This suggests that while there may be fluctuations in the daily returns, they were generally not substantial. Overall, the analysis of average returns provides valuable insights into the volatility and stability of the companies' stocks.

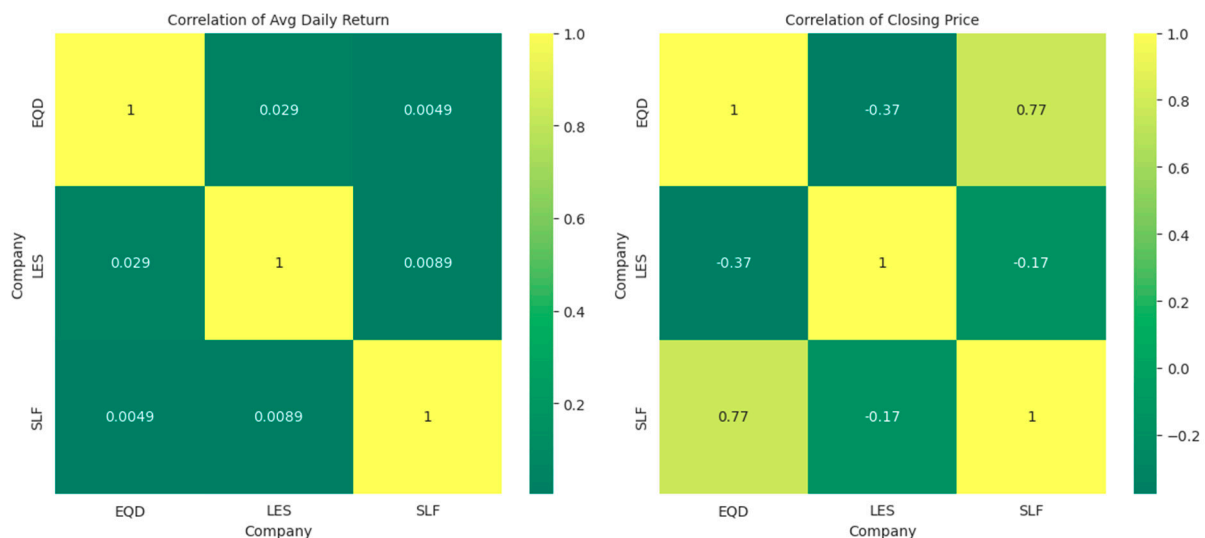
The pronounced fluctuations in returns for LES and SLF, coupled with comparatively smaller changes for EQD, highlight the diverse nature of their performance and the potential for further investigation into the underlying factors driving these fluctuations.

After analyzing the returns of the three companies, it became evident that there existed a correlation among them, characterized by relatively small yet notable changes in daily returns. To validate this observation, we conducted further analysis by examining the correlation between the returns of the three companies. Figure 3 presents two correlation matrices: one for the delayed returns and the other for the closing prices.





**Figure 2.** Average daily returns.



**Figure 3.** Correlation between companies.

Upon examination, we observed a significant correlation between the closing prices of EQD and SLF, exceeding 77% ( $p$ -value = 0.002). This finding is not surprising, considering that the companies operate within the same sector, namely consumer credit. However, when analyzing the delayed returns, the correlation between LES and SLF was found to be merely 0.029 ( $p$ -value = 0.82). This demonstrates a noteworthy divergence from the anticipated correlation, suggesting that there may be limited synchronicity in the performance of LES and SLF in terms of their delayed returns. Furthermore, the modest correlation observed among the three companies in delayed returns indicates a lack of interrelation in the consumer credit sector. This implies that improvements or deteriorations in one

company's performance do not necessarily coincide with those of the others, underscoring the independent nature of their operations.

In summary, the analysis of the correlation matrices revealed distinct patterns in the relationships between the companies' returns and closing prices. While a strong correlation existed between the closing prices of EQD and SLF, suggesting sector-related coherence, the limited correlation in delayed returns between LES and SLF implies a degree of independence in their performance within the consumer credit sector.

## 2.2. Developed Models

In this study, we conducted a comprehensive benchmarking analysis of three prominent models in the field of sequential data analysis. The first model under examination was the ARIMA model (Wahyudi 2017), a classical time series forecasting technique widely used for its simplicity and effectiveness in capturing linear relationships within sequential data. The ARIMA model operates by decomposing the time series data into trend, seasonality, and residual components, and then using autoregression and moving-average components to model the data's behavior over time.

The ARIMA model is represented mathematically in Equation (1):

$$Y_t = c + \phi_1 * Y_{t-1} + \phi_2 * Y_{t-2} + \dots + \phi_p * Y_{t-p} + \theta_1 * \epsilon_{t-1} + \theta_2 * \epsilon_{t-2} + \dots + \theta_q * \epsilon_{t-q} + \epsilon_t \quad (1)$$

where:

$Y_t$  is the value of a stationary time series at time  $t$ .

$c$  is the constant term or intercept.

$\phi_1, \phi_2, \dots, \phi_p$  are the autoregressive coefficients.

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$  are the lagged values of the time series.

$\theta_1, \theta_2, \dots, \theta_q$  are the moving-average coefficients.

$\epsilon_t$  is the error term or white noise at time  $t$ .

$\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-q}$  are the lagged values of the error term.

To ensure that the time series was stationary, we conducted the Augmented Dickey–Fuller (ADF) test. The ADF values and their corresponding  $p$ -values for each of the price series are as follows:

EQD (Company 1): ADF value (differenced) =  $-41.219$ ,  $p$ -value (differenced) =  $0.0$

LES (Company 2): ADF value (differenced) =  $-20.189$ ,  $p$ -value (differenced) =  $0.0$

SLF (Company 3): ADF value (differenced) =  $-40.777$ ,  $p$ -value (differenced) =  $0.0$

These results confirmed that the differenced price series for each company are stationary, as indicated by the significant  $p$ -values (less than  $0.05$ ). Stationarity is crucial in time series analysis, as it ensures that the statistical properties of the series, such as mean and variance, remain constant over time, facilitating more reliable modeling and forecasting.

The second model was a recurrent neural network (RNN), specifically an LSTM (Long Short-Term Memory) network. Unlike traditional statistical models, such as ARIMA, LSTM networks are capable of capturing long-term dependencies and nonlinear relationships within sequential data. LSTM networks feature recurrent connections that enable them to retain memories of past information, making them well suited for time series forecasting tasks (Fang et al. 2021). At the core of an LSTM network are memory cells, which are equipped with mechanisms to selectively remember or forget information over time. This ability to retain information over long sequences enables LSTM networks to capture long-term dependencies in sequential data.

One distinctive feature of LSTM cells is the presence of gates, which regulate the flow of information within the network. The first one is called the Forget Gate, which controls the extent to which the previous cell state should be retained or forgotten. It takes as input the previous cell state ( $C_{t-1}$ ) and the current input ( $x_t$ ), and outputs a Forget Gate vector ( $f_t$ ) with values between  $0$  and  $1$ . A value of  $1$  indicates that the corresponding element in the cell state should be retained, while a value of  $0$  indicates that it should be forgotten. The second gate is called the Input Gate, which is responsible of determining the extent

to which new information should be incorporated into the cell state. It takes as input the previous cell state ( $C_{t-1}$ ) and the current input ( $X_t$ ), and outputs an Input Gate vector ( $I_t$ ) with values between 0 and 1. This gate controls the update of the cell state by modulating the contributions of the new input and the previous cell state.

The next gate is the Output Gate, which determines the extent to which the current cell state should influence the output. It takes as input the previous cell state ( $C_{t-1}$ ), the current input ( $X_t$ ), and the current cell state ( $C_t$ ), and outputs an Output Gate vector ( $O_t$ ) with values between 0 and 1. This gate regulates the output of the LSTM cell by controlling the information flow from the cell state to the output. These gates are equipped with activation functions, typically sigmoid functions, that squish the gate values between 0 and 1, allowing for fine-grained control over the information flow.

In Figure 4, we illustrate the architecture of an LSTM cell, highlighting the key components discussed above. This visualization provides a clear understanding of how the Forget Gate, Input Gate, and Output Gate interact within the LSTM cell to regulate the flow of information and retain memory over long sequences.

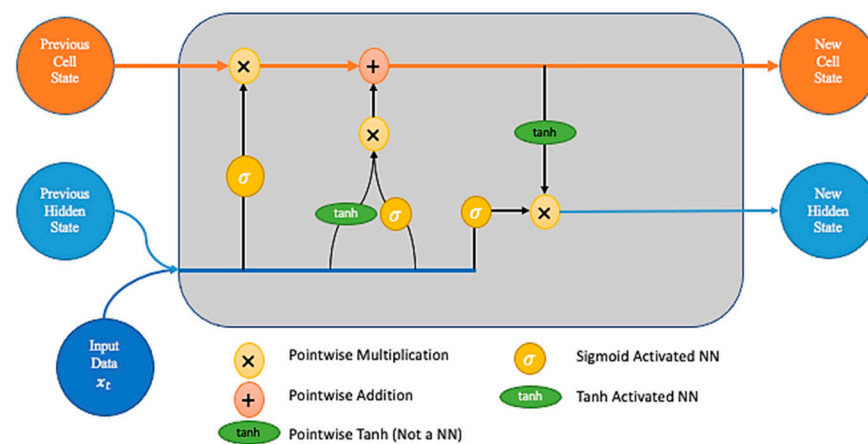


Figure 4. LSTM cell.

In our study, we developed our own RNN based on LSTM layers. The model architecture, as illustrated in Figure 5, comprises three LSTM layers with progressively decreasing hidden units (64, 32, and 16). This strategic reduction in hidden units serves to manage model complexity and mitigate overfitting, thereby enhancing the model's generalization capacity. Additionally, dropout layers are strategically inserted after each LSTM layer to impose regularization, further fortifying the model against overfitting by randomly deactivating a fraction of neurons during training.

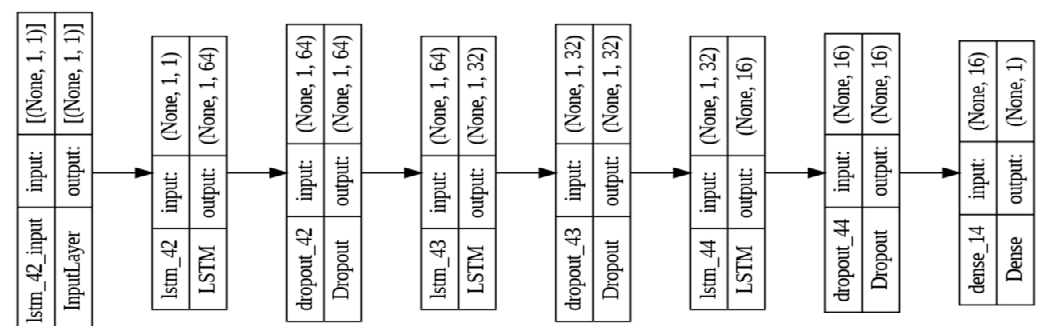


Figure 5. RNN-LSTM model architecture.

The third model in our study was based on transformers, which are a revolutionary architecture in the domain of sequential data processing (Vaswani et al. 2017). Unlike the other models, such as ARIMA and LSTM, which rely on recurrent connections or



convolutions, transformers adopt a fundamentally different approach by employing self-attention mechanisms. This allows them to capture dependencies between input elements across varying distances more efficiently, making them particularly adept at handling long-range dependencies in sequential data. The transformer architecture consists of an encoder–decoder structure, where the encoder processes the input sequence, and the decoder generates the output sequence. Notably, transformers have demonstrated superior performance in natural language processing tasks, achieving state-of-the-art results in machine translation, text generation, and other language-related tasks (Lin et al. 2022).

In our study, we explored the capabilities of transformers for time series forecasting, leveraging their ability to capture complex temporal patterns and dependencies. Figure 6 illustrates the architecture of the transformer model, highlighting its distinctive components and illustrating the flow of information through the network.

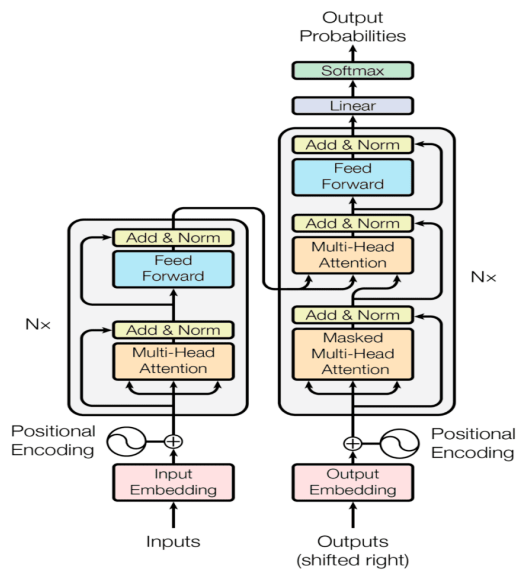


Figure 6. Transformer architecture.

In our implementation, we defined the transformer model using the TensorFlow Keras API (Bisong 2019). The model architecture is instantiated with parameters such as the number of layers, model dimensionality, number of attention heads, and feed-forward network dimension. These parameters are crucial for determining the model's capacity and performance. The architecture of our transformer model, illustrated in Figure 7, consists of several key components. It includes multi-head self-attention mechanisms, feed-forward neural networks, layer normalization, and positional encoding. These components enable the model to efficiently capture complex dependencies within sequential data.

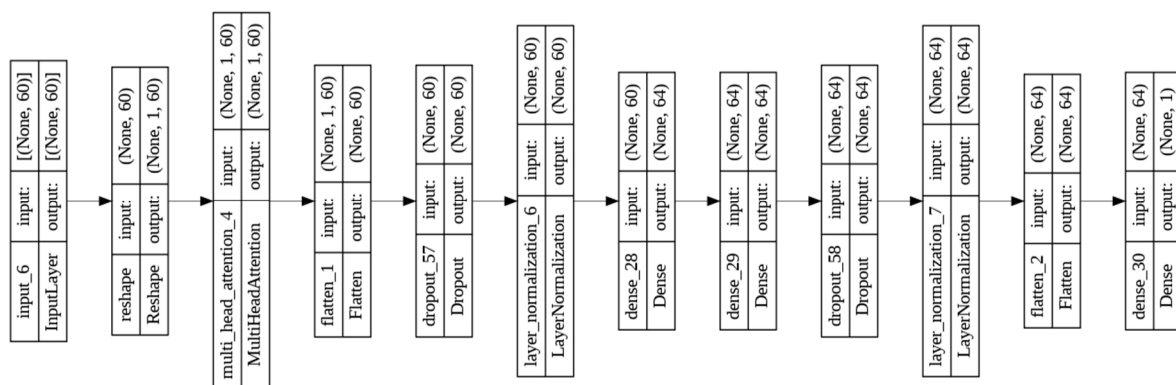


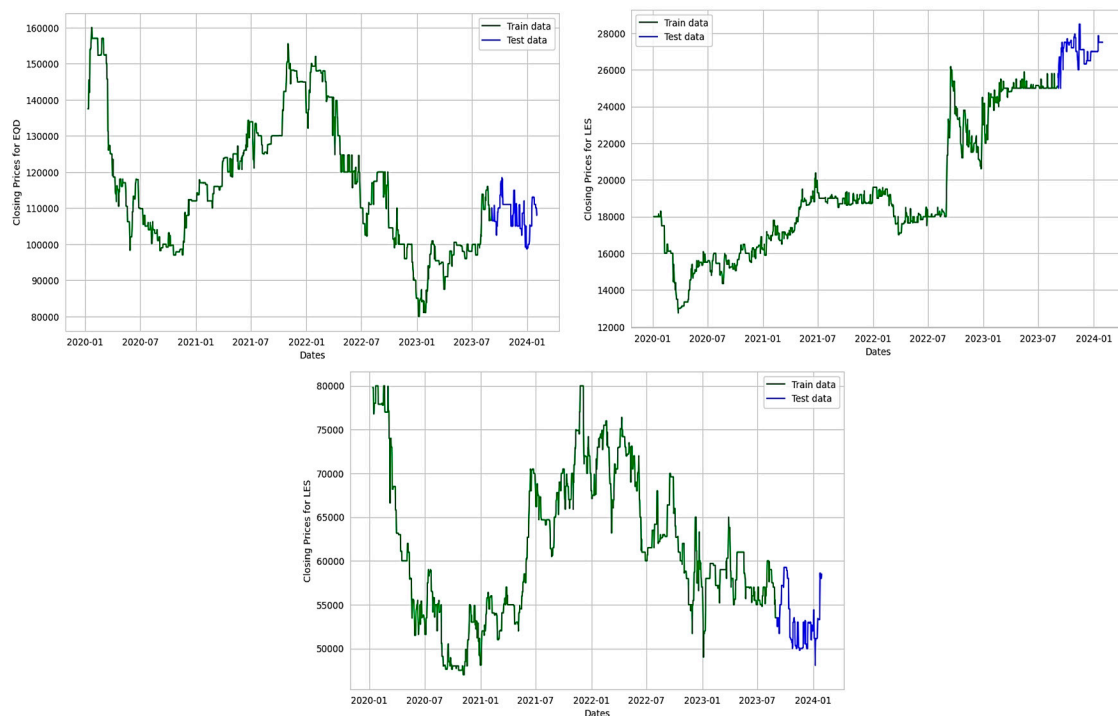
Figure 7. Transformer model architecture.

The transformer model class encapsulates the model's architecture. It comprises multiple layers, each containing a multi-head self-attention mechanism, followed by a feed-forward neural network (FFN). The input and target sequences are concatenated and passed through the model, with self-attention and FFN layers processing the information iteratively. The instantiation of the model involves specifying parameters such as the number of layers, model dimensionality, number of attention heads, and feed-forward network dimension. These parameters dictate the model's architecture and determine its capacity to learn from the data.

### 2.3. Training and Evaluating the Models

To effectively train and evaluate our models, we adopted a data-splitting strategy tailored to the nature of our sequential data. Considering the inherent dependency on chronological order and the preservation of temporal relationships, traditional cross-validation methods were not suitable. Instead, we partitioned our data into training and testing subsets, allocating the last 10% of the data for testing purposes and reserving the initial 90% for model training. This approach ensured that our models were trained on historical data while being evaluated on unseen future data, facilitating a more realistic assessment of their predictive performance.

Figure 8 illustrates this data-partitioning strategy, depicting the training and testing portions in green and blue colors, respectively, across the three companies represented in our dataset. This delineation allows for the assessment of each model's forecasting capabilities using the last six months of data, providing insights into their effectiveness across different temporal contexts.



**Figure 8.** Partitioning of data into training and testing sets for model evaluation.

In the training phase of our models, we began by normalizing our data using the MinMaxScaler, which scales the data to a range between 0 and 1, facilitating convergence and enhancing the performance of the models. For the ARIMA model, we employed a grid search approach to determine the optimal parameters for the model, including the order ( $p$ ,  $d$ ,  $q$ ). This iterative process involves fitting multiple ARIMA models with different parameter combinations to the training data and selecting the configuration that minimizes the error. In contrast, for the LSTM and transformer models, training was conducted

using the Adam optimizer with MSE loss. The Adam optimizer updates the parameters iteratively based on the gradient of the loss function. The update rule for Adam is shown in Equation (2):

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{v_t} + \epsilon} \cdot m_t \quad (2)$$

where  $\theta_t$  represents the parameters at time step  $t$ ,  $\eta$  is the learning rate,  $m_t$  is the estimate of the first moment of the gradients,  $v_t$  is the estimate of the second moment of the gradients, and  $\epsilon$  is a small constant to prevent division by zero.

During the training phase, input sequences were sequentially passed through the model, and the model's parameters were adjusted to minimize the prediction error. Also, to prevent overfitting, early stopping and model checkpoint callbacks were implemented, allowing the training process to halt when the model's performance on a validation set ceased to improve significantly. The training data were iterated over multiple epochs, with each epoch comprising batches of data, to optimize the model's parameters. Additionally, the model's generalization ability was monitored by evaluating its performance on a validation set throughout the training process. Once training was complete, the models' performance was evaluated using various metrics, such as MSE, MAE, and R-squared, on the testing set, which comprises a portion of the data reserved exclusively for model evaluation.

These evaluation metrics provided insights into the models' accuracy and effectiveness in forecasting future values. The MSE, MAE, and R-squared are presented in Equations (3)–(5):

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{y}_i| \quad (4)$$

$$R^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5)$$

where  $y_i$  corresponds to the actual values, and  $\hat{y}_i$  corresponds to the predicted values.

### 3. Results and Discussion

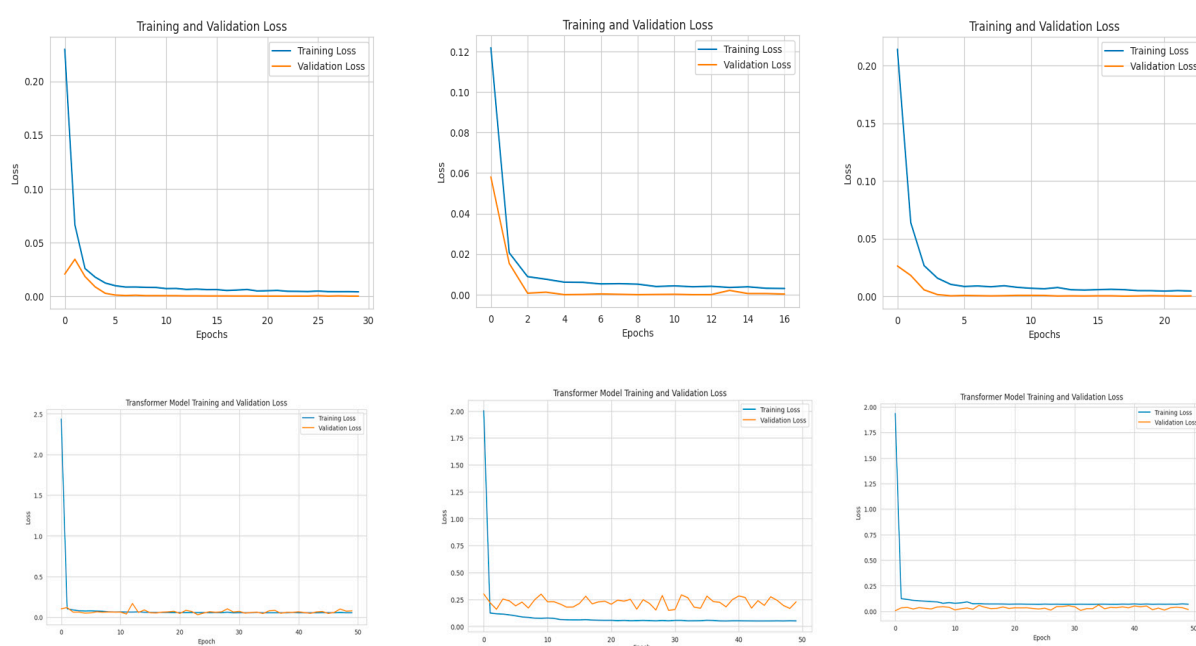
After training the three models, we obtained interesting results for each one. As depicted in Table 1, the performance metrics MSE, MAE, and  $R^2$  for the three companies, EQD, LES, and SLF, showed variations across the different models. ARIMA yielded a notable  $R^2$  score of 0.85 for SLF's data, indicating its effectiveness in capturing the underlying patterns. In contrast, LSTM demonstrated impressive results, with  $R^2$  scores exceeding 0.99 for EQD and LES data and 0.95 for SLF data, underscoring its robustness in financial data forecasting. However, the transformer model struggled to produce satisfactory results, with negative  $R^2$  scores indicating poor performance. This discrepancy can be attributed to transformers' reliance on large datasets, typically more prevalent in text-based tasks, unlike financial time series data. Comparing ARIMA and LSTM, while ARIMA performed reasonably well, LSTM's superior performance across all metrics highlights its suitability for capturing the complex dynamics inherent in financial data. These results underscore the significance of accurate predictions for financial decision-making. The high  $R^2$  scores attained by LSTM indicate its potential for enhancing forecasting accuracy, thereby aiding stakeholders in making informed investment decisions for the three companies evaluated.

The training and validation loss curves for the LSTM and transformer models are illustrated in Figure 9. Notably, both the training and validation loss curves exhibited parallelism and followed similar trajectories. This alignment suggests that our models neither suffered from overfitting nor underfitting, indicating a balanced learning process. Furthermore, the fluctuations in the loss curves revealed that certain models may halt training prematurely, such as the LSTM model for the first company after 30 epochs or the transformer model for the second company after 16 epochs. This observation underscores the effectiveness of our early stopping and checkpoint mechanisms, which effectively

terminated model training when the error stabilized or when signs of overfitting emerged. The synchronization between the training and validation loss curves, coupled with the timely cessation of training, enhanced our confidence in the robustness and generalization capability of the developed models.

**Table 1.** Stock market forecasting results.

Model	ARIMA			LSTM			Transformer		
Scores	MSE	MAE	R <sup>2</sup>	MSE	MAE	R <sup>2</sup>	MSE	MAE	R <sup>2</sup>
EQD	0.00081	0.0144	0.7317	0.000006	0.00246	0.9978	0.04019	0.19265	−12.07
LES	0.00076	0.0124	0.5844	0.000005	0.0015	0.9977	0.27211	0.51993	−151.8
SLF	0.0011	0.0163	0.8527	0.000032	0.0151	0.9592	0.13496	0.35644	−16.05



**Figure 9.** Training and validation losses.

Lastly, Figure 10 presents a visual comparison of the obtained results through comparative graphs depicting the prediction values (highlighted in red) alongside the actual values (depicted in green). As previously discussed, the LSTM model demonstrated exceptional forecasting accuracy, particularly for the first two companies. Additionally, ARIMA yielded commendable results across all three companies, with a notable observation in the third company, where ARIMA visually outperformed LSTM. However, it is important to note that while ARIMA may excel in certain instances, the overall performance, as quantified by the average score, indicated LSTM's superior stability and proximity to the actual values across various forecasting scenarios. Conversely, transformer visibly underperformed within our study, signaling a limitation in its effectiveness. This underscores the preference for LSTM, especially when dealing with Moroccan financial data, offering researchers valuable insights for model selection and future exploration.

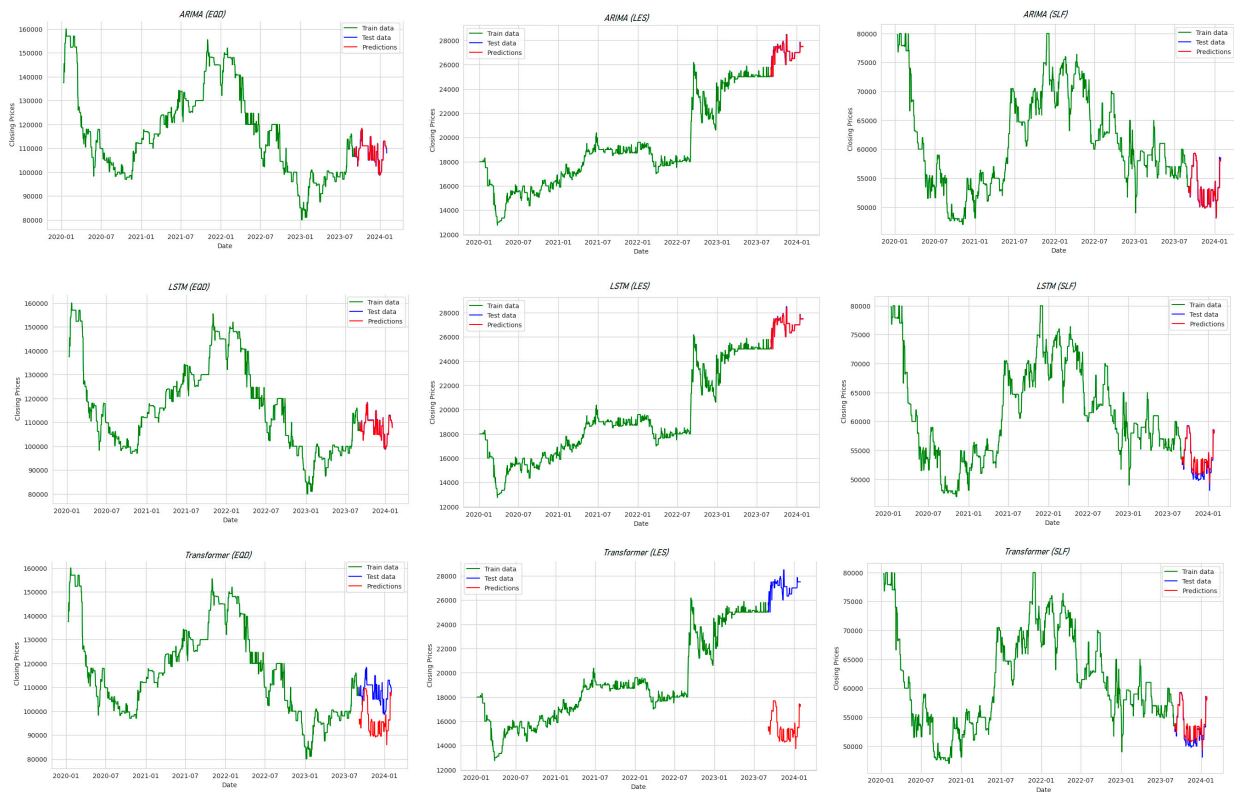


Figure 10. Forecasting results.

#### 4. Conclusions

In this comprehensive study, we explored the predictive capabilities of ARIMA, LSTM, and transformers using data from three prominent Moroccan credit companies listed on the Casablanca Stock Exchange. Each model was meticulously tailored to the unique characteristics of the company data, and evaluation was conducted based on the last 10% of the dataset. The results obtained from our analysis revealed the remarkable performance of LSTM, underscoring the effectiveness of recurrent neural networks specifically designed for time series (sequential) data. This finding highlights the potential for using advanced forecasting techniques in the Moroccan Stock Market.

However, it is important to acknowledge several limitations inherent in our approach. Firstly, our study relied on historical stock price data, which assumes that future market conditions will resemble those observed in the past. This assumption may not always hold true, particularly in volatile or rapidly changing markets. Secondly, while we employed rigorous model evaluation techniques, such as cross-validation and hyperparameter tuning, the performance of our models could be affected by factors such as data quality, market anomalies, and external economic events not explicitly accounted for in our analysis. Thirdly, the generalizability of our findings beyond the specific companies and timeframe studied may be limited, considering the variability in market dynamics across different sectors and periods. Additionally, the nominal measures of forecasting performance, particularly MSE differences, were not explicitly tested in our study, which could be considered a limitation of our approach. Lastly, our study primarily focused on one-step-ahead predictions, which, while applicable to short-term investment decision-making, may not reflect the needs of investors with longer investment horizons. Future research could explore longer investment horizons to provide a more comprehensive analysis of model performance over different time spans.

The predictability of stock prices using daily data can be attributed to several rational factors, including the structure and trading restrictions of the Moroccan CSE, as well as the possibility that investors are being rewarded for taking risks. This predictability suggests



that financial asset returns in the Moroccan market may be somewhat consistent with empirical evidence from other financial markets, though it does not directly address market efficiency. Future research could further investigate these aspects to provide deeper insights into the underlying mechanisms of stock price movements.

Despite these limitations, our research contributes significantly to the field of financial forecasting in Morocco, providing actionable insights that can inform strategic decisions and drive positive outcomes for companies operating within the Moroccan market. By leveraging advanced predictive models and harnessing the power of data-driven insights, businesses in Morocco can gain a competitive edge and thrive in the dynamic landscape of the Casablanca Stock Exchange. Furthermore, our study serves as a valuable resource for investors seeking to capitalize on the potential of the Casablanca Stock Exchange, offering empirical evidence and reliable forecasting models to guide decision-making processes. In essence, our study not only advances the understanding of stock market forecasting in the Moroccan context but also lays the foundation for future research endeavors aimed at unlocking the full potential of the Moroccan Stock Exchange. As the market continues to evolve and mature, our findings serve as a catalyst for innovation and growth, inspiring companies and investors alike to embrace the opportunities presented by the burgeoning Moroccan financial landscape.

**Author Contributions:** Conceptualization, K.L. and M.B.; methodology, K.L. and M.B.; software, K.L. and M.B.; validation, K.L. and M.B.; formal analysis, K.L. and M.B.; investigation, K.L. and M.B.; resources, K.L. and M.B.; data curation, K.L. and M.B.; writing—original draft preparation, K.L. and M.B.; writing—review and editing, K.L. and M.B.; visualization, K.L. and M.B.; supervision, K.L. and M.B.; project administration, K.L. and M.B.; funding acquisition, K.L. and M.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data supporting this study's findings are available at <https://www.investing.com/>.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Ahmed, Abdullahi D., and Rui Huo. 2021. Volatility transmissions across international oil market, commodity futures and stock markets: Empirical evidence from China. *Energy Economics* 93: 104741. [CrossRef]
- Azzam, Henry T. 2015. *The Emerging Middle East Financial Markets*. Bloomington: AuthorHouse.
- Baali, Boubker, Brahim Elmorchid, and Brahim Mansouri. 2023. Determinants of the variation in the liquidity behavior of the casablanca stock exchange: A global econometric analysis on time series. *Finance & Finance Internationale* 1: 1–22.
- Ballings, Michel, Dirk Van den Poel, Nathalie Hespeels, and Ruben Gryp. 2015. Evaluating multiple classifiers for stock price direction prediction. *Expert Systems with Applications* 42: 7046–56. [CrossRef]
- Bao, Wei, Jun Yue, and Yulei Rao. 2017. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS ONE* 12: e0180944. [CrossRef] [PubMed]
- Bhowmik, Roni, and Shouyang Wang. 2020. Stock market volatility and return analysis: A systematic literature review. *Entropy* 22: 522. [CrossRef] [PubMed]
- Bisong, Ekaba. 2019. Google Colaboratory. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*. Berkeley: Apress, pp. 59–64.
- Brière, Marie, Matthieu Keip, and Tegwen Le Berthe. 2022. Artificial Intelligence for Sustainable Finance: Why It May Help. Available online: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4252329](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4252329) (accessed on 21 May 2024).
- Chen, Yu, Ruixin Fang, Ting Liang, Zongyu Sha, Shicheng Li, Yugen Yi, Wei Zhou, and Huilin Song. 2021. Stock price forecast based on CNN-BiLSTM-ECA Model. *Scientific Programming* 2021: 2446543. [CrossRef]
- Dibiah, Tobin Obari, and Ogechukwu Rita Mojekwu. 2023. Test of random walk on selected stock markets in Africa. *GPH-International Journal of Business Management* 6: 1–32.
- Dridi, Salim. 2021. Supervised learning—A systematic literature review. *Preprint*. Available online: <https://osf.io/gq5tc/> (accessed on 21 May 2024).
- Fang, Wei, Yupeng Chen, and Qiongying Xue. 2021. Survey on research of RNN-based spatio-temporal sequence prediction algorithms. *Journal on Big Data* 3: 97. [CrossRef]
- Khan, Salman, Muzammal Naseer, Munawar Hayat, Syed Waqas Zamir, Fahad Shahbaz Khan, and Mubarak Shah. 2022. Transformers in vision: A survey. *ACM Computing Surveys CSUR* 54: 1–41. [CrossRef]

- Kompella, Subhadra, and Kalyana Chakravarthy Chilukuri. 2020. Stock market prediction using machine learning methods. *International Journal of Computer Engineering and Technology* 10: 2019. [\[CrossRef\]](#)
- Kuvshinov, Dmitry, and Kaspar Zimmermann. 2022. The big bang: Stock market capitalization in the long run. *Journal of Financial Economics* 145: 527–52. [\[CrossRef\]](#)
- Li, Hao, Yanyan Shen, and Yanmin Zhu. 2018. Stock price prediction using attention-based multi-input LSTM. Paper presented at Asian Conference on Machine Learning, PMLR, Beijing, China, November 14–16; pp. 454–69.
- Lin, Tianyang, Yuxin Wang, Xiangyang Liu, and Xipeng Qiu. 2022. A survey of transformers. *AI Open* 3: 111–32. [\[CrossRef\]](#)
- Low, Pi Rey, and Eric Sakk. 2023. Comparison between autoregressive integrated moving average and long short term memory models for stock price prediction. *IAES International Journal of Artificial Intelligence (IJ-AI)* 12: 1828. [\[CrossRef\]](#)
- Løkken, Vegard Nordgård, and Ørjan Østensen Aas. 2020. Volatility Spillover between Commodities and Equities-a Study of Oil, Steel, and Cotton. Master's dissertation, University of Stavanger, Stavanger, Norway.
- Lu, Wenjie, Jiazheng Li, Jingyang Wang, and Lele Qin. 2021. A CNN-BiLSTM-AM method for stock price prediction. *Neural Computing and Applications* 33: 4741–53. [\[CrossRef\]](#)
- Ma, Rui, Hamish D. Anderson, and Ben R. Marshall. 2016. International stock market liquidity: A review. *Managerial Finance* 42: 118–35. [\[CrossRef\]](#)
- Mishra, Sasmita, and Sudarsan Padhy. 2019. An efficient portfolio construction model using stock price predicted by support vector regression. *The North American Journal of Economics and Finance* 50: 101027. [\[CrossRef\]](#)
- Mohammad, Maha A. Alellah. 2023. Ensemble learning: A review. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 14: 517–23.
- Nosratabadi, Saeed, Amirhosein Mosavi, Puhong Duan, Pedram Ghamisi, Ferdinand Filip, Shahab S. Band, Uwe Reuter, Joao Gama, and Amir H. Gandomi. 2020. Data science in economics: Comprehensive review of advanced machine learning and deep learning methods. *Mathematics* 8: 1799. [\[CrossRef\]](#)
- Nti, Isaac Kofi, Adebayo Felix Adekoya, and Benjamin Asubam Weyori. 2020. A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review* 53: 3007–57. [\[CrossRef\]](#)
- Patel, Jigar, Sahil Shah, Priyank Thakkar, and Ketan Kotecha. 2015. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications* 42: 259–68. [\[CrossRef\]](#)
- Prasad, Vivek Kumar, Darshan Savaliya, Sakshi Sanghavi, Vatsal Sakariya, Pronaya Bhattacharya, Jai Prakash Verma, Rushabh Shah, and Sudeep Tanwar. 2022. Stock Price Prediction for Market Forecasting Using Machine Learning Analysis. Paper presented at International Conference on Computing, Communications, and Cyber-Security, Dalian, China, October 17–19; Berlin and Heidelberg: Springer, pp. 477–92.
- Pulungan, Dolly Parlagutan, Sugeng Wahyudi, Suharnomo Suharnomo, and Harjum Muharam. 2018. Technical analysis testing in forecasting socially responsible investment index in Indonesia stock exchange. *Investment Management and Financial Innovations* 15: 135–43. [\[CrossRef\]](#)
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention Is All You Need. *arXiv* arXiv:1706.03762.
- Wahyudi, Setyo Tri. 2017. The ARIMA Model for the Indonesia Stock Price. *International Journal of Economics & Management* 11: 223–36.
- Wang, Haiyao, Jianxuan Wang, Lihui Cao, Yifan Li, QiuHong Sun, and Jingyang Wang. 2021. A stock closing price prediction model based on CNN-BiSLSTM. *Complexity* 2021: 5360828. [\[CrossRef\]](#)
- Wang, Shuzhen. 2023. A stock price prediction method based on BiLSTM and improved transformer. *IEEE Access* 11: 104211–23. [\[CrossRef\]](#)
- Wu, Jimmy Ming-Tai, Zhongcui Li, Norbert Herencsar, Bay Vo, and Jerry Chun-Wei Lin. 2023. A graph-based CNN-LSTM stock price prediction algorithm with leading indicators. *Multimedia Systems* 29: 1751–70. [\[CrossRef\]](#)
- Zaimi, Wiam. 2022. An empirical analysis of a stock market index of a developing country: Case of the main index of The Casablanca Stock Exchange MASI. *Global Business Finance Review* 27: 1–16. [\[CrossRef\]](#)
- Zhang, Cheng, Nilam N. A. Sjarif, and Roslina B. Ibrahim. 2022. Decision fusion for stock market prediction: A systematic review. *IEEE Access* 10: 81364–79. [\[CrossRef\]](#)

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.