International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 15, No.1 (Jan-24)

http://dx.doi.org/10.12785/ijcds/150103

# Global Stock Price Forecasting during a Period of Market Stress Using LightGBM

**Omaima Guennioui[1], Dalila Chiadmi[1] and Mustapha Amghar[1]**

[1]*Mohammadia School of Engineering, Mohammed V University in Rabat, Rabat, Morocco*

*Received 4 Apr. 2023, Revised 23 Nov. 2023, Accepted 23 Dec. 2023, Published 1 Jan. 2024*

**Abstract:** Accurately predicting future stock prices is crucial for investors, particularly during market stress, enabling informed decisions to mitigate losses and reduce financial exposure. While machine learning techniques have shown promise in this field, most studies have focused on local models. Global forecasting models, which are trained on a variety of time series, have demonstrated encouraging outcomes in outperforming local models. This article presents a global forecasting approach utilizing Light Gradient-Boosting Machine (LightGBM) to predict stock prices in the Moroccan market during highly volatile period. The study includes a sector-wise analysis, with the most volatile sector being evaluated, and correlated stocks from other sectors were considered to enhance the data. Our findings highlight the effectiveness of utilizing a global forecasting approach with an optimized LightGBM model for Moroccan stock price prediction, outperforming models such as Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU) and Convolutional Neural Networks (CNN). Additionally, incorporating data from multiple stocks improves the accuracy of stock price prediction.

**Keywords:** Stock market, Market stress, LightGBM, Global Forecasting Models

## 1. Introduction

Predicting stock market trends with precision is a significant and arduous task involving time series, especially during periods of market stress. A stock market stress period is characterized by a time of considerable disturbance or uncertainty within the financial markets, typically marked by a rapid decline in stock prices and increased market volatility. In these periods, investors may face heightened fear and uncertainty, prompting more selling activity and exacerbating the price decline. Precise prediction of future stock prices is essential for investors, particularly during such times, as it allows them to make informed decisions to minimize potential losses and decrease their financial risk. Consequently, accurate stock price prediction serves as a priceless asset for executing effective investment strategies.

This paper's case study centers on the Covid-19 pandemic, a period marked by significant turbulence and uncertainty in the financial markets. Due to the outbreak of the Covid-19 virus, a prolonged global health and economic crisis was anticipated, causing financial asset prices to plummet sharply beginning in late February 2020. Subsequently, the search for safety by investors resulted in widespread selling, which significantly impacted the equity market. The volatility levels witnessed in the middle of March 2020 in the United States stock market were comparable to, or even surpassed, those experienced in October 1987, December 2008, and earlier, in the late 1920s and early 1930s [1]. The Moroccan stock market, like its global counterparts, was jolted by high levels of volatility and a significant drop in the MASI index (Moroccan All Shares Index) within a few days. Between March 2, 2020, when Morocco reported its first COVID case, and March 31, 2020, the MASI experienced a steep decline of over -20% (-16% for the S&P 500 Index and Euro Stoxx 50 Index -US market and UE market-).

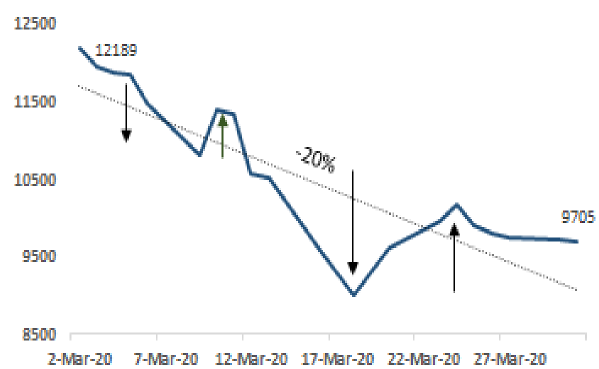


Figure 1. MASI closing price during market stress period

In general, there are two primary methods of analyzing stocks. The first, known as fundamental analysis, involves

*E-mail address: omaimaguennioui@research.emi.ac.ma, chiadmi@emi.ac.ma, amghar@emi.ac.ma* 

evaluating the intrinsic value of a stock. This approach implies an in-depth analysis while exploiting balance sheet, strategic initiatives and microeconomic indicators of the studied company. The second type is technical analysis, which involves studying the statistical data generated by market activity, such as past prices and volumes, to determine the stock's evolution.
The growing prevalence of machine learning in various industries [2] over the past few decades has resulted in a surge of interest in and use of machine learning techniques. Since then the scientific community has studied the applicability of these methods to the problem of predicting stock prices and has proven the relevance of this approach [3]. However, the majority of this research has primarily focused on implementing local models to forecast stock prices, which entails utilizing a distinct set of parameters for each individual stock.

Lately, Global Forecasting Models, which are trained on numerous time series and built using a single model based on all available data, have shown promising results [4] in both real-world applications and forecasting contests [5][6], outperforming many local forecasting methods. Unlike local models, which concentrate on specific data subsets or regions within a dataset and train separate models for each group, global forecasting models possess the ability to gather cross-series information during the training process. This capability allows them to handle complexity and over fitting on a global level [7].
Light Gradient-Boosting Machine (LightGBM), as an enhanced version of the gradient boosting algorithm created by Microsoft, is specifically designed to handle a large number of data instances and features. This makes LightGBM a suitable choice for building models that can be generalized across multiple datasets with cross-correlations, making it easier to develop accurate solutions for complex problems involving large datasets [8].

This article presents a global forecasting approach using LightGBM to predict future prices of different stocks in the Moroccan market during a period of market stress which is in our case study the COVID-19 pandemic period. The LightGBM model is optimized to improve accuracy. The study also includes a sector-wise analysis, with the model being applied to the stocks of the most volatile sector. To enhance the data, correlated stocks from other sectors were also considered. Afterwards, the predictive outcomes are evaluated against those generated by state-of-the-art deep learning models such as Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU) and Convolutional Neural Networks (CNN), all of them renowned for their precision in solving time series forecasting problems.

This research assesses the efficacy of a global forecasting approach using LightGBM for the Moroccan stock market. The primary contribution of this study is the introduction of an accurate and comprehensive method capable of predicting multiple stocks simultaneously during periods of heightened market volatility.

The remainder of this paper is organized as follows. The second section reviews related works on stock price prediction and global forecasting models. In section 3, we describe the methodology applied, followed by our experimental results, analysis and discussion in Section 4. In Section 5, we summarize some conclusions and suggest directions for future research

## 2. Related Works

There have been a number of works in the use of machine learning methods for stock market prediction. Guo, Li and Xu [9] proposed a hybrid financial time series model called LSTM_LightGBM. They used the opening, closing, highest, lowest prices, trading volume, and adjusted closing price as inputs to the model for prediction. They conclude that the model proposed is stable and feasible in the stock price fluctuation forecast. Vijh, Chandola, Tikkiwal and Kumar [10] applied two popular machine learning techniques, Artificial Neural Network (ANN) and Random Forest (RF), to predict the next day's closing stock prices for five companies from different sectors of operation. To create the models, the financial data (Open, High, Low, and Close prices) of each company were processed to generate new variables that serve as inputs to the models. The results show that both ANN and RF models are efficient in predicting the next day's closing price for the stocks. In [11], Hu, Qin and Zhang selected twelve technical indicators for stock price trend prediction based on correlation analysis. Three models: random forest, logistic regression, and LightGBM were used to forecast three selected stocks. The performance of different models was compared using accuracy, recall, precision, and f1 value in various time windows. This finding has some guiding significance for the improvement of long-term and short-term forecasting performance. This study recommends LightGBM as the preferred model for medium and long-term share price forecasting due to its performance. Li, Xu and Li [12] explored the impact of five indicator categories on stock returns: valuation, profit, scale, risk, and growth. Three algorithms : RF, Extreme Gradient Boosting (XGBoost), and LightGBM are tested using a rolling test, with LightGBM demonstrating the best performance. The paper conducts backtesting on processed data and selects a portfolio strategy comprising 10 stocks. Five investment strategy evaluation indicators are calculated, revealing that the selected portfolio outperforms the benchmark CSI 300 index.
However, in these cited articles and many of the existing literature, they predominantly rely on stock-related data as inputs to their machine learning algorithms for prediction. This approach tends to overlook significant information about market and other relevant entities, which can ultimately limit the accuracy and effectiveness of the models.

On a different aspect, online news data are increasingly being used in the study of stock price prediction, with

researchers applying machine learning techniques to analyze their impact on stock movements. In [13], Mohan et al. employed time series models, neural networks, and a hybrid of both with news data as inputs. The models showed a strong relationship between financial news and stock prices, with the RNN model outperforming others. Despite its effectiveness, the model has demonstrated limitations in its ability to handle low or highly volatile stocks. The study [14] focuses on predicting sudden changes in the closing price of virtual currencies using LightGBM. Natural language information from social media is utilized to create features for prediction. The process involves generating sentence embeddings from tweets using Sentence-BERT, labeling the dataset based on closing price changes, and aggregating the embeddings. The authors' conclusion indicates that incorporating linguistic features from tweet data can be beneficial in predicting sudden changes in price labels. Ko and Chang in [15] analyzed news data and forum posts to conduct sentiment analysis and forecasted the next day's stock opening price using the LSTM neural network. Results indicated that combining sentimental information from both sources improved model accuracy. In [16] Sharaf, Hemdan, El-Sayed and El-Bhnasawy presented a survey of sentiment analysis and stock analysis, followed by the introduction of an efficient proposed system that combines sentiment analysis of social news and historical data analysis. The proposed system faced limitations during the Covid-19 period.
Despite its potential benefits, sentiment analysis has limitations, depending on the attitudes of key figures or successful analysts, and thus, it can be biased. Furthermore, positive opinions may be based on past performance rather than future market potential.

In recent years, global models for time series have been developed to capitalize on similarities among related time series. From the M5 Kaggle's Forecasting Competitions, Makridakis, Spiliotis and Assimakopoulos [6] derived three crucial findings for forecasting: (1) LightGBM exhibits exceptional accuracy in predicting retail sales, (2) the inclusion of external adjustments and explanatory variables boosts precision, and (3) employing cross learning and cross-validation is advantageous. They also emphasized that machine learning methods have prevailed in recent competitions, underscoring the significance of combining statistics and data science. In this article, we have incorporated the recommendations provided, particularly those focused on enhancing forecasting accuracy.

In a context where global models have proven effective in addressing large-scale data challenges, this paper seeks to apply this approach to the stock market using optimized LightGBM during a period of market instability. Furthermore, our research explores the significance of the correlation between the closing prices of stocks across various sectors, particularly those with a strong correlation with the sector under investigation. Our numerical findings substantiated this hypothesis.

## 3. Methodology

### A. Model Dataset

We retrieved the data from Casablanca Stock Exchange website (www.casablanca-bourse.com) for the period from 01/12/2019 to 31/12/2020. The main variables of interest were the closing prices of individual stocks and their corresponding sectors. Additionally, we retrieved for the same time frame the closing prices of sectoral indices : MASI Building and Construction, MASI Banking, MASI Insurance, MASI Food/Production, MASI Chemicals , MASI Oil & Gas, Funding & Financial Services, MASI Beverages, MASI Distributors, MASI Electricity, MASI Pharmaceuticals, MASI Mining, MASI Materials & IT, MASI Telecom, MASI ESG, MASI Ingineering AND Equipment, MASI Hotels, MASI Real Estate Development, MASI Transportation Services, MASI Real Estate Investment and Promotion, MASI Holdings, MASI Forestry & Paper. Figure 2 displays the evolution of the closing prices for four different sectors of the MASI during the studied period.

#### 1) Historical Volatility

The objective of this paper is to predict the stock price of the most volatile sector during Covid-19 pandemic. We used sectoral closing prices retrieved from our source in

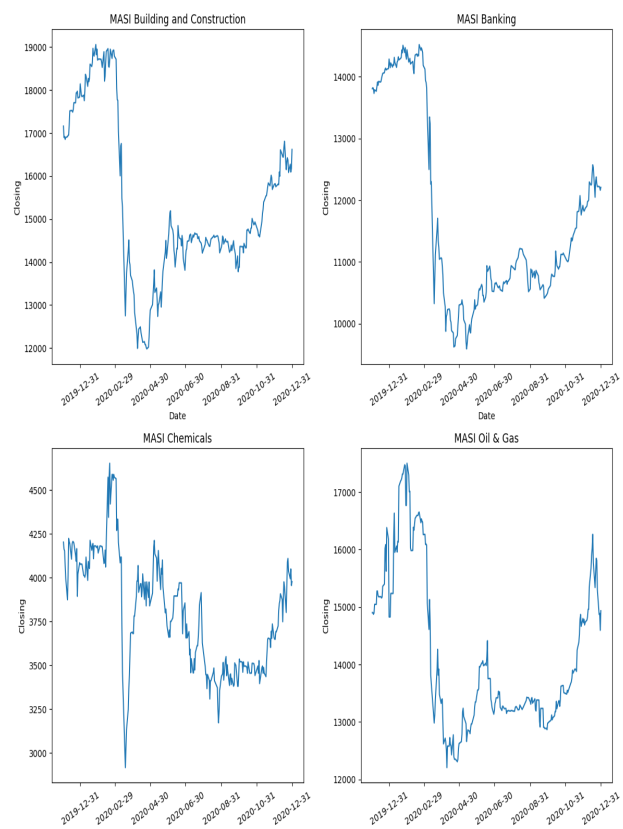


Figure 2. Closing prices of four sectoral indices in MASI

order to compute historical volatility as follows:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(r_i - m)^2}{n-1}} \quad (1)$$

where :

$$r_i = \frac{P_{i+1} - P_i}{P_i} \quad and \quad m = \frac{\sum_{i=1}^{n} r_i}{n} \quad (2)$$

In the provided formula, the symbol $\sigma$ represents the standard deviation, which is a measure of historical volatility. The variable $r_i$ represents the individual returns at each data point, m represents the mean of the returns, $P_i$ represents the closing price of sectoral indices, and n denotes the total number of data points.

Our analysis of Moroccan sectoral indices for the year 2020 - marked by the Covid-19 pandemic - revealed a notably high annualized volatility in the Chemical index at 38%. Details are given in Table I.
When computing volatility, we excluded sectors that constitute less than 0.3% of the MASI's total composition, as we consider this level too low. Consequently, our study and subsequent tests will concentrate on the Chemical sector, which comprises the following stocks: SNEP and MAGHREB OXYGENE.

*2) Sectoral Correlation*

In our applied approach, the global model will perform a cross-learning of the stocks within the Chemical index, which is the sector under study. Subsequently, we will assess the potential benefit of incorporating the stocks from the two most highly correlated sectors with the Chemical index into our prediction set.
Figure 3 displays the correlation among all indices through a heat map, which uses distinct colors to clearly represent the relationship between different attributes. As shown in the figure, Funding & Financial Services Index, as well as Forestry & Paper, show a high correlation with the sector under study. The values in the figure were computed using Pearson's correlation coefficient (3), which is a formula for determining correlation strength.

$$r(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} \quad (3)$$

TABLE I. Top 5 of the most volatile Indices during 2020

| Index | Annualized Volatility |
|---|---|
| Chemicals | 38% |
| Real Estate Investment and promotion | 36% |
| Transportation Services | 33% |
| Mining | 30% |
| Building and Construction | 28% |

We retain stocks of the two indices Funding & Financial Services and Forestry & Paper for the enrichment of our prediction set.The stocks that constitute these two indices are as follows: DIAC SALAF, EQDOM, MAGHREBAIL, MAROC LEASING, and SALAFIN for Funding & Financial Services, and MED PAPER for Forestry & Paper.

*B. Global Optimized LightGBM Model*

*1) LightGBM*

LightGBM is a machine learning algorithm released by Microsoft in 2017 that is part of the family of gradient boosting decision trees. Specifically, it is designed for efficient training of large-scale supervised learning tasks, such as regression, classification, and ranking. LightGBM achieves high accuracy on these tasks by iteratively constructing a set of decision trees and optimizing the leaf-wise split strategy [17]. This strategy selects the optimal feature value to split on for each tree leaf in order to maximize the overall gain in accuracy [8].
One key advantage of LightGBM is its speed and scalability. It is designed to handle large datasets with millions of examples and features. LightGBM achieves this by employing several techniques. Firstly, it uses a histogram-based approach for feature discretization, which allows it to quickly transform continuous features into discrete values. Secondly, it uses a leaf-wise data layout rather than a level-wise layout, which means that it only computes the gradients and hessians of the examples that fall within the same leaf, reducing the computational cost. Finally, it supports parallel and distributed training, allowing it to scale to even larger datasets.

*2) Model Overview*

We implemented the model structure illustrated in Figure 4, which begins with data preprocessing. In our case, this step involved handling missing values and normalizing the data prior to training the model. However, due to the large number of parameters associated with LightGBM, which control model's behavior, the default parameter settings might not yield optimal performance for the dataset used in the study. To address this issue, we employed a hyperparameter tuning algorithm to adjust the model parameters and identify the best combination of settings for the dataset. Once the optimal parameters are identified, the model is trained and used to make predictions.

*3) Hyperpameters Tuning*

LightGBM model contains several internal hyperparameters that can significantly impact the accuracy of its predictions. However, using the default values for these hyperparameters may not necessarily yield the best results for predicting stock prices. To identify the optimal combination of hyperparameters for this specific dataset, we employed a hyperparameter tuning technique called Grid Search. This method involves systematically evaluating the performance of the model for each combination of hyperparameters within a predefined grid [18]. By evaluating all possible combinations, we selected the hyperparameter set

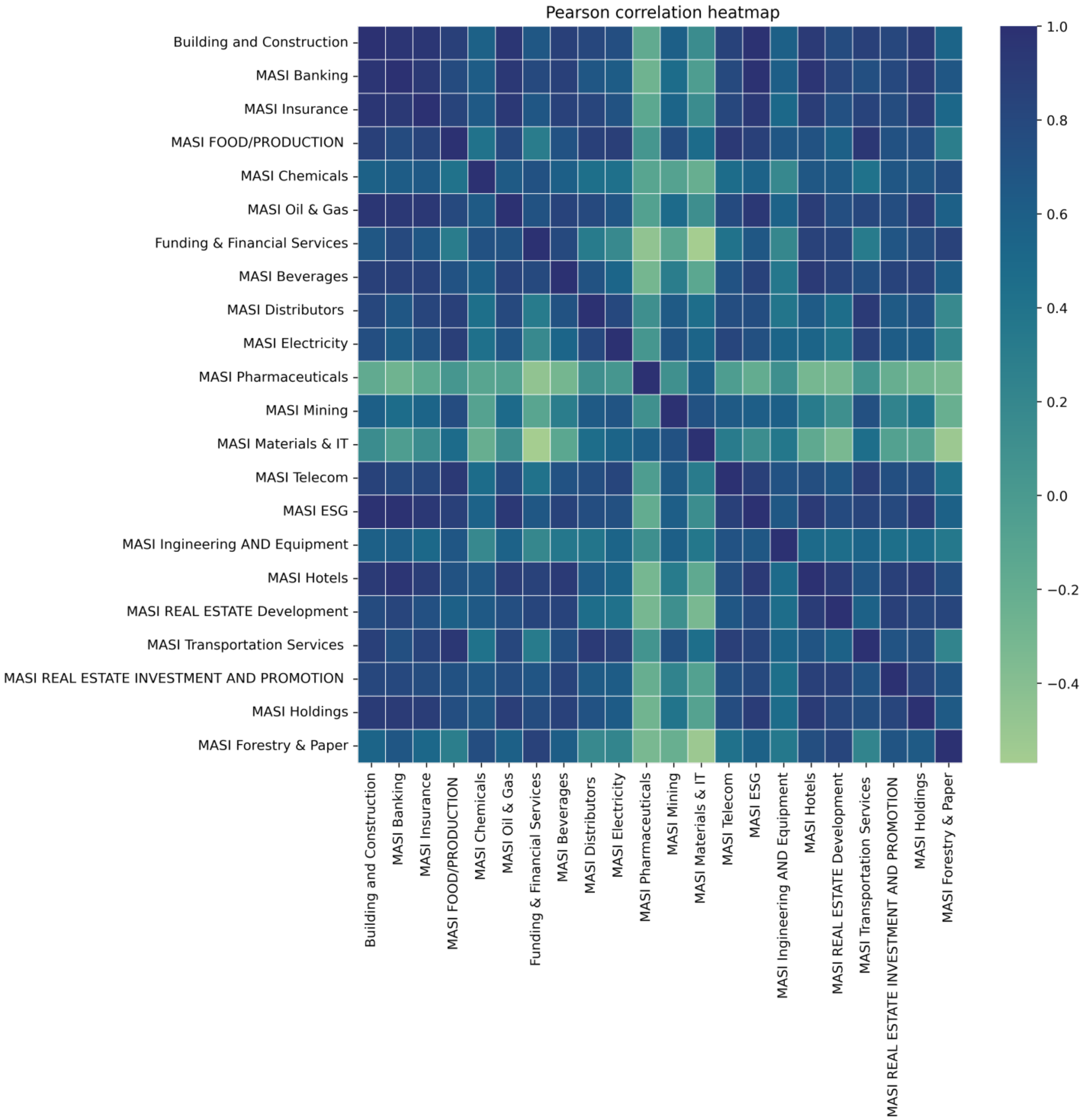


Figure 3. Heat map between indices

that yielded the best model performance. Grid search, a commonly used technique for optimizing machine learning models, was employed alongside cross-validation to enhance the model's generalization performance in our study. Table II provides comprehensive information on the hyperparameters and their respective definitions, along with the search range for each hyperparameter. This search range was employed to explore a various range of values during the tuning procedure.

## 4. Results and Discussion

### A. *Datasets & Baselines*

In this project we focused on predicting the next day closing price of the selected stocks. The dataset used for this purpose consists of daily closing prices for the chosen stocks during the period from 01/12/2019 to 31/12/2020, with the data adjusted for bank holidays and aligned across stocks.

To train the predictive model, the raw dataset was split into training and testing sets in a ratio of 0.8 and 0.2,

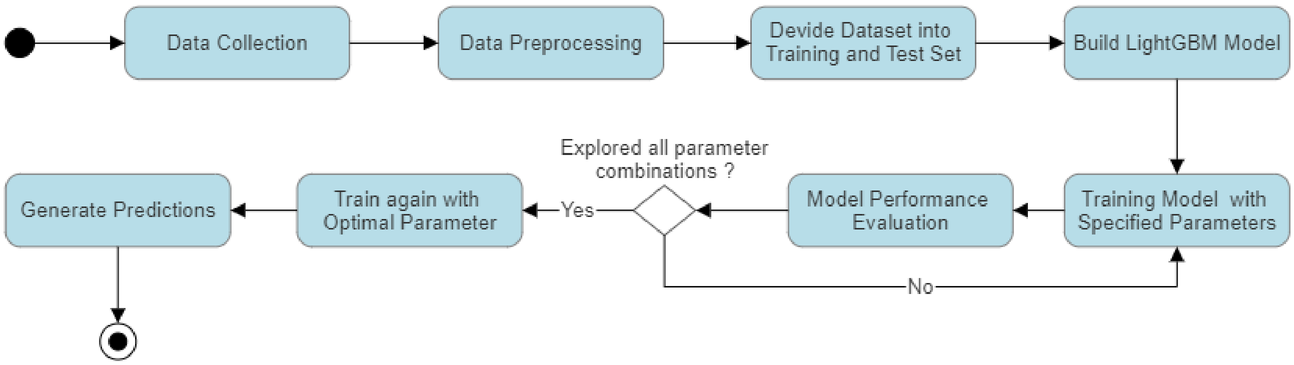


Figure 4. Applied process chart

TABLE II. Hyperparameter tuning details for LightGBM

| | | |
|---|---|---|
| n_estimators | Number of iterations | (100, 200, 300) |
| learning rate | Shrinkage coefficient of each tree | (0.001, 0.02, 0.04) |
| bagging_fraction | Percentage of data used per iteration | (0.3, 0.6, 0.9) |
| reg_lambda | Regularization on weights to avoid overfitting | (0.01, 0.1, 0.3) |

respectively.
As baseline, we evaluated the performance of deep learning models that are known for their effectiveness in time series forecasting [19]. Specifically, Long Short-Term Memory (LSTM) networks, a type of Recurrent Neural Networks, were chosen due to their ability to retain information over longer time periods. Convolutional Neural Networks (CNN) were also considered as they can capture temporal and spatial patterns in high-dimensional data. Another type of recurrent neural network, Gated Recurrent Unit (GRU), was also included in our analysis, as it has become popular for time series prediction.
Regarding the programming language, we employed the Python programming language, conducting the development on Google Colab. To implement the deep learning models, we made use of the TensorFlow and Keras packages. Finally, we did not utilize any GPU or cloud computing resources in our experimentation.

## B. Performance Predictor

We evaluated prediction performance of our regression problem with three different evaluation metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and R-squared (R2). Calculated as follow:

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}(y_j - \tilde{y}_j)^2} \tag{4}$$

$$MAE = \frac{1}{n}\sum_{j=1}|y_j - \tilde{y}_j| \tag{5}$$

$$R2 = \sqrt{1 - \frac{\sum_{j=1}(y_j - \tilde{y}_j)^2}{\sum_{j=1}(y_j - \bar{y})^2}} \tag{6}$$

In this context, the symbol y denotes the actual closing price of the test set, while the symbol $\tilde{y}$ represents the predicted value of that price. The symbol $\bar{y}$ represents the mean value of y and n denotes the number of data points in the test set.

## C. Experimental Results

Once the LightGBM model was optimized using the GridSearch algorithm, the resulting parameters were set as follows: the learning rate was set to 0.02 and the number of base learners, n_estimators, was set to 200. Additionally, the bagging_fraction, which controls the fraction of data used in each iteration, was set to 0.3, while reg_lambda, which adds a penalty term to the loss function to prevent over-fitting, was set to 0.1. These optimized parameters were then used as input for the LightGBM model to predict stock prices.
For each stock within the Chemical sector, baseline models were developed using a local approach, in which the models were trained on past observations of the same stock. The reported prediction error represents the average of the errors calculated for each individual prediction. Table III shows the evaluation indicators of the prediction results of the various models with metrics such as RMSE, MAE, and R2 score. It is worth noting that the term "LightGBM" represents the model without parameter tuning, while "Optimized LightGBM" refers to the model that has been tuned with parameter adjustments.
Among the models tested, Optimized LightGBM with cor-

related indices produced the lowest RMSE and MAE values at 0.033 and 0.02, respectively, while also achieving a high R2 score of 0.99. The Optimized LightGBM model, while not using correlated indices, also performed well, with an RMSE of 0.037 and an R2 score of 0.86. In addition, the default hyperparameters of LightGBM are not suitable for the prediction of stock prices. RMSE and MAE calculated for the default hyperparameters predictions are the least accurate of all the experiments.
Regarding deep learning models, trained for each stock. The LSTM and GRU models had similar performance, with both achieving RMSE and MAE values of around 0.038 and 0.029, respectively, but lower R2 scores of 0.66 and 0.64, respectively. Finally, the CNN model produced an RMSE of 0.0375, an MAE of 0.027, and an R2 score of 0.69, indicating that it performed better than the LSTM and GRU models but not as well as the optimized LightGBM models. Optimized LightGBM with correlated indices outperformed all the other models. These findings suggest that incorporating data from multiple stocks can enhance the accuracy and generalization ability of LightGBM models for global stock price forecasting in a context of high volatility of stock market.
We also observed that the LightGBM model was notably faster, completing parameter optimization, training, and prediction for all Chemical sector stocks in just 8.23 seconds. On the other hand, both the LSTM and GRU models took the longest time to predict the two stocks in the studied sector. The CNN model had an intermediate time performance, taking about 37 seconds to predict each stock in the sector. Table IV provides details of the number of predicted stocks, along with the adopted approach and execution time for each model.The R2 score provided represents a summary of the results obtained and presented in table III.
Figure 5 displays a multiple chart of the top-performing algorithms in both baseline and proposed methods to forecast future stock prices for each stock of the MASI Chemicals Index, with only a section of the data plotted on the abscissa. The graph provides a more intuitive visual representation of the volatility of the stocks composing MASI Chemicals Index. In addition, the chart gives an idea about of the model's predictive effectiveness, for both MAGHREB OXYGENE and SNEP stocks. From the figure, it is apparent that in most cases, the optimized LightGBM model is better suited to fit the fluctuation trend of stock prices, with predicted values closely approximating actual values.

To validate our approach, we conducted an evaluation of our model using additional data. Our focus was on the Building and Construction sector, which was one of the most volatile sectors during the studied period (Table I), and represented over 10% of the total market capitalization in the Moroccan stock market. The stocks included in this sector were AFRIC INDUSTRIES, ALUMINIUM DU MAROC, CIMENTS DU MAROC, COLORADO, JET CONTRACTORS, LAFARGEHOLCIM MAR and SONASID.
To conduct our evaluation, we utilized two distinct sets

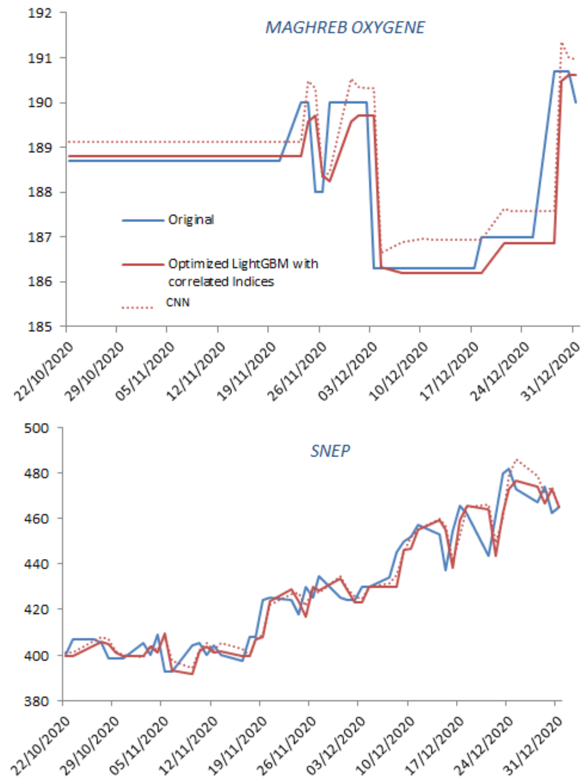


Figure 5. Predicted stocks closing price of MASI Chemicals Index

of hyperparameters: the default parameters of the model and the optimized parameters derived from the optimization process conducted specifically for the Chemical sector. By employing the same hyperparameters used in the previous evaluation, we aimed to assess the model's generalization capability specifically for the Moroccan market. Additionally, we improved our model by incorporating stocks from the MASI Banking and MASI Oil & Gas sectors, which are highly correlated with the Building and Construction sector. This inclusion allowed us to predict a total of 17 stocks using our model, providing valuable insights into the significance of incorporating the most correlated sector in the data. The summarized results of this evaluation are presented in Table V. The model labeled as "LightGBM" represents the model without parameter tuning, "Optimized LightGBM" refers to the model that underwent parameter adjustments based on the previous evaluation, and "Optimized LightGBM with correlated indices" denotes the prediction of stocks from the Building and Construction index with the inclusion of the MASI Banking and MASI Oil & Gas sectors. The results obtained from both Table III and Table V exhibit consistency, demonstrating similar magnitudes for the accuracy metrics. It is evident that the optimized LightGBM model with correlated indices surpassed other methods by achieving the lowest values for RMSE and MAE, as well as the highest R2 score among the evaluated approaches. Moreover, the incorporation of tuned

TABLE III. Performance results for MASI Chemicals Index

| Method | RMSE | MAE | R2 score |
|---|---|---|---|
| LightGBM | 0.043 | 0.031 | 0.82 |
| Optimized LightGBM | 0.037 | 0.026 | 0.86 |
| Optimized LightGBM with correlated indices | 0.033 | 0.020 | 0.99 |
| LSTM | 0.038 | 0.028 | 0.66 |
| GRU | 0.039 | 0.029 | 0.64 |
| CNN | 0.0375 | 0.027 | 0.69 |

TABLE IV. Performance and Execution Time comparison for MASI Chemicals Index

| Method | Approach | Predicted Stocks | Execution Time (minute) | R2 score |
|---|---|---|---|---|
| LightGBM | Global | 2 | 0.002 | 0.82 |
| Optimized LightGBM | Global | 2 | 0.14 | 0.86 |
| Optimized LightGBM with correlated indices | Global | 8 | 0.17 | 0.99 |
| LSTM | Local | 2 | 2.44 | 0.66 |
| GRU | Local | 2 | 2.10 | 0.64 |
| CNN | Local | 2 | 1.24 | 0.69 |

TABLE V. Performance results for Building and Construction Index

| Method | RMSE | MAE | R2 score |
|---|---|---|---|
| LightGBM | 0.049 | 0.034 | 0.93 |
| Optimized LightGBM | 0.044 | 0.030 | 0.96 |
| Optimized LightGBM with correlated indices | 0.040 | 0.027 | 0.98 |
| LSTM | 0.044 | 0.032 | 0.66 |
| GRU | 0.043 | 0.031 | 0.68 |
| CNN | 0.043 | 0.031 | 0.73 |

hyperparameters suggests the potential for generalization across the entire Moroccan market. Finally, these findings strongly support the notion that incorporating data from multiple stocks can significantly enhance the accuracy and generalization capabilities of LightGBM model for global stock price forecasting, especially in the context of high stock market volatility.

## 5. Conclusion

This paper introduces a global forecasting approach that uses LightGBM to predict stock prices in the Moroccan market during a period of market stress. The LightGBM model, optimized for accuracy, was applied to the most volatile sector, considering correlated stock data from other sectors. The study evaluated the predictive outcomes and execution time of this model against deep learning models, and the optimized LightGBM model demonstrated superior performance in both accuracy and execution time, outperforming baseline models. The findings suggest that incorporating data from multiple stocks can enhance the accuracy and generalization ability of LightGBM models for global stock price forecasting. The proposed approach provides the first evaluation of the effectiveness of a global forecasting method based on LightGBM for the Moroccan equity market during a period of significant market volatility.

There are several directions that can be explored following this project. For instance, the approach could be extended to cover the entire market. Furthermore, exploring the potential impact of integrating non-equity market data on the predictions could be considered. Additionally, it may be worthwhile to explore a comparison with contemporary machine learning algorithms to further enhance the analysis.

## References

[1] S. R. Baker, N. Bloom, S. J. Davis, K. Kost, M. Sammon, and T. Viratyosin, "The Unprecedented Stock Market Reaction to COVID-19," *The Review of Asset Pricing Studies*, vol. 10, no. 4, pp. 742–758, 07 2020.

[2] E. Alpaydin, "Introduction to Machine Learning," MIT Press, Cambridge, 2014.

[3] J. J. R. T. H. Strader, Troy J.; Rozycki and Y.-H. J. Huang, "Machine Learning Stock Market Prediction Studies: Review and Research Directions," *Journal of International Technology and Information Management*, vol. 28, no. 3, 2020.

[4] H. Hewamalage, C. Bergmeir, and K. Bandara, "Global models for time series forecasting: A simulation study," *Pattern Recognition*, vol. 124, p. 108441, 2022.

[5] C. S. Bojer and J. P. Meldgaard, "Kaggle forecasting competitions: