

Analysis and development of a predictive outlook model for Negara Bank of Indonesia stock using machine learning data training (Python)

Gherbali Abdelhamid ¹, Dekkiche Djamel ², Maachoumustefa ³

¹ Ahmed Zebana University - Relizan - Laboratory Of Financial Markets Management By The Application Of Mathematics and Computer Science , Algeria, abdelhamid.gherbali@univ-relizane.dz

² Oran Graduate school of Economics, , Algeria, djamel.dekkiche@ese-oran.dz

³ Ibn khaldoun tiaret,Algeria ,Maachoumustefa@yahoo.fr

Received: 24/11/2024

Accepted: 01/02/2025

Published: 17/03/2025

Abstract:

Predicting future prices of securities is a very complex process due to several variables and conditions beyond control, making it very difficult to forecast. Our study aims to analyze current data for the purpose of discovering it from December 2018 to October 2024 to predict the future stock prices of Bank Negara using current (historical) data and building a future guiding model useful for relevant parties, making it easier for them to make informed investment decisions about buying or not buying stocks, all assuming normal conditions without emergencies, relying on the analytical and foresight approach. The study concluded that the outputs facilitate the investment decision-making process and also the model's value in future predictions and improving returns from data processing to decision-making, in addition to developing advanced features to handle temporal data and flexibility. Relying on the XGBoost library, which uses deep learning algorithms to achieve more accurate results than traditional methods.

Keywords: stock prices; forecast; Bank Negara; deep learning; decision.

Jel Classification Codes: G21, G110.

1. INTRODUCTION

The financial and economic systems depend heavily on financial markets, especially the equity markets (Houwayji , 2024) The latter is intricate and uncertain (Charandeep & Jyotsna , 2023) Stated differently, it is constantly moving(dynamic). (Chowdhury, et al., 2024) As a result, stock market forecasting is one of the primary avenues for project funding (Ghassan, 2022) and extremely complicated for all stakeholders (financial analysts, investors, and researchers) (Bassant A. , Saad M. , & Saleh M, 2024) (Manohar , 2017) One of the primary avenues for project funding is machine learning, a subfield of artificial intelligence that focuses on developing statistical models and algorithms that enable computers to do tasks without explicit programming.

Machine learning is a branch of artificial intelligence that focuses on creating statistical models and algorithms that let computers carry out tasks without explicit programming. (Rudolph, 2018) The final one allows for future prediction (John Paul & Luca , 2018) Future forecasting based on data (Mehryar , Afshin, & Ameet , 2018) For this purpose, Python has gained popularity due to its adaptability (Charles , 2009) It is frequently used for forecasting and data analysis, and it has many libraries that make it easier to use (Harris , 2023) The literature has detailed a variety of methods, including as long term memory (LSTM) networks and recurrent neural networks (RNNs), (Krishna , Akshima , & Kamal , 2024). in forecasting. Our study's most significant contribution to the rest is the development of an indicative future model that helps all pertinent parties decide whether or not to invest. We trained a machine learning model to achieve the intended goals of forecasting future share prices for Indonesian Bank of Negara and assisting in the process of making well-informed decisions.

2. Literature review :

Table1. Previous studies on the topic of the study.

authors	Title	Insight
(Y., CH., K., G. , & S., 2023)	Stock Market Prediction Compliance Using AI and ML	In an article published in the INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM), researchers sought to build a daily price prediction model that offers better efficiency by relying on Python and machine learning techniques
(Ping & Yanfen , 2023)	Analysis of international stock data based on python	Through an article published in the journal Advances in Economics and Management Research, the researchers sought to analyse the linkage of Istanbul's equity indices to the rest of the countries through Python using various types of regression for the purpose of ascertaining the interconnectedness of others and found a correlation between different equity markets in a range of countries.
(Orlunwo & Ledesi , 2021)	Predicting Stock Price in Python Using TensorFlow and Keras	The researchers proposed a machine learning technique and trained it using publicly available inventory for the purpose of obtaining information that is used for accurate prediction in Python and by building an LSTM model based on data from the Yahoo platform.
(Dinesh, Nikhil ,	Stock Market	Through an article published in the International Journal of

Sanket , Piyush, Mahesh, & Dinesh, 2019)	Prediction Using Machine Learning In Python	Innovative Research in Technology, researchers sought to predict the future of the methods of equity costs using Python and LSTM, trying to find out the difference of this technique from other technologies of machine learning. The study concluded that 55.9% of the accuracy of the stock prediction is not high.
--	---	--

Source: Prepared by researchers based on site academic.

3. Materials and methods:

In order to accomplish the goals of the study, which was based on Python based models for a sample of Bank Negara Indonesia securities with 71 observations from December 2018 to October 2024 utilizing the Kaggle website, descriptive and analytical methodologies were employed.

4. Analyse the Data.

We will work on data analysis by presenting it, checking for missing data, understanding its components, conducting descriptive analyses, and inferring relationships through a series of steps sequentially.

4.1 Loading the Data:

The data under study contains 71 rows and 7 columns.

Table2. Study variables

	Date	Open	High	Low	Close	Adj Close	Volume
0	2018-12-31	4400.0	4700.0	4325.0	4537.5	3834.552246	996016400
1	2019-01-31	4550.0	4662.5	4325.0	4400.0	3718.353516	662752400
2	2019-02-28	4412.5	4825.0	4262.5	4700.0	3971.877686	517327000
3	2019-03-31	4712.5	5125.0	4700.0	4800.0	4056.385986	633546200
4	2019-04-30	4800.0	4800.0	3912.5	4200.0	3549.337402	1347774600
...	...	...	...	...	...	...	...
66	2024-06-30	4690.0	5150.0	4630.0	4970.0	4970.000000	1255034700
67	2024-07-31	4940.0	5550.0	4880.0	5350.0	5350.000000	1227187800
68	2024-08-31	5350.0	5850.0	5300.0	5350.0	5350.000000	1201449900
69	2024-09-30	5375.0	5650.0	5100.0	5650.0	5650.000000	567301000
70	2024-10-18	5550.0	5650.0	5550.0	5650.0	5650.000000	58870300

71 rows × 7 columns

Source: Prepared by researchers based on Jupyter.

4.2 Data Structure:

Date: The date on which the data point was documented (usually the final day of the month)

Open: The stock's price at the start of the trading session for that day.

High: The peak price the stock achieved during that trading session.

Low: The lowest price the stock reached during that trading session.

Close: The stock's price at the conclusion of the trading session.

Adj Close: The closing price adjusted for dividends and stock splits, often used for a more accurate historical performance analysis.

Volume: The total number of shares that were traded during that period.

4.3 Column Explanation:

Table 3: Categories of data

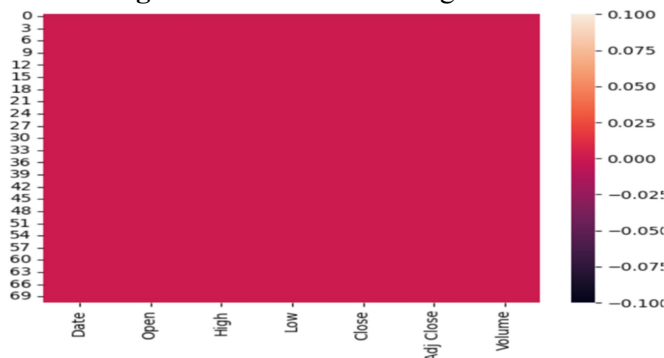
<class 'pandas.core.frame.DataFrame'>			
RangeIndex: 71 entries, 0 to 70			
Data columns (total 7 columns):			
#	Column	Non-Null Count	Dtype
--- ---			
0	Date	71 non-null	object
1	Open	71 non-null	float64
2	High	71 non-null	float64
3	Low	71 non-null	float64
4	Close	71 non-null	float64
5	Adj Close	71 non-null	float64
6	Volume	71 non-null	int64
dtypes: float64(5), int64(1), object(1)			
memory usage: 4.0+ KB			

Source:Jupyter Notebook

We have 7 columns with 71 statements and the data type varies between integers and names and numbers that have a decimal.

4.4 Ensuring the presence of missing data

Figure1: Check For Missing Values



Source: Prepared by researchers based on Jupyter.

Since the color is the same, there is no missing data

Are there duplicate values:

None False

Are there missing values:

None False

Through the application on Python, we find that there are no missing or duplicate data.

4.5 Summary Statistics:

Table4: ANALYSE DESCRIPTIVE

	count	mean	std	min	25%	50%	75%	max
Open	71.0	3.984155e+03	1.057730e+03	1.915000e+03	3.062500e+03	4.225000e+03	4.682500e+03	5.925000e+03
High	71.0	4.253486e+03	1.049318e+03	2.050000e+03	3.450000e+03	4.612500e+03	4.912500e+03	6.250000e+03
Low	71.0	3.736972e+03	1.058833e+03	1.485000e+03	2.793750e+03	3.937500e+03	4.418750e+03	5.725000e+03
Close	71.0	3.995845e+03	1.077858e+03	1.910000e+03	3.043750e+03	4.237500e+03	4.687500e+03	6.000000e+03
Adj Close	71.0	3.672184e+03	1.105766e+03	1.652788e+03	2.713525e+03	3.705716e+03	4.372304e+03	5.726346e+03
Volume	71.0	1.324455e+09	6.905044e+08	5.887030e+07	9.217491e+08	1.183421e+09	1.510818e+09	4.525855e+09

Source: Prepared by researchers based on Jupyter.

4.6 Price Statistics (Open, High, Low, Close)

Open : The stock was opened at an average price of 3984.15

High : Average of 4253.49 reached the peak price,

Low: 3736.97 was the minimum price.

Average closing: 3995.85

Day Low Day High : 1915 5925

Closes ranged from a low of 1910 to 6000. This kind of range indicates that there has been a significant fluctuation in the price over time. The degree of volatility. High standard deviation values in our example, 1077 indicate that prices are highly erratic, meaning that the stock price may deviate significantly from the norm. Because of the risk, investors should be concerned about this Dividends and stock splits with adjusted closing

prices Modification: 3672.18 is the split adjusted average closing price, That price is a more accurate representation of the stock price and accounts for dividends and stock splits Volume of Trades.

Average daily volume: Over the previous month, 1.32 billion shares were exchanged, which is a respectable volume of trading activity.

Volume: The range of shares is from 58.87 million to 4.53 billion. Trading tactics and liquidity may be impacted by this variation. Impact: Increased price stability is typically the outcome of high volumes, which show interest in the shares. Analysis on a Monthly and Annual Basis, Month: June is an average month (6.44). This implies that while performance varies monthly, the data set covers the full calendar year. Year: The dataset is very recent and takes into account the state of the market, with an average year of about 2021.

Risk and Variance: We know that there are good trades, but there are also concerns when there are significant price movements (high price std. dev.).

Trend identification: By graphing the closing price at various temporal intervals, one can visually ascertain whether the stocks are ascending (bull market), descending (bear market), or exhibiting lateral movement (range bound). Correlation- Understand market dynamics by comparing the open, high, low, and close prices in relation and make informed trading decisions.

Table5 : Correlation Matrix

	Open	High	Low	Close	Volume
Open	1.000000	0.978526	0.961026	0.942911	-0.536675
High	0.978526	1.000000	0.967760	0.973138	-0.470573
Low	0.961026	0.967760	1.000000	0.984578	-0.593621
Close	0.942911	0.973138	0.984578	1.000000	-0.530377
Volume	-0.536675	-0.470573	-0.593621	-0.530377	1.000000

Source: Prepared by researchers based on Jupyter.

4.7 Interpretation:

4.7.1 Diagonal Elements: Each variable perfectly correlates with itself (value of 1.0).

4.7.2 Positive Correlations:

Open, High, Low, and Close prices show strong positive correlations with each other (values range from 0.94 to 0.98). This indicates that as one of these prices increases, the others tend to increase as well.

4.7.3 Negative Correlations:

There are negative correlations between price variables (Open, High, Low, Close) and Volume. This suggests that higher trading volumes are associated with lower price levels, and vice versa.

Volume has a negative correlation with opening and closing prices. For example, the value of 0.536675 means that increased trading volume is associated with lower opening price.

6.4 General interpretation

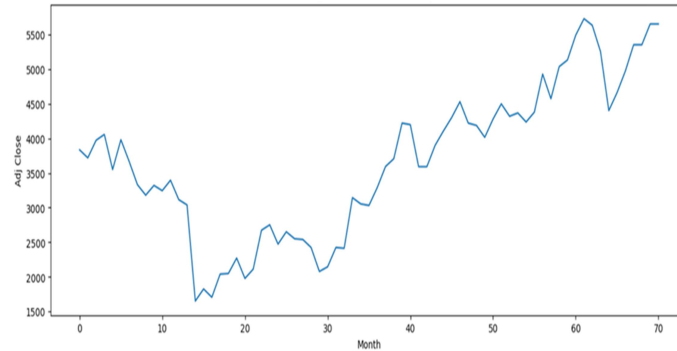
Price Relations: There is a strong correlation between opening, higher, lower, and closing prices. This means that price movements in the market are usually synchronized.

Trading Size: The negative relationship between trading volume and prices may indicate that in periods of strong price rise, there may be a decline in trading volume, or vice versa. This can be explained by the fact that in times of high volatility, trading volume may decrease, as investors may be reluctant to make purchasing or selling decisions.

5. Visualization.

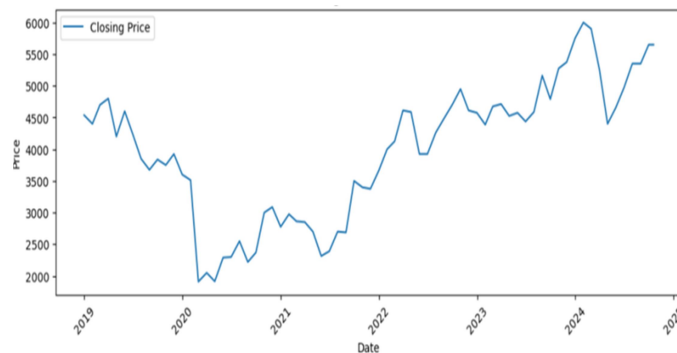
Analyzing the closing price curve of a particular bank's stock can provide information reflecting its performance in the market. By understanding trends, fluctuations and comparison with competitors, it makes it easier for investors to make decisions.

Figure2: Closing prices for Bank Negara (Monthly).



Source: Prepared by researchers based on Jupyter.

Figure3: Closing prices for Bank Negara (Yearly)

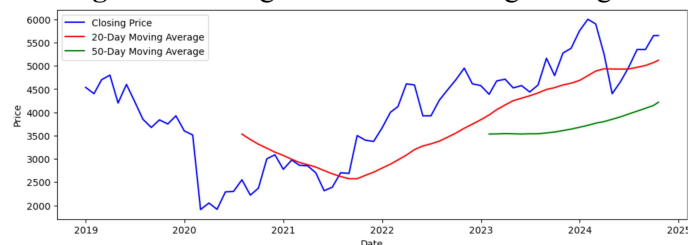


Source: Prepared by researchers based on Jupyter

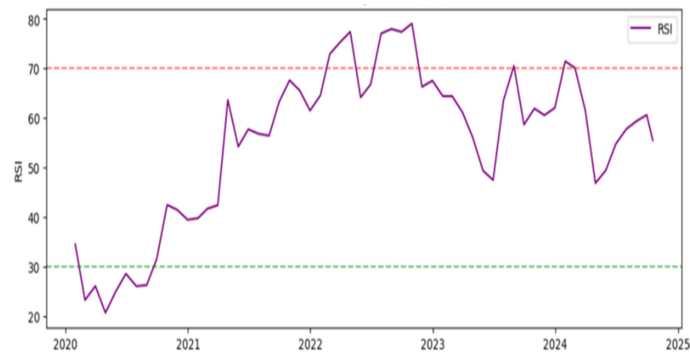
It is evident from the above data that closing prices are prone to significant fluctuations and disruptions, indicating instability brought on by a variety of causes, including banking policy changes and financial crises. Since the latter can be used to identify patterns, we employed the 20-day and 50-day moving averages for more clarity. In turn, this suggests an upward trend because the closing price was higher than the moving average.

Index of Relative Strength (RSI): The stock may be overpriced if the RSI is above 70. Because there was little demand for it during the Corona pandemic, its value was less than thirty. It surpassed 70 in 2023, which is a result of rising demand for it. rbought (overbought state), whilst a value below 30 would suggest an oversold state.

Figure 4: Closing Prices with Moving Averages



Source: Prepared by researchers based on Jupyter

Figure6:Relative Strength Index(RSI)

Source: Prepared by researchers based on Jupyter

6. Predicting future stock prices based on historical prices with a buy or hold decision:

We trained the model with machine learning techniques in order to reach the desired goals, as shown

Steps to build the model

Import libraries

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from xgboost import XGBRegressor
import matplotlib.pyplot as plt
```

Load the data

```
file_path = r"C:\Users\THINKPAD\Downloads\archive (1)\BBNI.JK_monthly.csv"
data = pd.read_csv(file_path)
```

Convert the 'Date' column to datetime

```
data['Date'] = pd.to_datetime(data['Date'])
data.set_index('Date', inplace=True)
```

Handle missing values

```
data.fillna(method='ffill', inplace=True) # or use data.ffill(inplace=True)
```

Feature engineering (lag features)

```
for i in range(1, 13): # Create lag features for the past 12 months
    data[f'lag_{i}'] = data['Close'].shift(i)
```

Drop NA values created by lagging

```
data.dropna(inplace=True)
```

Prepare features and target

```
X = data.drop(columns=['Close'])
```

```
y = data['Close']

# Split the data
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42,
shuffle=False)

# Train the model
model = XGBRegressor(n_estimators=100, learning_rate=0.1, max_depth=5)
model.fit(X_train, y_train)

# Predicting on test set
y_pred = model.predict(X_test)

# Evaluate the model
rmse = np.sqrt(mean_squared_error(y_test, y_pred))
print(f'RMSE: {rmse}')

# Future predictions for the next 12 months
last_data = data.iloc[-1].values[:-1] # Get the last row of data, excluding the target variable
future_predictions = []

for _ in range(12):
    # Prepare the next input data
    input_data = last_data.reshape(1, -1)
    pred = model.predict(input_data)
    future_predictions.append(pred[0])

    # Update last_data for the next iteration
    last_data = np.append(last_data[1:], pred) # Shift and append the new prediction

# Display future predictions
print("Future Predictions for the next 12 months:", future_predictions)

# Decide to buy or not based on predictions (arbitrary threshold)
threshold = np.mean(future_predictions) # Example threshold
decision = ['Buy' if pred > threshold else 'Hold' for pred in future_predictions]
print("Decision to buy or hold for the next 12 months:", decision)

# Plotting the results
plt.figure(figsize=(12, 4))
plt.plot(range(1, 13), future_predictions, marker='o', label='Future Predictions')
```

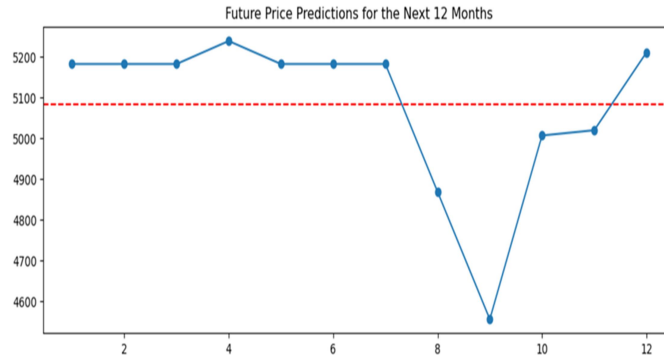
```
plt.axhline(y=threshold, color='r', linestyle='--', label='Threshold')
plt.title('Future Price Predictions for the Next 12 Months')
```

Table6:Decision to buy or hold for the next 12 months.

Date	Predicted Price	Decision
2024-10-31	5181.929	'Buy'
2024-11-30	5181.929	'Buy'
2024-12-31	5181.929	'Buy'
2025-01-31	5238.311	'Buy'
2025-02-28	5181.929	'Buy'
2025-03-31	5181.929	'Buy'
2025-04-30	5181.929	'Buy'
2025-05-31	4868.2974	'Hold'
2025-06-30	4555.71	'Hold'
2025-07-31	5006.499	'Hold'
2025-08-31	5019.436	'Hold'
2025-09-30	5210.272	'Buy'

Source: Prepared by researchers based on Jupyter

Figure7: Future Price Prediction for Next 12 Months



Source: Prepared by researchers based on *Jupyter*

7.Conclusion:

In conclusion, it can be said that the process of predicting future stock prices is very complex due to several factors beyond our control. We have built an indicative predictive futures model that is useful for decision-making based on historical data, which is unprecedented.

8.References :

- 1.Bassant A. , A., Saad M. , D., & Saleh M, E. (2024, January 12). Enhancing the Prediction of Stock Market Movement Using Neutrosophic-Logic-Based Sentiment Analysis. *journal of Theoretical and Applied Electronic Commere Research*(19), 116-134.
- 2.Charles , R. (2009). *Python for Everybody Exploring Data Using Python 3*.
- 3.Chowdhury, M., Nabi, N., Md Nasir, U., Mujiba , S., Hammed, E., Anik , M., et al. (2024, April 02). Deep Learning Models for Stock Market Forecasting: A Comprehensive Comparative Analysis Analysis. *Journal of Business and Management Studies*, 95-99.
- 4.Harris , W. (2023). *Introduction to Computer programming with Python*. Canada: Remix Athabasca University.
- 5.John Paul, M., & Luca , M. (2018). *Artificial intelligence for dummies*. Canada: willey.
- 6.Mehryar , M., Afshin, R., & Ameet , T. (2018). *Foundations of Machine Learning* (second edition ed.). Francis Bach, Editor.
- 7.Ping , D., & Yanfen , S. (2023). Analysis of international stock data based on python. *Advances in Economics and Management Research*, 05.
- 8.Rudolph, R. (2018). *Machine Learning Step-by-Step Guide To Implement Machine Learning Algorithms with Python*.
- 9.Charandeep , S., & Jyotsna , S. (2023, May 24). *Stock Market Prediction Using Python*.
- 10.Dinesh, A., Nikhil , P., Sanket , K., Piyush, N., Mahesh, K., & Dinesh, K. (2019, August). *Stock Market Prediction Using Machine Learning In Python*. *INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN TECHNOLOGY*, Volume 6(Issue 3).
- 11.Ghassan, A. (2022). *Equity Trading*. Abu Dhabi, United Arab Emirates: Arab Monetary Fund Building .
- 12.Houwayji , S. (2024, Feb 08). The Influence of Risk Management Practices on Financial Market Stability: Insights from Lebanon. *Dutch Journal of Finance and Management*, 7((1)).
- 13.Krishna , A., Akshima , A., & Kamal , K. (2024, july 10). *Predicting Stock Market Prices*

and Provide Recommendations. (©. C. Publishing, Ed.) International Journal of Computer Information Systems and Industrial Management Applications, 16, pp. 293-308.

14. Manohar , S. (2017). Mastering Machine Learning with Python in Six Steps. Bangalore, Karnataka,, India : Apress.
15. Orlunwo, P., & Ledesi , K. (2021, June). Predicting Stock Price in Python Using TensorFlow and Keras. International Journal of Research and Scientific Innovation (IJRSI), VIII(VI).
16. Y., S., CH., R., K., .., G. , R., & S., D. (2023, March). Stock Market Prediction Compliance Using AI and ML. INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM), 07(03).