



Pemodelan Matematis,
Departemen Matematika UI Semester Genap TA
2021/2022

Stock Trading Robot

Kelompok B3

Anatasya Oktaviani Handriati | Ashley Kainama |
Christantina Ethan Agustya | Halwatunnisa |
Kezya Samantha Sherryn | Marsha Putri Mahira | M. Shiqo Filla |
Muhammad Ichsanudin | Muhammad Reza Maullanna |
Prasetya Nugroho Hutomo | Richardy Lobo' Sapan

Daftar isi



Pembukaan

Latar Belakang Masalah
Rumusan Masalah
Tujuan
Asumsi & Batasan Masalah

Penutup

Analisis dan Kesimpulan
Progres saat ini
Lampiran

Isi

Metrik Evaluasi
Metode yang digunakan
ARIMA
LSTM

Latar Belakang Masalah

Saham memiliki harga yang fluktuatif, secara sekilas bergerak secara acak. Untuk membuat sebuah Robot trader, diperlukan sebuah metode untuk bisa memprediksi pergerakan harga saham.

Rumusan Masalah dan Tujuan



Rumusan Masalah

1. Metode apa yang Lebih baik digunakan oleh Stock Trading Robot dalam memprediksi harga saham?
2. Bagaimana cara memvisualisasikan hasil *forecasting* harga saham dengan metode tersebut?

Tujuan

1. Menentukan metode yang lebih baik digunakan oleh Stock Trading Robot dalam memprediksi harga saham
2. Memperlihatkan hasil forecasting harga saham dengan metode tersebut.

Parameter Keberhasilan

1. Mampu menjelaskan Metode yang digunakan
2. Mampu menampilkan *forecasting* saham 3 Perusahaan dan membandingkan dengan harga saham sesungguhnya

Asumsi & Batasan Masalah



Asumsi

1. Harga saham dipengaruhi oleh harga historisnya
2. Harga saham yang digunakan adalah harga penutupan
3. Keadaan suatu perusahaan akan terefleksi pada keadaan sahamnya (*technical analysis*)
4. Saham dari perusahaan sifatnya sah (bukan investasi bodong, dkk) terdaftar pada bursa efek Indonesia/yahoo finance

Batasan Masalah

1. Periode pengecekan setiap hari penutupan
2. Contoh diambil 3 saham penutupan dari perusahaan berbeda (BCA JKT, SIDOMUNCUL JKT, UNILEVER JKT)
3. Data diambil 10 tahun terakhir
4. Hari libur tidak dihitung pada pengambilan *data close*

Input dan Output Stock Trading Robot



Input:

- CSV data historis harga saham

Output:

- Tren dari saham
- Hasil *Forecasting* harga saham

Parameter: Rentang Pengambilan Data, Perbandingan Data Train:Test

Metode yang Digunakan Oleh Stock Bot



Terlepas dari volatilitas, harga saham bukan hanya angka yang dihasilkan secara acak. Jadi, mereka dapat dianalisis sebagai **barisan data waktu-diskrit**; dengan kata lain, pengamatan deret waktu (**time-series**) yang diambil pada titik waktu yang berurutan (biasanya setiap hari)

1. ARIMA (AutoRegressive Integrated Moving Average)

Persamaan forecasting **ARIMA** untuk **time series stasioner** adalah persamaan linier (yaitu, tipe regresi) di mana prediktor terdiri dari **lag variabel dependen dan/atau lag dari kesalahan ramalan**.

2. Long Short-term Memory (LSTM)

Singkatnya, komponen kunci untuk memahami model LSTM adalah **Cell State (Ct)**, yang mewakili ingatan *short term* dan *long term* internal sel.



Metode yang Digunakan Oleh Stock Bot

1. **ARIMA** (Auto Regressive Integrated Moving Average)

Model AR

- Model Auto Regressive (hanya AR) murni adalah model di mana Y_t hanya bergantung pada lag-nya sendiri. Artinya, **Y_t adalah fungsi dari 'lag dari Y_t '**.

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_1$$

I (Integrated)

- Integrated** di sini bermakna **mengurangi nilai saat ini dari sebelumnya** dan dapat digunakan untuk mengubah deret waktu menjadi **deret waktu yang stasioner**.
 - Differencing **orde pertama** → tren linier → transformasi $z_i = y_i - y_{i-1}$.
 - Differencing **orde kedua** → tren kuadrat → $z_i = (y_i - y_{i-1}) - (y_{i-1} - y_{i-2})$, dan seterusnya.

Model MA

- Demikian pula model Moving Average (hanya MA) adalah model di mana **Y_t hanya bergantung pada error perkiraan yang lag**

$$Y_t = \alpha + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q}$$

Metode yang Digunakan Oleh Stock Bot



1. ARIMA (Auto Regressive Integrated Moving Average)

- Model ARIMA adalah model di mana **deret waktu di-differencing setidaknya sekali ($d \geq 1$) untuk membuatnya stasioner dan menggabungkan suku-suku AR dan MA**. Sehingga persamaannya menjadi:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q}$$

- Prediksi **Y_t = Konstanta + Kombinasi linier Lag dari Y (hingga p lag) + Kombinasi Linier dari kesalahan prakiraan Lagged (hingga q lag)**
- Tiga bilangan bulat (p, d, q) biasanya digunakan untuk parametrize model ARIMA.
 - p**: jumlah suku autoregresif (order AR)
 - d**: jumlah differencing (order difference)
 - q**: jumlah suku Moving Average (order MA)

Metode yang Digunakan Oleh Stock Bot



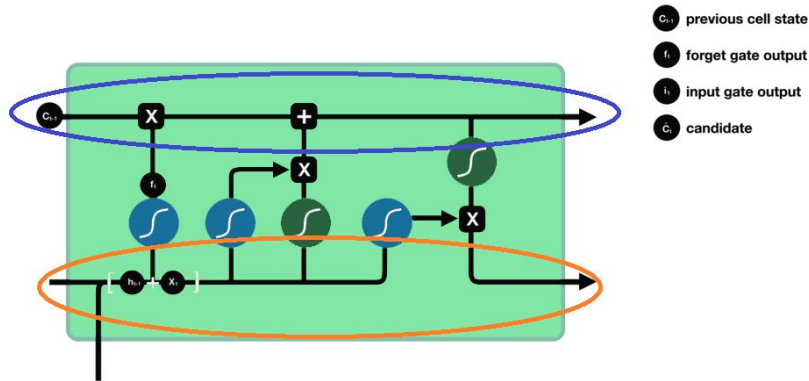
2. Long Short-term Memory (LSTM)

LSTM adalah singkatan dari **Long-Short Term Memory** yang mana adalah suatu pendekatan neural networking berulang. **LSTM** dapat memiliki beberapa lapisan tersembunyi dan saat melewati setiap lapisan, informasi yang relevan disimpan dan semua informasi yang tidak relevan dibuang di setiap sel. **LSTM** punya 4 gerbang utama: **FORGET** gate, **INPUT** gate, **OUTPUT** gate, dan **Cell** gate

Metode yang Digunakan Oleh Stock Bot



2. Long Short-term Memory (LSTM)



Cara Kerja

- Pada Metode LSTM, bagian **Short term** adalah data yang diambil pada proses iterasi pembelajaran data oleh AI yang nantinya akan dinilai oleh AI apakah data tersebut memiliki korelasi dengan data yang lainnya.
- Sedangkan **Long Term** pada LSTM, adalah hasil pembelajaran yang terakumulasi (hidden cell) yang menentukan apakah data dari short term dapat diteruskan untuk diakumulasikan atau tidak.

Metrik Evaluasi



- Karena prediksi harga saham pada dasarnya adalah masalah regresi, kita bisa menggunakan **RMSE (Root Mean Squared Error)** untuk menghitung tingkat akurasi model.
- Formulanya sbg berikut, dimana **N = jumlah titik waktu**, **At = harga saham aktual/sebenarnya**, **Ft = nilai prediksi/perkiraan**.

$$RMSE = \sqrt{\frac{1}{N} * \sum_{t=1}^N (At - Ft)^2}$$

- RMSE memberikan **selisih antara nilai prediksi dan nilai sebenarnya**
- **Makin kecil** nilai RMSE, maka **semakin baik** model



Pre Processing Data

1. Menginstall yfinance

```
[ ] pip install yfinance --upgrade --no-cache-dir
```

```
Collecting yfinance
  Downloading yfinance-0.1.70-py2.py3-none-any.whl (26 kB)
Requirement already satisfied: numpy>=1.15 in /usr/local/lib/python3.7/dist-packages (from yfinance) (1.21.5)
Requirement already satisfied: multitasking>=0.0.7 in /usr/local/lib/python3.7/dist-packages (from yfinance) (0.0.10)
Collecting lxml>=4.5.1
  Downloading lxml-4.8.0-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.manylinux_2_24_x86_64.whl (6.4 MB)
  |#####| 0.4 MB 12.7 MB/s
Requirement already satisfied: pandas>=0.24.0 in /usr/local/lib/python3.7/dist-packages (from yfinance) (1.3.5)
Collecting requests>=2.26
  Downloading requests-2.27.1-py2.py3-none-any.whl (63 kB)
  |#####| 63 kB 28.1 MB/s
Requirement already satisfied: python-dateutil>=2.7.3 in /usr/local/lib/python3.7/dist-packages (from pandas>=0.24.0->yfinance) (2.8.2)
Requirement already satisfied: pytz>=2017.3 in /usr/local/lib/python3.7/dist-packages (from pandas>=0.24.0->yfinance) (2018.9)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-packages (from python-dateutil>=2.7.3->pandas>=0.24.0->yfinance) (1.15.0)
Requirement already satisfied: charset-normalizer<=2.0.0 in /usr/local/lib/python3.7/dist-packages (from requests>=2.26->yfinance) (2.0.12)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests>=2.26->yfinance) (2021.10.8)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests>=2.26->yfinance) (2.10)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests>=2.26->yfinance) (1.24.3)
Installing collected packages: requests, lxml, yfinance
Attempting uninstall: requests
  Found existing installation: requests 2.23.0
  Uninstalling requests-2.23.0:
    Successfully uninstalled requests-2.23.0
Attempting uninstall: lxml
  Found existing installation: lxml 4.2.6
  Uninstalling lxml-4.2.6:
    Successfully uninstalled lxml-4.2.6
```

2. Download Resource di Yahoo Finance secara online

Yang diambil SIDO.JK, BBKA.JK, UNVR.JK

Cara pengambilan dataframe dari ke 3 perusahaan saham tersebut

```
[ ] import yfinance as yf
import pandas as pd

data_unilever=yf.download("UNVR.JK")
data_sidomuncu=yf.download("SIDO.JK")
data_BCA=yf.download("BBKA.JK")
```

data_unilever

```
[*****100*****] 1 of 1 completed
[*****100*****] 1 of 1 completed
[*****100*****] 1 of 1 completed
```

	Open	High	Low	Close	Adj Close	Volume
Date						
2003-09-03	615.0	675.0	605.0	675.0	383.379425	43162500

Pre Processing data



3. Cleansing dan Print Data Saham

Cleansing data

```
print(data_unilever["Close"].isnull().value_counts(), "\n")
print(data_sidomuncul["Close"].isnull().value_counts(), "\n")
print(data_BCA["Close"].isnull().value_counts(), "\n")
```

```
False    4629
Name: Close, dtype: int64

False    2062
Name: Close, dtype: int64

False    4430
True         1
Name: Close, dtype: int64
```

```
[ ] data_BCA=data_BCA.dropna()
```

```
print(data_BCA["Close"].isnull().value_counts(), "\n")
```

```
False    4430
Name: Close, dtype: int64
```

data_unilever

```
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
[*****100%*****] 1 of 1 completed
```

	Open	High	Low	Close	Adj Close	Volume
Date						
2003-09-03	615.0	675.0	605.0	675.0	383.379425	43162500
2003-09-04	690.0	750.0	675.0	685.0	389.059052	82850000
2003-09-05	685.0	700.0	680.0	690.0	391.898956	24055000
2003-09-08	690.0	690.0	665.0	680.0	386.219116	18007500
2003-09-09	680.0	680.0	660.0	665.0	377.699554	14075000
...
2022-04-01	3660.0	3690.0	3600.0	3620.0	3620.000000	26735700
2022-04-04	3630.0	3660.0	3540.0	3570.0	3570.000000	36823300
2022-04-05	3570.0	3690.0	3570.0	3680.0	3680.000000	30816700
2022-04-06	3680.0	3680.0	3560.0	3600.0	3600.000000	25193400
2022-04-07	3600.0	3610.0	3550.0	3610.0	3610.000000	7590100

Pre Processing data



4. Plotting Data

```
import matplotlib.pyplot as plt

plt.figure(figsize=(10,10))
plt.grid(True)
plt.plot(data_unilever["Close"], color="Green", label="Ini plot unilever")
plt.xlabel("Waktu")
plt.ylabel("Harga saham Closing")
plt.legend()
plt.show()
```

Hasil plot saham



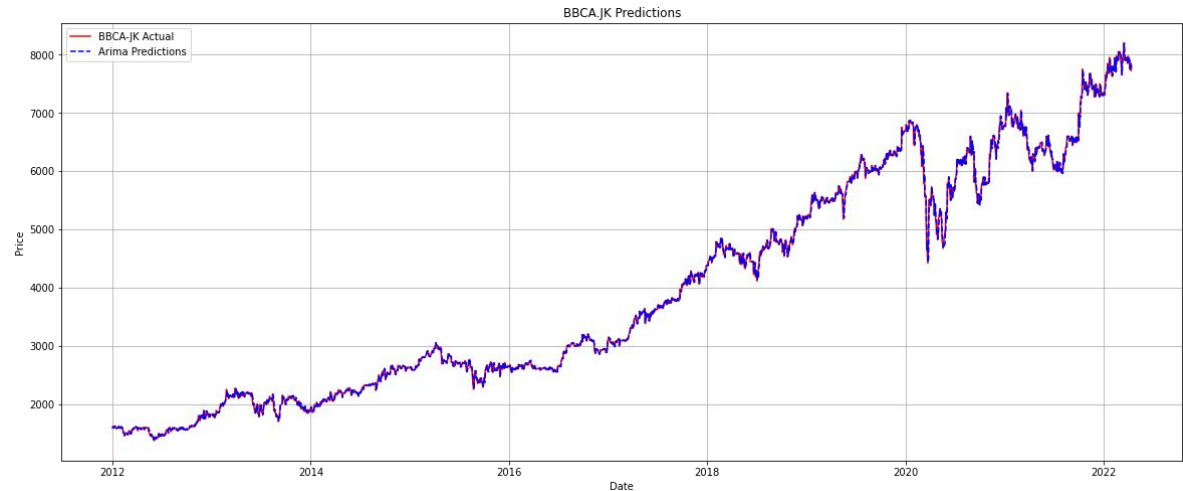
Hasil Forecasting harga saham dengan ARIMA



1. ARIMA

Plot perbandingan data asli dengan hasil prediksi menggunakan model ARIMA pada data test (2012 - sekarang)

(Saham **BBCA.JK**)



Root Mean Squared Error: **65.81**

Prediksi Harga Saham Tanggal 13/04/2022 : **7799.49**

Hasil Forecasting harga saham dengan ARIMA



1. ARIMA

Plot perbandingan data asli dengan hasil prediksi menggunakan model ARIMA pada data test (2012 - sekarang)

(Saham **UNVR.JK**)



Root Mean Squared Error: **131.44**

Prediksi Harga Saham Tanggal 13/04/2022 : **3520.07**

Hasil Forecasting harga saham dengan ARIMA



1. ARIMA

Plot perbandingan data asli dengan hasil prediksi menggunakan model ARIMA pada data test (2012 - sekarang)

(Saham **SIDO.JK**)



Root Mean Squared Error: **9.19**

Prediksi Harga Saham Tanggal 13/04/2022 : **907.73**

Hasil Forecasting harga saham dengan LSTM

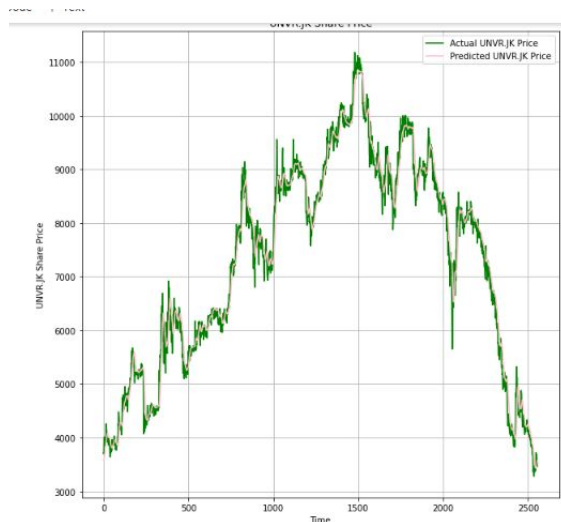


2. Long Short-term Memory (LSTM) Program

Input Pengambilan Data dari LSTM

1. Diambil dari 60 hari terakhir untuk memprediksi nilai saham berikutnya (bisa beberapa hari kedepan)

Hasil Running Program (UNVR.JK)



Maka hasil forecasting saham untuk tanggal ke 2022-04-13 adalah:
ini adalah hasil prediksinya [[3577.0105]]

Root Mean Squared Error:

294.5548

Prediksi Harga Saham

Tanggal 2022-04-13 :

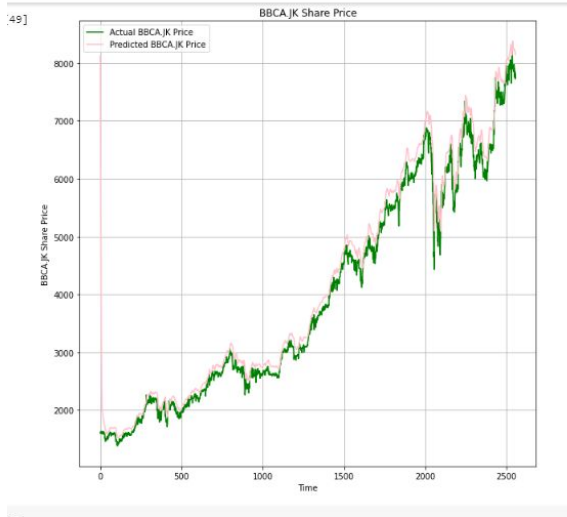
[[3577.3816]]

Hasil Forecasting harga saham dengan LSTM



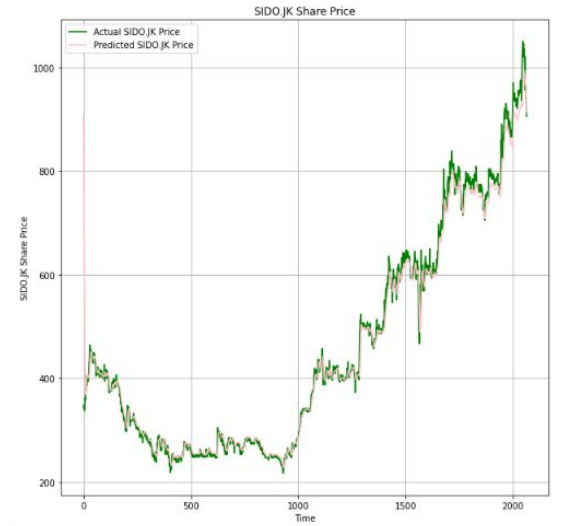
2. Long Short-term Memory (LSTM)

Hasil Running Program (BBCA.JK dan SIDO.JK)



Root Mean Squared Error: **298.415**

Prediksi Harga Saham Tanggal 2022-04-13 : **[[7816.318]]**



Root Mean Squared Error: **32.0572**

Prediksi Harga Saham Tanggal 2022-04-13 : **[[925.62836]]**

Perbandingan ARIMA dan LSTM



Perbandingan prediksi harga saham kedua metode dengan harga asli saham (Rupiah)

	BBCA.JK	UNVR.JK	SIDO
ARIMA	7799.49	3520.07	907.73
LSTM	7727.7373	3577.3816	925.62836
Harga Asli	7800	3440	905

Perbandingan RMSE kedua metode

	BBCA.JK	UNVR.JK	SIDO
ARIMA	65.81	131.44	9.19
LSTM	250.5662	294.5548	32.0572

Analisis & Kesimpulan



Terdapat dua metode yang digunakan oleh *stock trading bot*. Yang pertama adalah **AutoRegressive Integrated Moving Average (ARIMA)** dan yang kedua adalah **Long Short Term Memory (LSTM)**. ARIMA memanfaatkan ***persamaan linier tipe regresi**. LSTM memanfaatkan **komponen kunci yang disebut cell state (Ct)** yang mewakili **ingatan jangka panjang dan jangka pendek** internal sel.

Berdasarkan ketiga data historis harga saham penutupan yang digunakan, dapat dilihat bahwa baik untuk BBCEA.JK, UNVR.JK dan SIDO, **Arima lebih mendekati harga saham asli** dibandingkan dengan LSTM.

Dari metrik evaluasi sendiri, **ARIMA memiliki tingkat RMSE lebih baik daripada LSTM**, sehingga dapat disimpulkan bahwa **ARIMA** adalah metode yang cocok untuk pendekatan robot trading dalam memprediksi harga saham.

*

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q}$$

Progress Saat Ini



Bagian	Progress	Status	Comments
Pembagian Kelompok	100%	Selesai	Sudah membagi kelompok dan tugas
Perumusan Masalah	100%	Selesai	Sudah merumuskan masalah
Studi Literatur	100%	Selesai	Sudah melakukan studi literatur dari jurnal dan artikel online
Pengumpulan Data	100%	Selesai	Data dari yfinance
Pembuatan Program	80%	Ongoing	Sudah ada Program LSTM dan ARIMA, tetapi belum ada hasil analisis mengenai jual/beli saham bagi pengguna
Pengisian Lembar Kontrol Kelompok & Individu	66%	Ongoing	Sudah dilakukan sesuai dengan ketentuan
Laporan	15%	Ongoing	Sudah ada kerangka dan beberapa bagian pendahuluan

Referensi



1. Shekhar, Shraddha. 2021. "LSTM for Text Classification in Python". <https://www.analyticsvidhya.com/blog/2021/06/lstm-for-text-classification/#:~:text=it%20be%20used%3F-,What%20is%20LSTM%3F,patterns%20LSTMs%20perform%20fairly%20better>. Diakses pada 11-4-2022
2. (2021, August). Machinelearningplus.com. <https://www.machinelearningplus.com/time-series/arma-model-time-series-forecasting-python/>
3. Baheti, Pragati. 2022. "12 Types of Neural Network Activation Functions: How to Choose?". [https://www.v7labs.com/blog/neural-networks-activation-functions#:~:text=Tanh%20Function%20\(Hyperbolic%20Tangent\)&text=In%20Tanh%2C%20the%20larger%20the,output%20will%20be%20to%20%2D1.0](https://www.v7labs.com/blog/neural-networks-activation-functions#:~:text=Tanh%20Function%20(Hyperbolic%20Tangent)&text=In%20Tanh%2C%20the%20larger%20the,output%20will%20be%20to%20%2D1.0). Diakses pada 11-4-2022
4. Saeed, Mehreen. 2021. "A Gentle Introduction to Sigmoid Function". <https://machinelearningmastery.com/a-gentle-introduction-to-sigmoid-function/#:~:text=A%20sigmoid%20unit%20in%20a,the%20weighted%20sum%20of%20inputs>. Diakses pada 11-4-2022
5. *How to Build ARIMA Model in Python for time series forecasting?* (2022). ProjectPro. <https://www.projectpro.io/article/how-to-build-arma-model-in-python/544>
6. *How to Build ARIMA Model in Python for time series forecasting?* (2022). ProjectPro. <https://www.projectpro.io/article/how-to-build-arma-model-in-python/544>
7. Agung. (2017, March 21). *ARIMA SARIMA : Si Kembar dari Time Series*. Agung Budi Santoso | Membumikan Ilmu Pengetahuan; Agung Budi Santoso. <https://agungbudisantoso.com/arma-sarima-si-kembar-dari-time-series/>
8. Maklin, C. (2019, May 25). *ARIMA Model Python Example — Time Series Forecasting*. Medium; Towards Data Science. <https://towardsdatascience.com/machine-learning-part-19-time-series-and-autoregressive-integrated-moving-average-model-arma-c1005347b0d7>
9. (2021, August). Machinelearningplus.com. <https://www.machinelearningplus.com/time-series/arma-model-time-series-forecasting-python/>
10. *Predicting Stock Prices Using Machine Learning - neptune.ai*. (2021, July 29). Neptune.ai. <https://neptune.ai/blog/predicting-stock-prices-using-machine-learning>

The background of the image is a blurred, high-tech financial interface. It features a smartphone in the center-left, tilted slightly. The phone's screen shows a trading application with a candlestick price chart at the top. Below the chart, trade details are visible: 'Amount 2', 'Price 143.7684', 'Estimated cost: 143.1333', 'Profit: 0.0000%', and 'Available currency: 14'. At the bottom of the screen are large red 'SELL' and green 'BUY' buttons. The phone is set against a dark blue background with glowing green and red lines, suggesting a digital market or data visualization. Overlaid on the right side of the phone is the text 'Thank you' in a large, white, sans-serif font.

Thank you