APPLICATION OF MACHINE LEARNING FOR BITCOIN EXCHANGE RATE PREDICTION AGAINST US DOLLAR

**Ninuk Wiliani[1]; Rizki Hesananda[2]; Nidya Sari Rahmawati[3]
Erdham Hestiadhi Prianggara[4]**

Information System and Technology[1*)]; Information Technology[2,3,4,]
BRI Institute
http://bri-institute.ac.id/
ninukwiliani15@gmail.com[1*)]; hessananda@gmail.com[2];
nidyasarirahmawati24@gmail.com[3]; erdhamz@gmail.com[4]

(*) Corresponding Author

Abstract—Predicting a currency Exchange rate and performing analysis is an action to try to determine the price valuation of a currency or other financial instrument traded on an exchange platform. Bitcoin is a consensus network that enables new payment systems and fully digital money. Bitcoin is the first decentralized peer to peer payment network that is fully controlled by its users without any central authority or intermediary. From the user's point of view, Bitcoin is like cash in the internet world. Bitcoin can also be viewed as the most prominent triple bookkeeping system in existence today. The change in Bitcoin's behavior against the US dollar is influenced by many factors. Basic or economic factors that may be affected include inflation rates and money supply. In this study, data was collected by obtaining all data through the API provided by binance.com and labeled with the specified attribute. The modeling is done by using the rapidminer application. The process begins by taking training data that has been provided previously. The next stage is the data testing process, all operators that have been previously determined are connected and tested using the Linear Regression operator. The purpose of testing this data is to predict stock prices from the testing data that has been made by the Split Data operator, which is 19% of the total data that has been prepared.

Keywords: Prediction, Training, Testing, Bitcoin

Abstrak—*Memprediksi sebuah nilai tukar suatu mata uang dan melakukan analisa adalah sebuah aksi untuk mencoba menentukan valuasi harga dari suatu mata uang tersebut atau instrument finansial lain yang diperdagangkan disebuah platform exchanger. Bitcoin adalah jaringan consensus yang memungkinkan sistem pembayaran baru dan uang yang sepenuhnya berbentuk digital. Bitcoin merupakan jaringan pembayaran peer to peer desentralisasi pertama yang dikontrol sepenuhnya oleh penggunanya tanpa ada otoritas sentral ataupun perantara. Dari sudut pandang pengguna, Bitcoin seperti uang tunai di dunia internet. Bitcoin juga dapat dipandang sebagai sistem pembukuan tiga pencatatan paling menonjol yang ada saat ini. Perubahan perilaku Bitcoin terhadap dolar AS dipengaruhi oleh banyak factor. Faktor faktor dasar atau ekonomi yang mungkin terpengaruh termasuk nilai inflasi dan peredaran uang. Pada penelitian ini, data dikumpulkan dengan cara mendapatkan seluruh data melalui API yang disediakan oleh binance.com dan diberi label atribut yang ditentukan. Pemodelan dilakukan dengan memanfaatkan aplikasi rapidminer. Proses diawali dengan mengambil data training yang telah disediakan sebelumnya. Tahap selanjutnya adalah prises pengujian data, semua metode yang sudah ditentukan sebelumnya, disambungkan dan di uji dengan menggunakan metode Linier Regresion. Tujuan dari pengujian data ini adalah untuk memprediksi harga saham dari data testing yang sudah dibuat oleh operator Split Data, yaitu 19% dari total data yang sudah disiapkan.*

Kata Kunci : *Prediksi, Training, Testing Bitcoin*

INTRODUCTION

The exchange rate is the exchange rate of a country's domestic currency against other foreign currencies. Foreign exchange rates are determined in the foreign exchange market, which is the market for trading different currencies[1]. Currency exchange rate is one of the important macroeconomic variables, because changes in exchange rates affect stability and economic activity. Exchange rate volatility can affect the flow of capital and trade, because more than one currency will be involved in international economic

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

transactions, namely the domestic currency and foreign currency[2].

Predicting a currency exchange rate and performing Analysis is an action to try to determine the price valuation of a currency or other financial instrument traded on an exchanger platform[3].

Crypto assets are digital assets that use cryptography, peer-to-peer networks and public ledgers to manage new unit generation, verify transactions and secure transactions without intermediary intervention[4]. The definition of crypto assets has the same elements as the definition of cryptocurrency by Merriam Webster

This happens because crypto assets started as cryptocurrencies when Bitcoin was launched in 2009, but there have been many developments in crypto technology so that cryptocurrencies are not able to cover all existing crypto products[2]. Crypto assets have many functions and uses, depending on the purpose of the owner of the crypto asset and the type of crypto asset itself. Crypto assets can be used as a means of payment and can also be used as an investment instrument because their value is highly volatile.

Bitcoin is a consensus network enabling new payment systems and fully digital money. Bitcoin is the first decentralized peer to peer payment network that is fully controlled by its users without any central authority or intermediary[5]. From a user's perspective, Bitcoin is like cash in the internet world. Bitcoin can also be viewed as the most prominent triple record opening system currently available.

The change in Bitcoin's behavior against the US dollar is influenced by many factors. Basic or economic factors that may be affected include the inflation rate, money supply, capital inflows and outflows[6].

Linear regression will generally make predictions based on pre-existing values. This method constructs a function of linear equations, which are formed by constants and which represent straight lines, when converted into graphical form. After obtaining a linear equation, predictions can be made by replacing the test variable in the linear equation. After that the predicted value of the expected variable will be obtained. The linear regression method is often used in research related to prediction or forecasting, because many libraries in various programming languages and applications have implemented the method, one of which is the RapidMiner application.

Neural Network or Artificial Neural Network (ANN), which is one of the methods of machine learning commonly used for discrete, real, and vector problems[7]. Neural Network is also a model of the human nervous system in carrying out certain tasks

Artificial Neural Network (ANN) often referred to as Neural Network (NN) is a mathematical model or a computational model inspired by the structure and function of the biological aspects of human neural networks. Neural Network consists of a group of interconnected artificial nerves that process information[8].

MATERIALS AND METHODS

In this study, data was collected by obtaining all data through the API provided by Binance.com via https://api.binance.com/api/v3/ticker/price/symbol=BTCUSDT. Some of the attributes contained in the data to be used are shown in Table 1.

Table 1. Atribut Data Nilai Curs

Column	Type	Notes
Date	Date	Date
Open	Numeric	BTC-USD exchange rate
High	Numeric	BTC-USD exchange rate
Low	Numeric	BTC-USD exchange rate
Close	Numeric	BTC-USD exchange rate
Volume	Numeric	BTC-USD exchange rate

The data obtained will be divided into 2 parts, namely as training data and testing data. The Training data consists of more than 100 data on daily exchange rate changes which are from July 1, 2021 to November 30, 2021. The training data will be used in the linear regression method to obtain equations or predictive models. The second part is data testing which consists of 31 daily crucible change data which is in the range of July 1, 2021 to November 30, 2021. The processing of training data and data testing will use the Rapid Miner application[9].

In this study, the attributes used to obtain predictive models using linear regression methods and artificial neural networks are the values of Open, High, Low, and Close[10]. The date attribute is not used in the prediction model because it does not have a large influence on the prediction results. However, it is helpful to use as the ID of the value that appears later. In addition, the date attribute cannot be included in the calculation of linear regression and neural network methods because the value of the date is not a numeric.

The modelling is done by utilizing the RapidMiner application[11]. The process starts by taking the training data that has been provided previously by pressing right on the main process screen, selecting New Operator, then Import, then Data, and finally Read Excel, because the training data collected is in the form of an excel file. This process can also be seen in Figure 3.
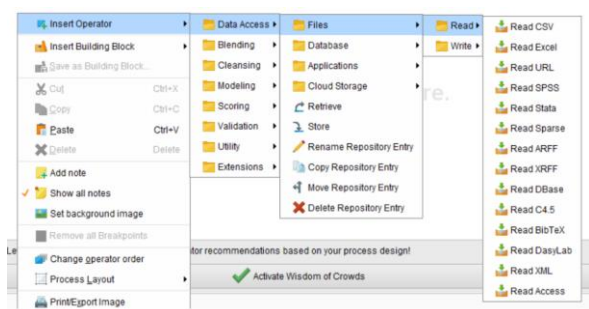
Figure 3. How to add Read Excel operator

After the Read Excel operator is added, the next step is to select the training data excel file that will be used. Click the Import Configuration Wizard button, as shown in Figure 4, then select the excel file to use. There are several steps after selecting a file, namely selecting data, annotations, and data attributes. Because the data used is as needed, there is no need for additional settings[12].
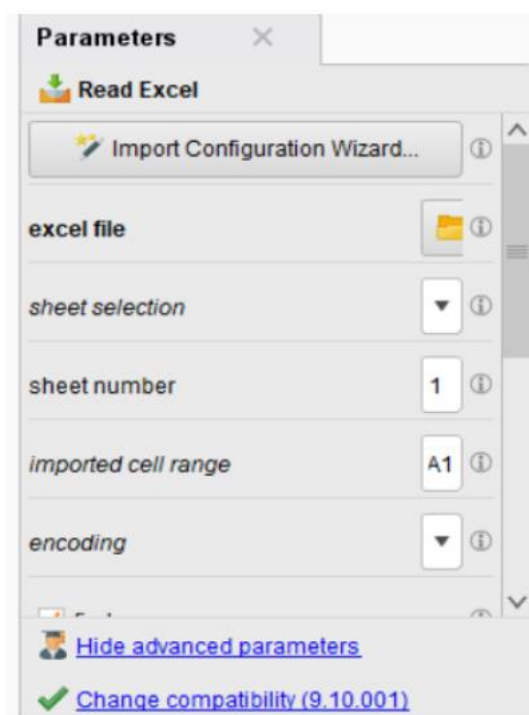

Figure 4. Selection of Training Data FIles

The next step after adding the Read Excel operator is to add the Linear Regression operator. The operator belongs to the Function category and is in the Modelling Folder then Predictive then to Function, as shown in Figure 5[13]. After adding the Linear Regression operator, the next step is to add the Split Data operator and then set the Enumeration that is in the operator as shown in Figure 6 Then add the ratio in the operator with a ratio of 0.81 and 0.19 to create training data and testing data.
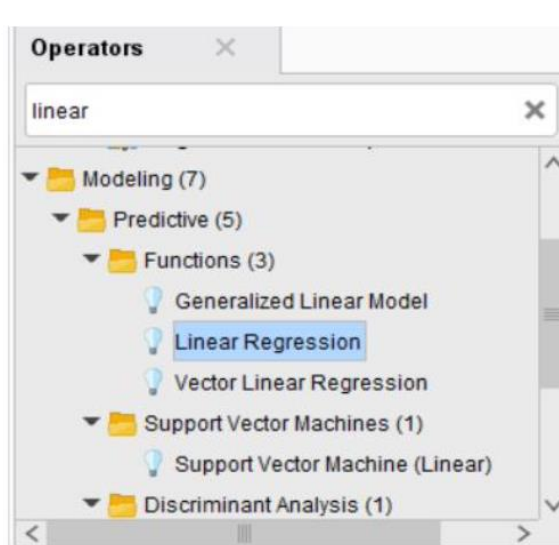
Figure 5. Liner Regression Operation Location
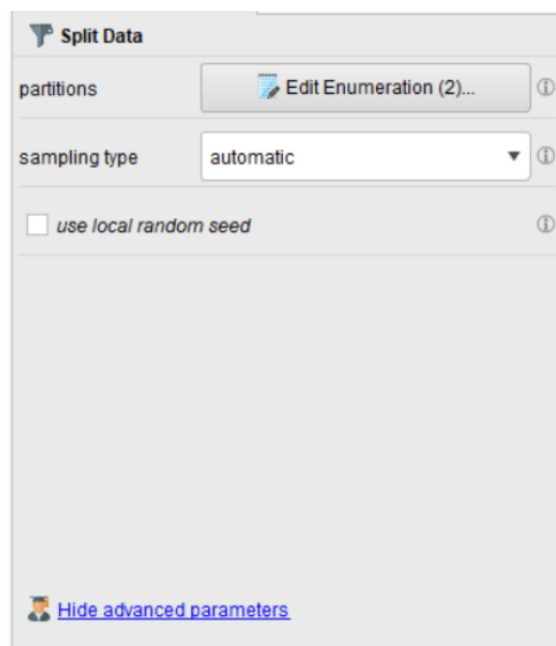
Figure 6. The Edit Enumeration Button

Then, add the Apply Model operator to apply the previously trained model using the training data on the unlabelled data (data testing). The goal is to get predictions on unlabelled data (testing data) that do not yet have a label. What needs to be considered is that the testing data must have the same order, type, and attribute roles as the training data[14].
And add the Performance operator. Because what we will predict is numerical data, the Performance operator used is the Performance (Regression) operator. The Performance (Regression) operator is used only for regression tasks. On the other hand, the Performance operator automatically determines the type of learning task and calculates

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

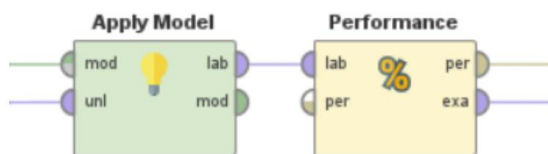the most common criteria for that type. The attribute display is shown in Figure 7.

Figure 7. Apply Model and Performance (Regression) Attributes

In addition to using Liner Regression, we also tested the data using the Neural Network operator (Neural Net). Neural Net is an operator that can be applied to an invisible data set to predict label attributes. Basically, the attributes used in this test are the same. It's just that in this test, the Linear Regression operator is replaced with the Neural Net operator as shown in Figure 8.
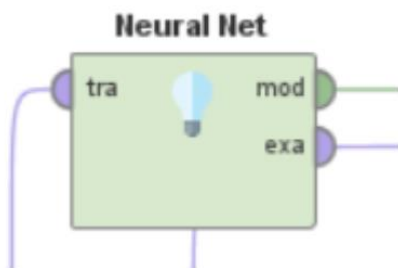
Figure 8. Neural Network Operator
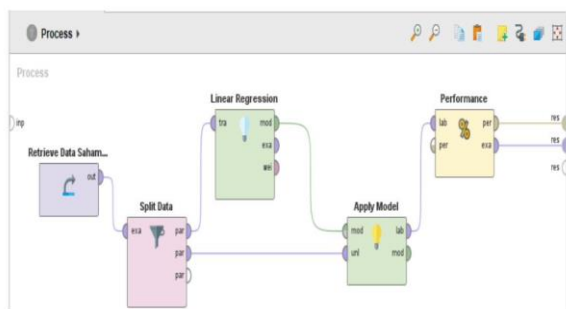
Figure 9. Linier Regression Testing

The next stage is the data testing process in RapidMiner. All operators that have been described previously, are connected as shown in Figure 9 above. The data will be tested using the Linear Regression operator. The purpose of testing this data is to predict stock prices from the testing data that has been made by the Split Data operator, which is 19% of the total data that we have prepared.

RESULTS AND DISCUSSION

Figure 10. The results of running data using Linear Regression

From this process, 29 examples prediction data were produced as shown in Figure 10. The data was then converted into Microsoft Excel Worksheet (.xlsx).
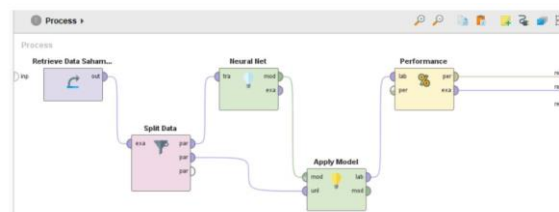

Figure 11. Neural Net Testing

The data testing process is carried out again using the Neural Net operator. All operators are connected as shown in Figure 11. The purpose of testing this data is the same as before, namely to predict stock prices from the testing data that has been made by the Split Data operator, which is 19% of the total data that we have prepared.

Figure 11. Results of running data using Neural Net

And from this process, 29 examples prediction data were generated as shown in Figure 10. As before, the data was then converted into Microsoft Excel Worksheet (.xlsx).

Figure 12. The results of statistical data using Linear Regression

From Figure 12 above, it is known that the prediction (Close) results obtained using the Linear Regression operator are at a minimum price of 31683,539 to a maximum of 64893,474 with an average price prediction of 48282,402. We can see the statistical development of the prediction data through the "Visualization" feature in the RapidMiner application as shown in Figure 13 below.

Figure 13. Results of data visualization using Liner Regression

Figure 14. The results of statistical data using Neural Net

From Figure 14 above, it is known that the prediction (Close) results obtained using the Neural Net operator are at a minimum price of 31509,256 to a maximum of 64029,358 with an average price prediction of 48271,341. We can see the statistical development of the prediction data through the "Visualization" feature in the RapidMiner application as shown in Figure 15 below.
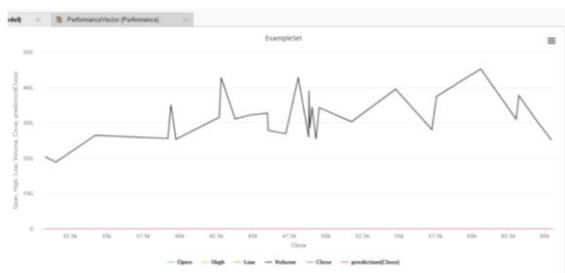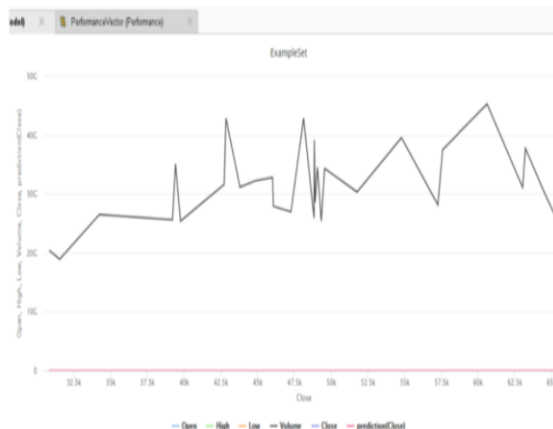

Figure 15. Results of data visualization using Neural Network

Due to the limitations of the visualization modeling in the RapidMiner application, we created our own script using the Python programming language (.py). Here is the content of the script that we made.

```
Algorithm: data-to-graph.py
import pandas as pd
import numpy as np
import plotly.graph_objects as go

data = pd.read_excel('./test-btc-usd.xlsx', parse_dates=True)

data.info()
print(data.head())

fig = go.Figure(data=[go.Candlestick(x=data['Date'],
open=data['Open'], high=data['High'], low=data['Low'],
close=data['Close'])])

fig.update_layout(yaxis_title="BTC-USDT", title="Normal
Data Chart")

fig.show()
```

Figure 16. Content of script

```
data2 = pd.read_excel('./btc-usdt-linear-regression.xlsx',
parse_dates=True)
fig2 = go.Figure(data=[go.Candlestick(x=data2['Date'],
open=data2['Open'], high=data2['High'], low=data2['Low'],
close=data2['Close'])])

fig2.update_layout(yaxis_title="BTC-USDT", title="Linear
Regression Prediction Data Chart")
fig2.show()

data3 = pd.read_excel('./btc-usdt-neural-network.xlsx',
parse_dates=True)
fig3 = go.Figure(data=[go.Candlestick(x=data3['Date'],
open=data3['Open'], high=data3['High'], low=data3['Low'],
close=data3['Close'])])
fig3.update_layout(yaxis_title="BTC-USDT", title="Neural
Network Prediction Data Chart")
fig3.show()
```

Figure 17. Content of script

From the script, it will produce 3 images obtained from the data that has been made by the RapidMiner application. Here are 3 image visualizations made from the script we created and used.

Figure 16. Normal Data Visualization

Figure 17. Data Visualization Using Linear Regression

Figure 18. Data Visualization Using Neural Network

Table 2. Comparison of Prediction (Close) Results

Operator	Minimum	Maximum	Average
Linier Regresion	31683,539	64893,474	48282,402
Neural Network	31509,256	64029,358	48271,341

From table 2 it can be seen that the prediction results of the Linear Regression operator and the Neural Net operator are not much different. However, the prediction result from Neural Network is smaller than Linear Regression.

Through the practicum that has been done, it can be said that The application of Neural Networks in making predictions takes a lot of time because they need to do a lot of experiments in determining the number of input data, the number of neurons and layers, the possibility of learning levels, testing with functions, and applying learning techniques to the formed network. So, compared to the use of Linear Regression, the use of Neural Network operators can be said to be the best result. Because with maximum Neural Network operator performance, a higher level of accuracy is obtained.

CONCLUSION

From the practical explanation above, it can be concluded that to predict stock prices, you can use the Forecasting method with the operators used are Linear Regression and Neural Network. To predict stock prices, you can use the Forecasting method with the operators used are Linear Regression and Neural Network. Prediction (Close) results obtained using the Linear Regression operator is at a minimum price of 31683,539 to a maximum of 64893,474 with an average price prediction of 48282,402. Prediction(Close) results obtained using the Neural Net operator are at a minimum price of 31509,256 to a maximum of 64029,358 with an average price prediction of 48271,341. The application of Neural networks is considered more effective than Linear Regression because, with a maximum performance from Neural Network operators, a higher level of accuracy is obtained.

REFERENCE

[1] R. Auer and R. Böhme, "Central bank digital currency: the quest for minimally invasive technology," *BIS Work. Pap.*, no. 948, 2021, [Online]. Available: www.bis.org.

[2] F. Diaz *et al.*, "Digital Menu Pada X Cafe Berbasis Desktop Graphical User Interface Dengan Visual Basic 2010 dan Microsoft Access," *J. Rekayasa Inf.*, vol. 6, no. 1, p. 43, 2017, doi:

10.1017/CBO9781107415324.004.

[3] Halimatusyakdiah, A. Kosim, and E. Meirawati, "Analisis Financial Distress pada Industri Kosmetik yang Terdaftar di Bursa Efek Indonesia (BEI) untuk Memprediksi Potensi Kebangkrutan Perusahaan," *Akuntabilitas J. Penelit. dan Pengemb. Akunt.*, vol. Vol 9 (2), no. 2, pp. 125–140, 2015.

[4] Y. Achsanty, H. Abrianto, and N. Wiliani, "Implementasi Metode Algoritma KVC Untuk Pengamanan Pesan," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.*, vol. 3, no. 3, p. 73, 2018, doi: 10.30645/jurasik.v3i0.62.

[5] U. Ravindran, "A Review Paper on Regulating Bitcoin Currencies," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 6, no. 4, pp. 4136–4140, 2018, doi: 10.22214/ijraset.2018.4682.

[6] G. Differences, "Real Time Defect Detection Method for rinted Images Based On Grayscale and Gradient Differences," vol. 11, no. 1, pp. 180–188, 2018, doi: 10.25103/jestr.111.22.

[7] L. Zhuang, Z. Zhang, and L. Wang, "The automatic segmentation of residential solar panels based on satellite images: A cross learning driven U-Net method," *Appl. Soft Comput. J.*, vol. 92, p. 106283, 2020, doi: 10.1016/j.asoc.2020.106283.

[8] I. P. Dhani and A. . G. S. Utama, "Pengaruh Pertumbuhan Perusahaan, Struktur Modal, Dan Profitabilitas Terhadap Nilai Perusahaan," *J. Ris. Akunt. Dan Bisnis Airlangga*, vol. 2, no. 1, pp. 135–148, 2017, doi: 10.31093/jraba.v2i1.28.

[9] K. Agroui, M. Pellegrino, and F. Giovanni, "Analysis Techniques for Photovoltaic Modules Based on Amorphous Solar Cells," *Arab. J. Sci. Eng.*, vol. 42, no. 1, pp. 375–381, 2017, doi: 10.1007/s13369-016-2050-5.

[10] H. Y. Kim and J. S. Cho, "Data governance framework for big data implementation with NPS Case Analysis in Korea," *J. Bus. Retail Manag. Res.*, vol. 12, no. 3, pp. 36–46, 2018, doi: 10.24052/jbrmr/v12is03/art-04.

[11] A. Sinha, O. S. Sastry, and R. Gupta, "Detection and characterisation of delamination in PV modules by active infrared thermography," *Nondestruct. Test. Eval.*, vol. 31, no. 1, pp. 1–16, 2016, doi: 10.1080/10589759.2015.1034717.

[12] T. K. Wen and C. C. Yin, "Crack detection in photovoltaic cells by interferometric analysis of electronic speckle patterns," *Sol. Energy Mater. Sol. Cells*, vol. 98, pp. 216–223, 2012, doi: 10.1016/j.solmat.2011.10.034.

[13] S. Deitsch *et al.*, "Automatic classification of defective photovoltaic module cells in electroluminescence images," *Sol. Energy*, vol. 185, pp. 455–468, Jun. 2019, doi: 10.1016/j.solener.2019.02.067.

[14] L. A. Sunjoyo, R. G. Santosa, and K. A. Nugraha, "Implementasi transformasi Haar Wavelet Untuk Deteksi Citra Jeruk Nipis Yang Busuk," *Informatika*, vol. 12, no. 2, pp. 165–173, 2016.