

Shares Price Forecasting using Simple Moving Average Method and Web Scrapping

Edra Arkananta Gultom^{1*}, Nurafni Eltivia², Nur Indah Riwijanti³

^{1,2,3} State Polytechnic of Malang

Corresponding Author: edraarkanantagultom@gmail.com¹⁾

Keywords : Shares Price, Forecasting, Simple Moving Average Method, Web Scrapping

Abstract:

The fluctuation of share prices in a secondary market allows investors/traders to gain profits through the difference in share prices (capital gain). In order to obtain these benefits, it is necessary to analyze before buying shares through fundamental and technical analysis. One of several methods in Technical Analysis is Simple Moving Average Method. This method can predict (forecast) share prices by calculating the moving average of the share price history. Historical share prices can be obtained in real-time using the Web Scrapper technique, so the results are more quickly and accurate. Using the MAPE (Mean Absolute Percent Error) method, the level of accuracy of forecasting can be calculated. As a result, the program could run successfully and display the value of forecasting and the level of accuracy for the entire data tested in LQ45. Besides, forecasting with a value of $N = 5$ has the highest level of accuracy, reaching 97,6 %, while the lowest one uses the value of $N = 30$, which is 95,0 %.

Introduction

Shares, according to the Big Indonesian Dictionary (KBBI), are proof of ownership of a limited liability company's share of capital, giving rights to dividends and others according to the size of the paid-up capital. Meanwhile, according to the Indonesia Shares Exchange, shares are a sign of the equity participation of a person or party (business entity) in a company or limited liability company. That party has claims on company income and claims on company assets and has the right to attend the General Meeting of Shareholders (GMS).

The primary market is where sales occur for the first time to the public or during an IPO (*Initial Public Offering*). While the secondary market occurs after the IPO, so transactions no longer directly involve the company, but only buying and selling between investors occur. The price of shares in the secondary market fluctuates. These price fluctuations allow investors to gain profits through the difference between the purchase and selling prices (*capital gain*). However, it also has the potential for loss if the selling price is less than the purchase price (*capital loss*).

Investors usually use two methods, namely Fundamental Analysis and Technical Analysis. Fundamental analysis is shares price analysis focusing on company performance and economic analysis affecting the company's future (Sutrisno, 2012). At the same time, Technical Analysis is a method for predicting future share prices and market trends by studying charts and market action in the past (Cahyono et al., 2011).

In this study, the authors use technical analysis because it is easier to implement, requiring only historical share price data. *Moving Average* is one part of the *time series* forecasting method (Nurlifa et al., 2017). The *Moving Average* method has two variants: the *Simple Moving Average* and the *Weighted/Exponential Moving Average*. The difference between the two is in the weighting technique. In the *Weighted Moving Average*, a higher weight is given to the period closer to the forecasted period (Sundari et al., 2015). This study will discuss making share price forecasting applications using the *Simple Moving Average* method. As for obtaining share price data, the *Web Scrapping* method will be used, namely, taking data directly from data provider sites.

Research Method

The object of research used in this study shares that they are members of LQ45, which can be seen in table 1.

Table 1. LQ45 Shares

Number	Code	Company Name
1	ADHI	Adhi Karya (Persero) Tbk.
2	ADRO	Adaro Energy Tbk
3	AKRA	AKR Corporindo Tbk.
4	ANTM	Aneka Tambang Tbk.
5	ASII	Astra International Tbk.
6	ASRI	Alam Sutera Reality Tbk.
7	BBCA	Bank Central Asia Tbk.

Number	Code	Company Name
8	BBNI	Bank Negara Indonesia (Persero) Tbk.
9	BBRI	Bank Negara Indonesia (Persero) Tbk.
10	BBTN	Bank Tabungan Indonesia (Persero) Tbk.
11	BKSL	Sentul City Tbk.
12	BMRI	Bank Mandiri (Persero) Tbk.
13	BSDE	Bumi Serpong Damai Tbk.
14	CPIN	Charoen Pokphand Indonesia Tbk.
15	ELSA	Elnusa Tbk.
16	EXCL	XL Axiata Tbk.
17	GGRM	Gudang Garam Tbk.
18	HMSP	HM Sampoerna Tbk.
19	ICBP	Indofood CBP Sukses Makmur Tbk.
20	INCO	Vale Indonesia Tbk.
21	INDF	Indofood SuksesMakmur Tbk.
22	INDY	Indika Energy Tbk.
23	INKP	Indah Kiat Pulp & Paper
24	INTP	Indocement Tunggal Prakasa Tbk.
25	ITMG	Indo Tambangraya Megah Tbk.
26	JSMR	Jasa Marga (Persero) Tbk.
27	KLBF	Kalbe Farma Tbk.
28	LPKR	Lippo Karawaci Tbk.
29	LPIF	Matahari Department Store Tbk.
30	MEDC	Medco Energi Internasional Tbk
31	MNCN	Media Nusantara Citra Tbk.
32	PGAS	Perusahaan Gas Negara (Persero) Tbk.
33	PTBA	Tambang Batubara Bukit Asam (Persero) Tbk.
34	PTPP	PP (Persero) Tbk.
35	SCMA	Surya Citra Media Tbk.
36	SMGR	Semen Indonesia (Persero) Tbk.
37	SRIL	Sri Rejeki Isman Tbk.
38	SSMS	Sawit Sumbermas Sarana Tbk.
39	TLKM	Telekomunikasi Indonesia (Persero) Tbk.
40	TPIA	Chandra Asri Petrochemical Tbk.
41	UNTR	United Tractors Tbk.
42	UNVR	Unilever Indonesia Tbk.
43	WIKA	Wijaya Karya (Persero) Tbk.
44	WSBP	Waskita Beton Precast Tbk.
45	WSKT	Waskita Karya (Persero) Tbk.

Source : idx.co.id/indeks-lq45

There are several methods for calculating forecasting accuracy, such as *Mean Forecast Error* (MFE), *Mean Absolute Deviation* (MAD), *Mean Square Error* (MSE), and *Mean Absolute Percent Error* (MAPE). *Mean Absolute Percent Error* (MAPE) will be used to calculate accuracy.

This is because this method is considered more accurate and easier to understand than other methods (Sanders et al., 2016).

The MAPE calculation formula is as follows:

$$\frac{1}{n} \sum \frac{|actual - forecast|}{|actual|} \times 100$$

n : the amount of data/record

actual : actual value

forecast : estimated value/forecasting

The calculation result is a percentage, so the smaller the percentage error, the more accurate the forecast will be. The formula can calculate the level of accuracy: Accuracy Level = 100% - MAPE.

Programs for Forecasting (*Machine Learning*)

The Programs for Forecasting (*Machine Learning*) using a flowchart can be seen in the following explanation:

- a) The program will ask the user to enter the shares code of the company that they want to forecast, the share price of which corresponds to the shares code on the Shares Exchange.
- b) Next, the program will access the shares price information provider site namely <https://finance.yahoo.com/> through the *Python* program.
- c) The program will download the intended site page and perform *Web Scrapping*.
- d) From process (c), the daily shares price of the company in question will be obtained, namely data date and shares closing. The data is then stored in the Array.
- e) Perform forecasting calculations using the *Simple Moving Average* method from previously obtained Array data.
- f) Perform forecasting accuracy calculations using the *MAPE* Method.
- g) Display forecasting results and forecasting accuracy levels.



Figure 1. Flowchart Programs for Forecasting (*Machine Learning*)

Web Scrapping

The steps for using *Web Scrapping*, in general, are as follows :

1. Create Scrapping Templates Study HTML documents from sites that will retrieve information.
2. Explore Site Navigation Learn navigation techniques on sites that will retrieve information to be embedded in the *Web Scrapper* application.
3. Automate Navigation and Extraction Automate the retrieval of information from the site based on the analysis in steps 1 and 2.
4. Extracted Data and Package History Obtain the desired results and store them in a local database for later processing as desired.

Shares price data is taken from the official Yahoo Finance website at the address: <https://finance.yahoo.com/>. Figure 2 below is an example of a screenshot of ADRO shares data taken from the Yahoo Finance website.

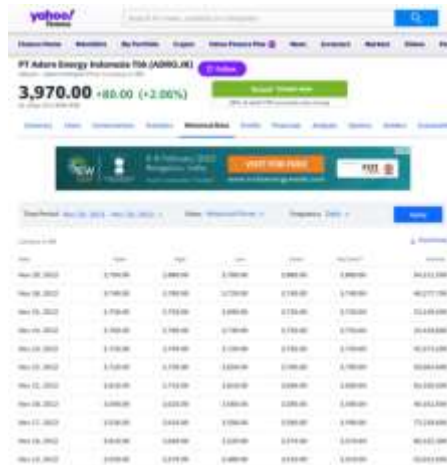


Figure 2. Example of a Yahoo Finance Site Screenshot

Result and Discussion

This study will divide the discussion into three sub-discussions: Program Running, Data Testing, and Forecasting Accuracy Analysis.

Program Running

The program that will run is a simple text-based program. The program uses the *Python* programming language and the *BeautifulSoup* library. *BeautifulSoup* is a library specifically used for *web scraping* in *Python*.

Some of the program code fragments are as follows:

```
import requests
import urllib.request
import time
from bs4 import BeautifulSoup
namaStok = input('Masukkan Kode Saham : ');
nilaiNString = input('Masukkan Nilai N : ');
nilaiN = int(nilaiNString)
url='https://finance.yahoo.com/quote/'+namaStok+'.JK/history?p='+namaStok+'.JK'
response = requests.get(url)
soup = BeautifulSoup(response.text, 'html.parser')
```

The program code above is a *script* to retrieve data from the Yahoo Finance site according to the share code entered.

```
for tbody in soup.select('tbody'):
    i = 0
    for tr in tbody.select('tr'):
        j = 0
        for td in tr.select('td'):
            data = td.text
            if (j==0):
                dataTgl = data
```

```

        if (j==5):
            dataHarga =
                data.replace(',','').
                replace('-','0')

            j = j+1

    if (dataHarga != '0') :
        hasil[0].append(dataTgl)
        hasil[1].append(dataHarga)
    i = i + 1

```

Next, the program will perform *web scrapping* and only retrieve the necessary data on the date and closing price of the shares. These two data are then entered into the Array variable for further forecasting calculations and the level of accuracy.

```

hasil[0].reverse()
hasil[1].reverse()
nilaiSum = 0
nilaiSumMAPE = 0
jmlSumMAPE = 0
nF = 0
nA = 0
hasil[3].append(0);
jmlMaks = (len(hasil[1]) - 1)
for k in range(0, jmlMaks):
    if (k < nilaiN) :
        nilaiSum= nilaiSum + float(hasil[1][k])
        hasil[2].append(nilaiSum)
        if (k == (nilaiN-1)):
            nF = float(nilaiSum/nilaiN)
        else :
            nF = 0
    else :
        nilaiSum = nilaiSum +
            float(hasil[1][k]) - float(hasil[1][k-nilaiN])
        hasil[2].append(nilaiSum)
        nF = float(nilaiSum/nilaiN)

    hasil[3].append(nF)

    if (k < nilaiN) :
        hasil[4].append(0)
        hasil[5].append(0)
    else :
        nF = float(hasil[3][k])
        nA = float(hasil[1][k])
        nilaiGap = abs(nF-nA)
        hasil[4].append(nilaiGap)
        hasil[5].append((nilaiGap/nA)*100)
        nilaiSumMAPE= nilaiSumMAPE+((nilaiGap/nA)*100)
        jmlSumMAPE = jmlSumMAPE + 1

```

$$k = k + 1$$

The code above is the program code to calculate the forecasting value and error rate (*MAPE*), and the average error rate. Next, the final forecasting results and accuracy level are displayed on the screen with the program code as follows:

```
print('-----')
print('Forecasting Simple Moving Average,
      dengan N = '+str(nilaiN))
print('Prediksi Untuk Besok :
      '+str(hasil[3][jmlMaks])+
      ', tingkat akurasi : %3.2f%%'%(100-(nilaiSumMAPE/jmlSumMAPE)) )
```

Data Testing

Program testing is carried out by calculating forecasting for all companies that are members of *LQ45* according to the list of companies in Table 1. The value of *N* is made differently, namely: *N* = 5, *N* = 10, *N* = 15, *N* = 20, *N* = 25, and *N* = 30.

Examples of test results are as follows:

Table 2. Forecasting Value with Different *N* Values

Shares Code	Forecasting Value					
	N=5	10	15	20	25	30
ADHI	498	497	510	527	537	532
ADRO	3970	4010	3990	3992	3990	3980
AKRA	1340	1419	14881	1544	1594	1660
ANTM	2040	2040	2047	2065	2086	2102
ASII	5700	5714	5756	5716	5727	5687
ASRI	167	172	176	178	183	187
BBCA	8650	8697	8716	8231	8286	8228
BBNI	9425	9455	9441	9415	9456	9445
BBRI	4970	4983	4923	4939	4962	4970
BBTN	1365	1319	1255	1219	1222	1215

Source : processed data, 2022

Based on the test results, it can be seen that the program can run well to calculate all *LQ45* shares.

Forecasting Accuracy Analysis

The accuracy of forecasting results can be measured using the *MAPE* (*Mean Absolute Percent Error*) method. Table 3 below shows the value of forecasting accuracy using the *MAPE* method. Accuracy calculations are carried out with different values of *N* so that it can be known which value of *N* is the most accurate.

Table 3. Forecasting Accuracy Level Using the MAPE Method

No	Shares Code	Forecasting Value					
		N=5	10	15	20	25	30
1	ADHI	97,8	96,8	96,4	96,3	96,2	95,9
2	ADRO	96,7	94,9	93,5	93,3	93,6	94,4
3	AKRA	97,5	96,2	95,3	95,0	94,9	94,9
4	ANTM	96,8	95,7	94,9	93,9	93,1	92,3
5	ASII	98,4	97,8	97,2	96,9	97,0	97,4
6	ASRI	98,3	97,3	96,6	96,3	95,8	95,1
7	BBCA	99,3	99,3	99,0	98,9	98,8	98,7
8	BBNI	97,8	96,9	96,1	95,9	95,8	95,7
9	BBRI	98,2	97,7	97,2	97,1	97,1	97,2
10	BBCA	97,0	95,6	94,5	94,1	93,8	93,3
11	BKSL	97,0	95,9	95,0	94,6	94,0	93,6
12	BMRI	98,2	97,8	97,2	96,9	96,6	96,3
13	BSDE	97,9	97,5	97,3	97,6	97,6	97,7
14	CPIN	96,3	95,5	94,3	93,3	92,4	91,7
15	ELSA	98,2	97,7	97,3	97,1	97,0	97,0
16	EXCL	97,9	97,9	98,0	97,9	97,9	97,7
17	GGRM	96,7	95,6	94,2	92,9	91,7	89,8
18	HMSP	97,0	96,3	95,5	94,3	93,1	91,3
19	ICBP	98,6	98,2	97,8	97,5	97,2	97,1
20	INCO	96,5	95,6	95,3	94,8	94,6	94,6
21	INDF	98,6	98,2	97,7	97,3	97,3	97,3
22	INDY	97,5	96,2	95,4	94,9	94,9	95,4
23	INKP	96,3	94,8	93,9	93,7	93,4	92,7
24	INTP	97,8	97,2	96,6	96,1	95,6	95,4
25	ITMG	96,5	94,9	93,4	92,4	92,8	93,8
26	JSMR	97,9	97,5	97,0	96,7	96,4	96,1
27	KLBF	98,6	98,1	94,4	96,9	96,2	95,7
28	LPKR	98,0	97,1	96,3	96,0	95,5	95,0
29	LPPF	96,7	95,2	94,2	93,0	92,0	91,3
30	MEDC	97,7	96,7	96,2	95,7	95,3	94,7
31	MNCN	97,9	97,3	96,7	96,7	96,5	96,3
32	PGAS	97,1	96,0	95,2	94,8	94,4	94,1
33	PTBA	97,3	96,3	95,4	94,6	94,3	94,4
34	PTPP	96,9	95,7	94,7	94,5	94,3	93,9
35	SCMA	96,5	95,3	94,4	94,0	93,8	94,2
36	SMGR	97,9	97,3	96,6	96,0	95,7	95,4
37	SRIL	98,2	97,3	96,6	96,0	95,1	94,1
38	SSMS	98,0	97,3	97,1	96,6	96,2	96,0
39	TLKM	98,7	98,4	98,1	98,0	98,0	97,9
40	TPIA	96,6	95,2	93,8	92,7	92,3	92,6
41	UNTR	97,2	95,7	94,5	93,8	93,7	94,5
42	UNVR	98,7	98,4	98,1	97,8	97,3	96,7

No	Shares Code	Forecasting Value					
		N=5	10	15	20	25	30
43	UNVR	98,7	98,4	98,1	97,8	97,3	96,7
44	WSBP	98,0	97,0	96,5	96,3	96,1	96,0
45	WSKT	97,2	96,1	94,9	94,5	94,1	93,6
	Average	97,6	96,7	95,9	95,5	95,2	95,0

Source : processed data, 2022

Conclusion

In this study, *web scrapping* using *Python* and the *BeautifulSoup* library went well. The accuracy rate of the *Simple Moving Average* forecasting calculation using the MAPE (*Mean Absolute Percent Error*) method on *LQ45* shares shows that calculations with a value of $N = 5$ provide the highest level of accuracy, namely 97.6%. At the same time, the lowest level of accuracy is using a value of $N = 30$, which is 95.0%. These results also show that the smaller the value of N , the better the accuracy.

References

- Departemen Pendidikan Nasional, Kamus Besar Bahasa Indonesia. Pusat Bahasa Edisi Ketiga, Jakarta: PT. Gramedia Pustaka Utama, 2005.
- Bursa Efek Indonesia "Saham, [Online]. Available ; <https://www.idx.co.id/produk/saham>.
- Sutrisno. Manajemen Keuangan Teori, Konsep dan Aplikasi. Yogyakarta: EKONISIA, 2012.
- Cahyono, Jaka. E, Strategi dan Teknik Meraih Untung di Bursa Saham. Jakarta: Elex Media Komputindo, 2011.
- Nurlifa, Alfian, Kusumadewi, Sri, Sistem Peramalan Jumlah Penjualan Menggunakan Metode Moving Average Pada Rumah Jilbab Zaky. Jurnal Inovtek Polbeng – Seri Informatika, Vol 2, No 1, 2017.
- Sundari, Shinta Siti, Susanto, Revianti, Wivia, Sistem Peramalan Persediaan Barang Dengan Weight Moving Average Di Toko The Kids 24. Konferensi Nasional Sistem & Informatika, 2015.
- Sulistiawan, Dedhy, Liliana, Analisis Teknikal Modern pada Perdagangan Sekuritas. Yogyakarta: Penerbit Andi, 2007.
- Sanders, Nada, Forecasting Fundamentals. New York: Business Expert Press, 2016.
- Turland, Matthew, Phparchitect's Guide to Web Scrapping with PHP. Alexandria: musketeers.me, 2010.
- Josi, A., Abdillah, L.A., & Suryayusra, Penerapan teknik web scrapping pada mesin pencari artikel ilmiah.ArXiv, abs/1410.5777, 2014.