

BAB IV

HASIL DAN PEMBAHASAN

4.1. Data

Data yang telah digunakan dalam penelitian ini diperoleh dari *website yahoo finance*. Situs tersebut memang menyimpan segala macam data saham perusahaan luar maupun dalam negeri. Dalam penelitian kali ini terdapat 5 saham perusahaan yang digunakan yaitu Telkom Indonesia dengan kode saham TLKM, PT. Kalbe Farma Tbk yang memiliki kode KLBF, Indofood CBP dengan kode saham ICBP, Aneka Tambang dengan kode saham ANTM, kemudian yang terakhir adalah Adaro Energy dengan kode saham ADRO. Data-data tersebut memiliki variabel *date*, *open*, *close*, *high*, *low*, *adjusting close*, dan *volume*. Masing-masing data berjumlah 248 data dalam jangka waktu 1 tahun dari Januari 2021 sampai Januari 2022. Tabel data saham yang digunakan dalam penelitian ini ditunjukkan pada Tabel 4.1 – 4.5.

Tabel 4.1. Data Saham TLKM

No	Date	Open	High	Low	Close	Adj Close	Volume
1	25/01/2021	3370.0	3410.0	3310.0	3370.0	3208.230469	118131800
2	26/01/2021	3360.0	3390.0	3260.0	3260.0	3103.510742	111414500
3	27/01/2021	3340.0	3380.0	3260.0	3380.0	3217.750488	159144700
4	28/01/2021	3280.0	3330.0	3230.0	3240.0	3084.470703	142648800
5	29/01/2021	3290.0	3290.0	3110.0	3110.0	2960.711182	191915100
6	01/02/2021	3130.0	3250.0	3040.0	3230.0	3074.950928	133929600
7	02/02/2021	3300.0	3330.0	3250.0	3270.0	3113.030762	114734600
8	03/02/2021	3250.0	3330.0	3220.0	3240.0	3084.470703	86433900
..
248	24/01/2022	4260.0	4310.0	4250.0	4300.0	4300.000000	70404500

Dari Tabel 4.1 yang merupakan tabel data saham Telkom Indonesia. Dapat dilihat pada tabel tersebut harga saham tersebut mengalami kenaikan dari harga Rp. 3370 menjadi yang terakhir adalah Rp.4.260. Hal tersebut memberi keuntungan kepada para pemegang saham Telkom Indonesia dalam waktu 1 tahun.

Tabel 4.2. Data Saham KLBF

No	Date	Open	High	Low	Close	Adj Close	Volume
1	25/01/2021	1615.0	1620.0	1535.0	1605.0	1574.939819	70622700
2	26/01/2021	1605.0	1620.0	1535.0	1545.0	1516.063599	79451300
3	27/01/2021	1545.0	1560.0	1480.0	1510.0	1481.719116	73126300
4	28/01/2021	1500.0	1565.0	1485.0	1515.0	1486.625488	80221200
5	29/01/2021	1525.0	1550.0	1455.0	1465.0	1437.561890	79625700
6	01/02/2021	1450.0	1515.0	1420.0	1490.0	1462.093628	71988500
7	02/02/2021	1510.0	1560.0	1490.0	1520.0	1491.531738	73428200
8	03/02/2021	1525.0	1575.0	1525.0	1560.0	1530.782593	58843900
..
248	24/01/2022	1705.0	1705.0	1620.0	1700.0	1700.000000	37893600

Pada Tabel 4.2 diatas yang menunjukkan data saham dari PT. Kalbe Farma Tbk menunjukkan bahwa mengalami kenaikan juga yang diduga disebabkan karena faktor pandemi yang mulai dilakukan penyesuai dan tidak separah pada tahun 2020 awal ketika saat pandemi baru dimulai. Sehingga berpengaruh terhadap harga saham PT. Kalbe Farma Tbk yang memang bergerak dalam bidang kesehatan.

Tabel 4.3. Data Saham ICBP

No	Date	Open	High	Low	Close	Adj Close	Volume
1	25/01/2021	9450.0	9575.0	9325.0	9500.0	9260.411133	8009500
2	26/01/2021	9575.0	9700.0	9400.0	9500.0	9260.411133	13587300
3	27/01/2021	9500.0	9650.0	9400.0	9500.0	9260.411133	9466000
4	28/01/2021	9500.0	9500.0	9350.0	9350.0	9114.193359	8804700
5	29/01/2021	9350.0	9400.0	9000.0	9100.0	8870.499023	14066800
6	01/02/2021	9000.0	9300.0	9000.0	9225.0	8992.346680	6846000
7	02/02/2021	9250.0	9575.0	9225.0	9500.0	9260.411133	12534000
8	03/02/2021	9525.0	9550.0	9400.0	9425.0	9187.302734	8510200
..
248	24/01/2022	8675.0	8725.0	8575.0	8600.0	8600.000000	2233500

Pada Tabel 4.3 merupakan data saham dari Indofood CBP. Perusahaan tersebut yang memproduksi produk mi instan yang terkenal, yaitu Indomie. Perusahaan tersebut tidak hanya memproduksi satu merek saja, banyak merek – merek lain yang diproduksi oleh Indofood CBP. Dari Tabel 4.3 tersebut terlihat bahwa harga *open* maupun *close* mengalami penurunan. Tentu saja memberikan kerugian bagi para pemegang saham ICBP. Faktor yang memungkinkan terjadinya

Fajar Akbardipura, 2022

PERBANDINGAN PERFORMA PERAMA PERAMALAN HARGA SAHAM 5 PERUSAHAAN PADA INDEKS LQ45 MENGGUNAKAN AUTOREGRESSIVE INTEGRATED MOVING AVERAGE DAN LONG SHORT – TERM MEMORY

UPN Veteran Jakarta, Fakultas Ilmu Komputer, S1 Informatika

[www.upnvj.ac.id – www.library.upnvj.ac.id – www.repository.upnvj.ac.id]

penurunan pada saham Indofood CBP adalah daya beli masyarakat yang menurun dikarenakan pandemi dan juga bahan baku yang naik.

Tabel 4.4. Data Saham ANTM

No	Date	Open	High	Low	Close	Adj Close	Volume
1	25/01/2021	2760.0	2960.0	2680.0	2870.0	2849.291504	1222991300
2	26/01/2021	2900.0	2930.0	2670.0	2670.0	2650.734619	786819300
3	27/01/2021	2620.0	2740.0	2490.0	2550.0	2531.600586	1341364800
4	28/01/2021	2470.0	2570.0	2380.0	2380.0	2362.827148	731854900
5	29/01/2021	2370.0	2440.0	2220.0	2220.0	2203.981689	450321200
6	01/02/2021	2090.0	2640.0	2070.0	2600.0	2581.239746	2186652700
7	02/02/2021	2600.0	2630.0	2420.0	2420.0	2402.538574	942434300
8	03/02/2021	2390.0	2520.0	2290.0	2330.0	2313.187988	948031000
..
248	24/01/2022	1950.0	1955.0	1875.0	1885.0	1885.000000	70289800

Tabel 4.4 diatas merupakan data saham dari perusahaan Aneka Tambang, perusahaan tersebut bergerak dalam bidang pertambangan. Harga saham dari Aneka Tambang cenderung mengalami penurunan. Dilihat juga dari volume transaksi saham tersebut mengalami penurunan dimana pada Januari 2021 transaksi saham ANTM lebih banyak dibandingkan Januari 2022.

Tabel 4.5. Data Saham ADRO

No	Date	Open	High	Low	Close	Adj Close	Volume
1	25/01/2021	1350.0	1350.0	1260.0	1295.0	1141.308228	142274300
2	26/01/2021	1290.0	1335.0	1230.0	1260.0	1110.462036	138670800
3	27/01/2021	1260.0	1285.0	1215.0	1245.0	1097.242188	119157300
4	28/01/2021	1230.0	1275.0	1190.0	1200.0	1057.582886	201003600
5	29/01/2021	1220.0	1255.0	1160.0	1200.0	1057.582886	192575300
6	01/02/2021	1205.0	1235.0	1165.0	1230.0	1084.022461	165569500
7	02/02/2021	1230.0	1265.0	1185.0	1200.0	1057.582886	176841000
8	03/02/2021	1200.0	1210.0	1155.0	1170.0	1031.143188	173080500
..
248	24/01/2022	2300.0	2350.0	2250.0	2270.0	2270.000000	94316300

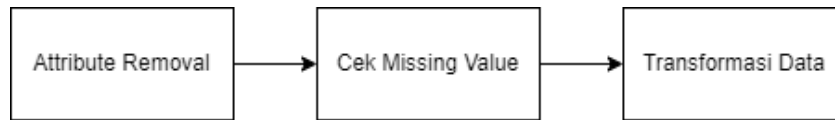
Pada Tabel 4.5 merupakan data saham Adaro Energy yang cukup terkenal karena dimiliki oleh Sandiaga Uno. Dari tabel data saham tersebut mengalami kenaikan harga yang sangat menguntungkan *trader* atau para pemegang saham

ADRO. Para perusahaan sekuritas sering sekali merekomendasikan emiten ADRO sebagai salah satu saham yang diperjual belikan karena mengalami kenaikan harga sehingga memberikan keuntungan bagi para pemegang saham tersebut. Kenaikan harga saham ADRO memiliki beberapa faktor seperti jumlah dividen atau bagi hasil yang menarik dan juga memiliki laporan keuangan yang positif.

Saat menganalisis data saham pada tabel di atas kita dapat melihat bahwa terdapat lima kolom harga yang berbeda dan satu kolom volume. Berikut penjelasan dari masing – masing atribut yang terdapat dalam Tabel 4.1 – 4.5.

1. Atribut *Open* mengacu pada harga di mana saham mulai diperdagangkan setiap hari.
2. Atribut *Close* adalah pesanan beli-jual terakhir yang dieksekusi untuk saham, kolom tersebut adalah harga yang disepakati setelah perdagangan sehari penuh. Bergantung pada likuiditas saham, ini mungkin terjadi dalam beberapa detik terakhir perdagangan atau jauh lebih awal di hari itu. Harga penutupan saham (*Close*) dianggap sebagai ukuran standar nilai untuk hari tertentu.
3. Atribut *High* menunjukkan harga tertinggi yang diperdagangkan pada hari transaksi yang ditentukan.
4. Demikian juga atribut *Low* menunjukkan harga terendah yang diperdagangkan pada hari transaksi yang ditentukan.
5. Volume adalah ukuran jumlah saham yang diperdagangkan dalam periode waktu tertentu. Setiap kali saham dibeli dan dijual, volume perdagangan dicatat. Jika saham yang sama dibeli dan dijual beberapa kali, volume saham dalam setiap transaksi dicatat. Volume tidak mencerminkan jumlah saham yang tersedia, itu adalah ukuran perputaran saham, hitungan semua transaksi yang terjadi pada setiap saham.
6. Kolom terakhir pada data harian di atas adalah *Adj Close*. Kolom tersebut menggabungkan perubahan harga ekuitas yang dihasilkan dari pembayaran dividen, pemecahan saham dan dari saham baru yang diterbitkan.

4.2. Pra-proses data



Gambar 4.1. Flowchart Pra-proses Data

Pengumpulan data yang diperlukan sudah dilakukan, kemudian langkah selanjutnya adalah melakukan *preprocessing* data yang akan diolah. Dari Gambar 4.1 dapat dilihat pada tahapan *preprocessing* untuk penelitian ini yang pertama adalah melakukan penghapusan atribut atau *attribute removal* terhadap variabel yang akan digunakan dalam penelitian, hal tersebut berlaku untuk kedua algoritma yang akan dibandingkan dalam penelitian ini.

Setelah proses *attribute removal* selesai, selanjutnya adalah memeriksa apakah data yang digunakan terdapat *missing value*. Secara kebetulan, data saham yang digunakan tidak memiliki *missing value*, sehingga tidak perlu mengatasi adanya *missing value* pada data. Kemudian untuk proses selanjutnya karena pada penelitian ini mengikuti alur proses dari KDD (*Knowledge Discovery in Database Process*) maka akan dilakukan proses transformasi data. Proses transformasi data umumnya dilakukan dengan mengolah data apapun menjadi format csv atau format yang dapat diolah oleh algoritma yang digunakan. Pada penelitian ini proses transformasi data tidak dilakukan karena sumber data saham yang diperoleh dari situs *yahoo finance* sudah berformat csv yang dimana format tersebut langsung dapat diolah oleh program yang akan digunakan untuk tujuan penelitian ini.

Atribut yang akan disimpan untuk dilakukan pemodelan adalah *date* dan *open*. Dari lima atribut yang dimiliki data aslinya hanya *date* dan *open* yang digunakan, hal ini dikarenakan dalam penelitian ini akan memprediksi harga pembuka (*open*) saham dan kedua atribut tersebut akan berpengaruh terhadap pemodelan. Atribut lainnya seperti *close*, *adj close*, dan *volume* tidak akan digunakan agar dapat menghasilkan harga prediksi yang akurat. Hasil dari *preprocessing* dapat dilihat pada Tabel 4.6.

Tabel 4.6. Hasil *attribut removal* dataset saham

No	Date	TLKM (Open)	KLBF (Open)	ICBP (Open)	ANTM (Open)	ADRO (Open)
1	25/01/2021	3370.0	1615.0	9450.0	2760.0	1350.0

No	Date	TLKM (Open)	KLBF (Open)	ICBP (Open)	ANTM (Open)	ADRO (Open)
2	26/01/2021	3360.0	1605.0	9575.0	2900.0	1290.0
3	27/01/2021	3340.0	1545.0	9500.0	2620.0	1260.0
4	28/01/2021	3280.0	1500.0	9500.0	2470.0	1230.0
5	29/01/2021	3290.0	1525.0	9350.0	2370.0	1220.0
6	01/02/2021	3130.0	1450.0	9000.0	2090.0	1205.0
7	02/02/2021	3300.0	1510.0	9250.0	2600.0	1230.0
8	03/02/2021	3250.0	1525.0	9525.0	2390.0	1200.0
...
248	24/01/2022	4260.0	1705.0	8675.0	1950.0	2300.0
Keterangan :						
TLKM		: Telkom Indonesia		ANTM		: Aneka Tambang
KLBF		: PT. Kalbe Farma Tbk		ADRO		: Adaro Energy
ICBP		: Indofood CBP				

4.3. Uji Stasioneritas

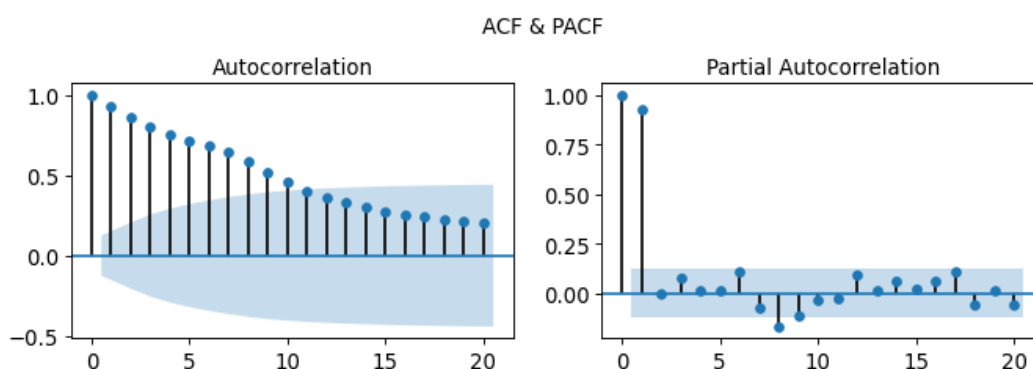
Untuk metode ARIMA tentu saja harus dilakukan uji stasioneritas terlebih dahulu supaya tidak terdapat data yang mengalami kenaikan dan penurunan yang terlalu mencolok. Dalam penelitian ini uji stasioneritas dilakukan dengan uji ADF menggunakan persamaan (7). Berikut adalah hasil uji data TLKM dengan ADF dapat dilihat pada Tabel 4.7.

Tabel 4.7. Uji stasioner ADF data saham

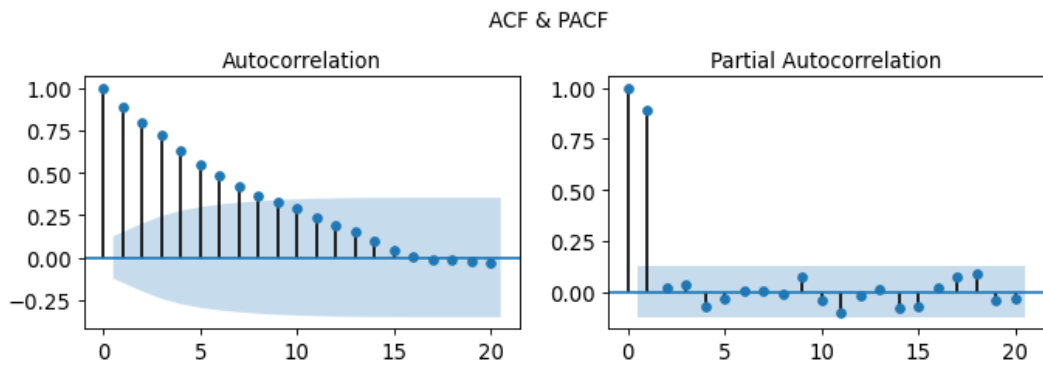
Result Dickey-Fuller Test	TLKM	KLBF	ICBP	ANTM	ADRO
Test Statistic	0.2245	-0.8742	-3.3396	-2.5965	0.0993
p-value	0.9736	0.7963	0.0131	0.0937	0.9660
Lags Used	2.0000	2.000	0.0000	3.0000	9.0000
Number of Observations Used	245.0000	245.0000	247.0000	244.0000	238.0000
Critical Value (1%)	-3.4574	-3.4573	-3.4571	-3.4574	-3.4581
Critical Value (5%)	-2.8734	-2.8734	-2.8733	-2.8734	-2.8737
Critical Value (10%)	-2.5731	-2.5730	-2.5730	-2.5731	-2.5732

Result Dickey-Fuller Test	TLKM	KLBF	ICBP	ANTM	ADRO
Keterangan :					
TLKM	: Telkom Indonesia		ANTM	: Aneka Tambang	
KLBF	: PT. Kalbe Farma Tbk		ADRO	: Adaro Energy	
ICBP	: Indofood CBP				

Seperti yang sudah dijelaskan pada Sub Bab 3.5.3 bahwa untuk melihat kestasioneran data perlu dilihat nilai statistik, nilai kritis, dan p-value. Apabila nilai statistik $ADF < \text{nilai kritis}$ dan $p\text{-value} < 0,05$ maka dinyatakan stasioner dan jika keadaan menunjukkan sebaliknya maka data tersebut tidak stasioner atau jika salah satu tidak memenuhi syarat untuk kestasioneran maka dilihat terlebih dahulu dari plot ACF dan PACF. Dapat dilihat pada Tabel 4.7, dataset TLKM, KLBF, ANTM, ADRO memiliki nilai $p\text{-value} > 0.05$. Data ANTM memiliki nilai statistik lebih kecil dibanding *critical value 10%*. Dapat terlihat juga bahwa nilai statistik ADF pada dataset ICBP bernilai lebih kecil dibandingkan *critical value 5%* dan *critical value 10*. Dari kondisi yang telah disebutkan sebelumnya, bahwa pada dataset ICBP dan ANTM harus dilakukan analisis pada plot ACF dan PACF untuk memastikan kestasioneran data yang dapat dilihat pada Gambar 4.2 dan 4.3.



Gambar 4.2. Plot ACF dan PACF Data ICBP

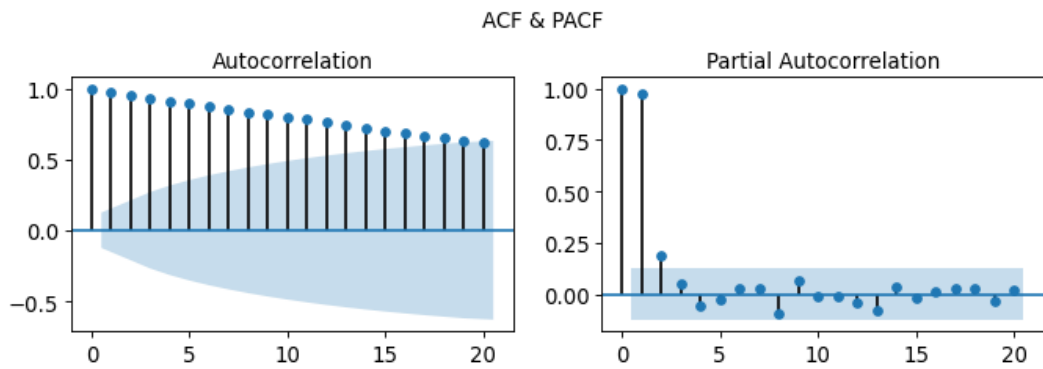


Gambar 4.3. Plot ACF dan PACF Data ANTM

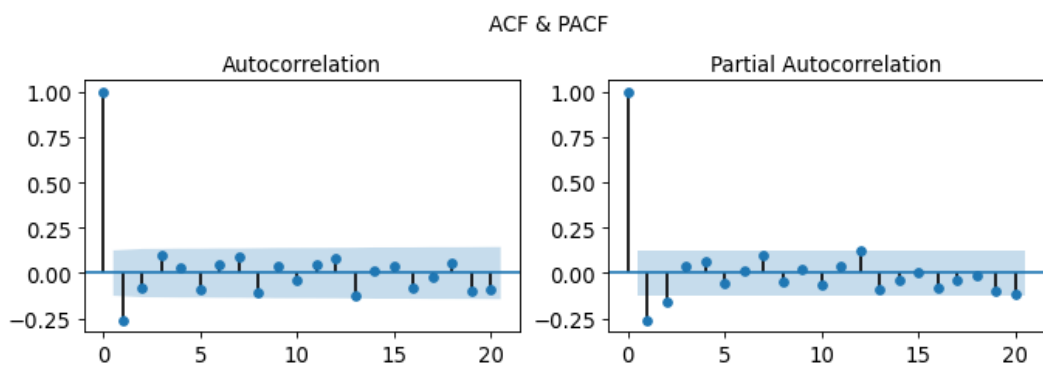
Dari Gambar 4.2 terlihat plot ACF mengalami penurunan yang lambat menuju nol, sehingga dari plot tersebut dinyatakan bahwa dataset ICBP memiliki data yang tidak stasioner. Untuk Gambar 4.3 terjadi penurunan secara cepat atau eksponensial menuju nol yang merupakan ciri dari suatu data yang stasioner dapat disimpulkan bahwa dataset ANTM memiliki data yang stasioner. Kemudian untuk sisa dataset lainnya dinyatakan tidak stasioner dari nilai statistik > nilai kritis dan $p\text{-value} > 0,05$.

4.4. Differencing

Setelah dilakukan pengujian stasioneritas maka sudah diketahui bahwa terdapat data yang tidak stasioner dan juga stasioner. Untuk kasus data yang tidak stasioner perlu dilakukan proses *differencing*. Plot ACF dan PACF pada data TLKM sebelum dilakukan proses *differencing* dapat dilihat pada Gambar 4.4, kemudian hasil dari proses *differencing* dapat dilihat pada Gambar 4.5. Proses plot ACF dan PACF itu sendiri menggunakan persamaan yang telah ditulis pada persamaan (5) dan persamaan (6). Sehingga, dari hasil persamaan tersebut diperoleh lah nilai – nilai yang ditandai oleh titik – titik biru pada plot ACF dan PACF yang nantinya titik – titik biru yang dinamakan lag akan digunakan untuk pemodelan.



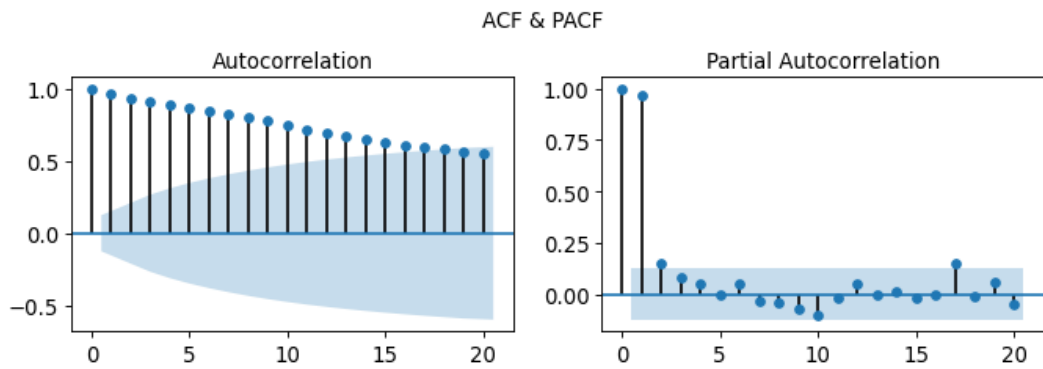
Gambar 4.4. ACF & PACF Sebelum Differencing Data TLKM



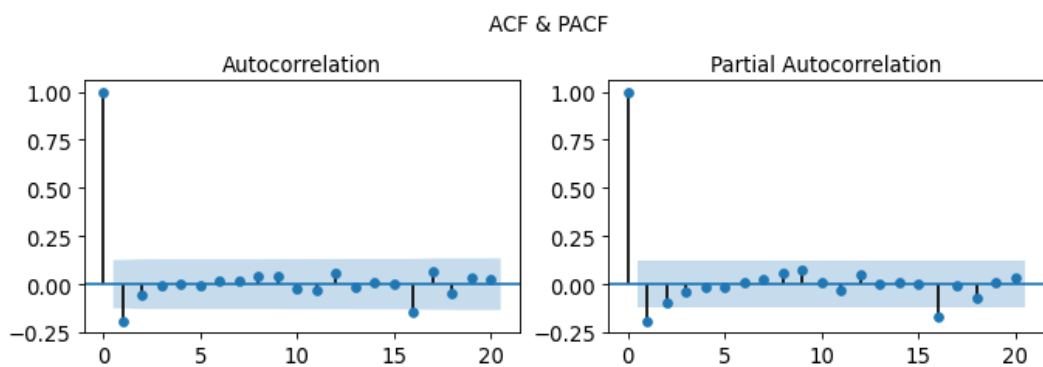
Gambar 4.5. ACF & PACF Setelah Differencing Pertama Data TLKM

Dapat dilihat perbedaannya sebelum *differencing* dan setelah *differencing*. Proses tersebut cukup dilakukan sekali saja karena dilihat dari Gambar 4.5 bahwa data sudah stasioner di proses *differencing* yang pertama, sehingga dari proses tersebut diperoleh parameter $d = 1$ yang akan digunakan dalam model ARIMA(p,d,q).

Selanjutnya untuk plot ACF dan PACF pada data KLBF sebelum dilakukan *differencing* dapat dilihat pada Gambar 4.6, lalu hasil dari proses *differencing* dapat dilihat pada Gambar 4.7.



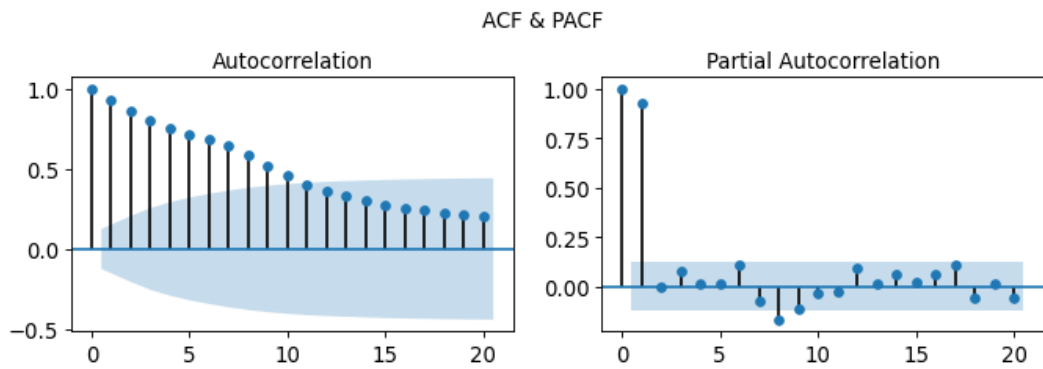
Gambar 4.6. ACF & PACF Sebelum Differencing Data KLBF



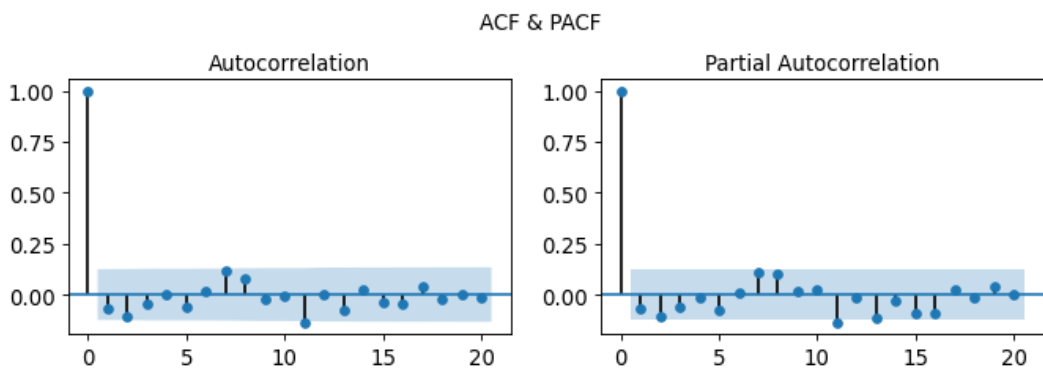
Gambar 4.7. ACF & PACF Setelah Differencing Pertama Data KLBF

Sama seperti plot data TLKM, proses *differencing* tersebut cukup dilakukan sekali saja karena dilihat dari Gambar 4.7 bahwa data sudah stasioner di proses *differencing* yang pertama, sehingga dari proses tersebut diperoleh parameter $d = 1$ yang akan digunakan dalam model ARIMA(p,d,q).

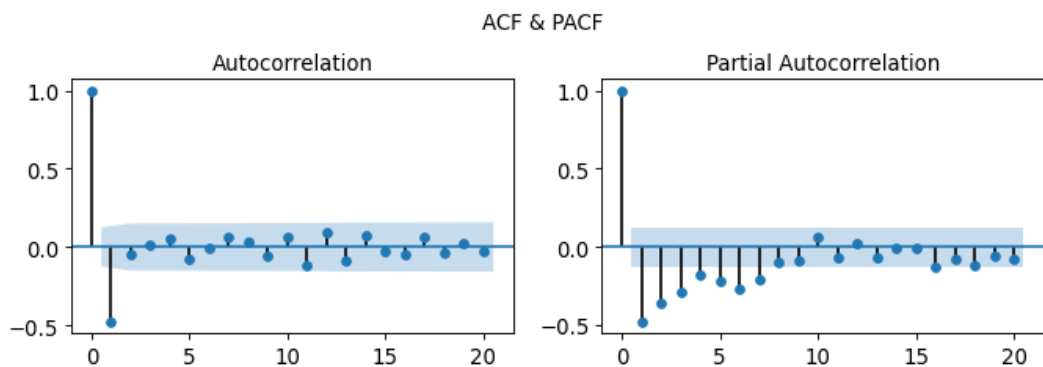
Selanjutnya untuk plot ACF dan PACF pada data ICBP sebelum dilakukan *differencing* dapat dilihat pada Gambar 4.8, lalu hasil dari proses *differencing* dapat dilihat pada Gambar 4.9 dan 4.10.



Gambar 4.8. ACF & PACF Sebelum Differencing Data ICBP



Gambar 4.9. ACF & PACF Setelah Differencing Pertama Data ICBP

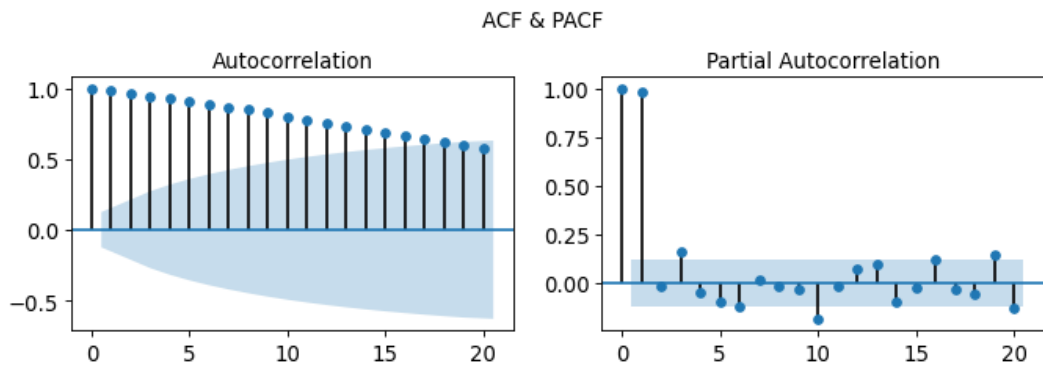


Gambar 4.10. ACF & PACF Setelah Differencing Kedua Data ICBP

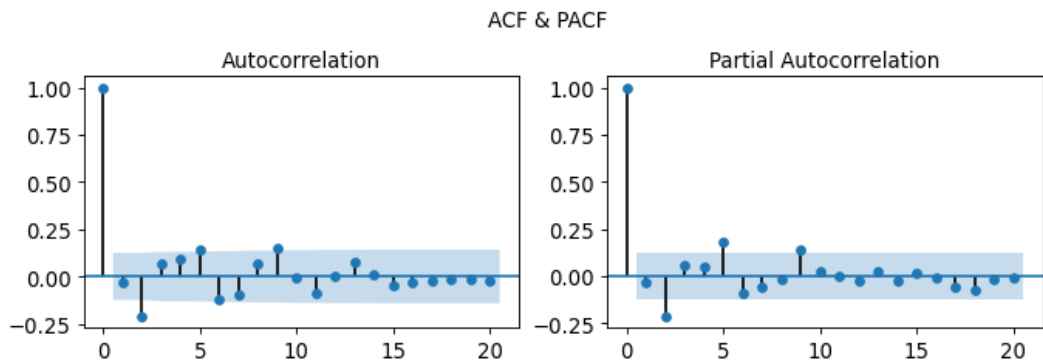
Dalam kasus data ICBP terdapat perbedaan dari data sebelumnya, proses *differencing* tersebut cukup dilakukan dua kali. Jika dilihat dari Gambar 4.9 bahwa data sudah stasioner di proses *differencing* yang pertama, tetapi tidak ada lag yang melewati area biru atau area batas signifikan selain lag ke-0, sehingga *differencing*

dilakukan sekali lagi. Dari proses *differencing* tersebut diperoleh 2 kasus dengan parameter $d = 1$ dan $d = 2$ yang akan digunakan dalam model ARIMA(p,d,q).

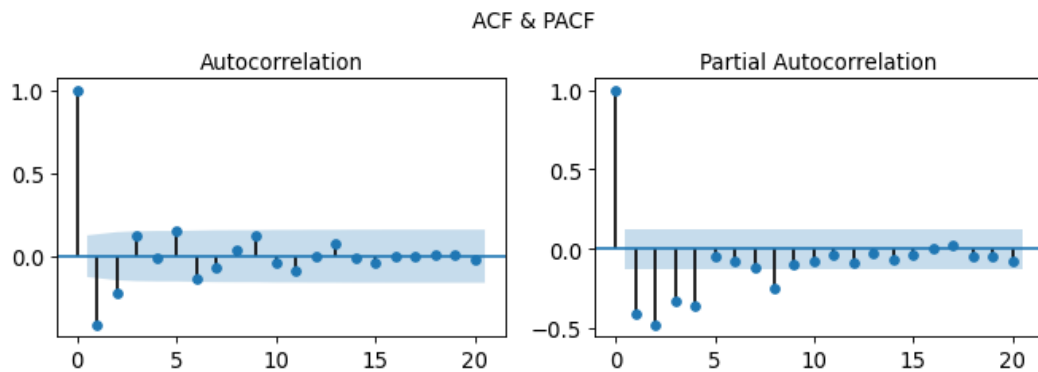
Selanjutnya untuk plot ACF dan PACF pada data ADRO sebelum dilakukan *differencing* dapat dilihat pada Gambar 4.11, lalu hasil dari proses *differencing* dapat dilihat pada Gambar 4.12 dan 4.13.



Gambar 4.11. ACF & PACF Sebelum Differencing Data ADRO



Gambar 4.12. ACF & PACF Setelah Differencing Pertama Data ADRO



Gambar 4.13. ACF & PACF Setelah Differencing Kedua Data ADRO

Dalam kasus data ADRO sama seperti data ICBP, proses *differencing* tersebut dilakukan dua kali. Jika dilihat dari Gambar 4.12 bahwa data sudah stasioner di proses *differencing* yang pertama, tetapi lag yang melewati area biru atau area batas signifikan tidak terlalu optimal, sehingga *differencing* dilakukan sekali lagi untuk melihat apakah pada *differencing* kedua terdapat lag yang lebih optimal. Dari proses *differencing* tersebut diperoleh parameter $d = 2$ yang akan digunakan dalam model ARIMA(p,d,q).

4.5. Normalisasi Data

Khusus untuk LSTM, data akan terlebih dahulu dilakukan normalisasi sebelum dilakukan proses *training* dan *testing*. Metode *MinMax* yang akan dimanfaatkan dalam proses normalisasi data dengan rentang 0 sampai 1. Data yang sudah dinormalisasi dapat dilihat pada Tabel 4.8 dengan menggunakan data transaksi pada TLKM.

Tabel 4.8. Atribut open sebelum dan sesudah normalisasi data TLKM

No	Open (Sebelum normalisasi)	Open (Setelah normalisasi)
1	3370.0	0.2677
2	3360.0	0.2598
3	3340.0	0.2440
4	3280.0	0.1968
5	3290.0	0.2047
6	3130.0	0.0787
7	3300.0	0.2125

8	3250.0	0.1732
...
248	4260.0	0.9685

Data yang dinormalisasikan tersebut diperoleh dari persamaan 14, berikut contoh perhitungannya.

$$X = 3370$$

$$X_{\min} = 3030$$

$$X_{\max} = 4300$$

$$X_n = \frac{x - X_{\min}}{X_{\max} - X_{\min}} = \frac{3370 - 3030}{4300 - 3030} = \frac{340}{1270} = 0,2677$$

4.6. Penentuan Plot Model

Khusus untuk ARIMA, setelah tahapan *differencing* dilakukan maka selanjutnya adalah menentukan parameter p dan q untuk dilakukan pemodelan. Identifikasi p dan q dilakukan dengan melihat plot PACF dan ACF yang dapat dilihat dalam tahapan uji stasioneritas dan *differencing* pada Gambar 4.3 – 4.13, lalu dari gambar – gambar tersebut dapat dibuat tabel p dan q dari plot PACF dan ACF yang dapat dilihat pada Tabel 4.9.

Tabel 4.9. Penentuan p dan q untuk pemodelan ARIMA

Parameter	TLKM		KLBF		ICBP		ANTM	ADRO
p	1	1	0	1	0	1	1	1
d	1	1	1	1	1	2	0	2
q	2	3	1	1	0	1	1	1

Data yang sudah stasioner tersebut terjadi sebuah kondisi yang dinamakan *cut off*. Kondisi *cut off* tersebut terjadi ketika lag pada plot ACF dan PACF melewati batas area berwarna biru yang dinamakan area batas signifikan. Jika terdapat lag

yang melewati area biru tersebut maka lag tersebut yang akan digunakan menjadi p atau q untuk pemodelan.

4.7. Pembagian Data

Dalam proses pembagian data ini yang terlebih dahulu melakukan pembagian data ialah model LSTM karena pada ARIMA dilakukan penentuan p dan q terlebih dahulu. Pembagian data menggunakan rasio 80:20, yaitu 80% yang menjadi data *test* dan 20% menjadi data *train*. Aturan yang menyatakan data *train* harus lebih besar sangat berlaku pada kedua model yang digunakan dalam penelitian ini karena model dapat mempelajari lebih baik apa yang akan dihasilkan nantinya. Berikut komposisi jumlah data *train* dan *test* pada proses ini.

Tabel 4.10. Data *train* dan data *test* dataset TLKM

Data Train	Data Test
80%	20%
195	50

4.8. Pemodelan ARIMA

Setelah semua tahapan yang diperlukan untuk pemodelan ARIMA telah dilakukan maka berdasarkan Tabel 4.9 kita akan menggunakan beberapa model ARIMA dengan ordo yang berbeda. Pada proses pemodelan ini akan dilihat terlebih dahulu *output* dari model ARIMA yang akan dicoba dan yang sesuai dengan dataset yang digunakan.

Pertama akan melihat *output* data TLKM dengan model ARIMA(1,1,2) dapat dilihat pada Gambar 4.14.

	coef	std err	z	P> z
const	2.1246	2.803	0.758	0.449
ar.L1.D.Open	-0.3945	0.370	-1.066	0.286
ma.L1.D.Open	0.1634	0.361	0.453	0.651
ma.L2.D.Open	-0.2246	0.101	-2.231	0.026

Gambar 4.14. Output Tabel Parameter ARIMA(1,1,2) Data TLKM

Dari Gambar 4.14 terdapat beberapa variabel seperti *coef*, *std err*, *z*, dan $P>|z|$. Berikut penjelasan untuk setiap variabel pada gambar tersebut, penjelasan variabel ini juga berlaku untuk gambar – gambar selanjutnya, yaitu Gambar 4.15 – 4.20.

Fajar Akbardipura, 2022

PERBANDINGAN PERFORMA PERAMA PERAMALAN HARGA SAHAM 5 PERUSAHAAN PADA INDEKS LQ45 MENGGUNAKAN AUTOREGRESSIVE INTEGRATED MOVING AVERAGE DAN LONG SHORT – TERM MEMORY

UPN Veteran Jakarta, Fakultas Ilmu Komputer, S1 Informatika

[www.upnvj.ac.id – www.library.upnvj.ac.id – www.repository.upnvj.ac.id]

1. Pada kolom “coef” merupakan signifikansi pada setiap fitur. Variabel “ar.L1.D.Open” mengacu pada istilah *autoregressive* dengan lag 1 dengan data “Open”. Variabel “ma.L1.D.Open” dan “ma.L1.D.Open” mengacu pada istilah *moving average* dengan lag 1 dan 2. Terdapat “D” pada variabel tersebut yang merupakan *difference*, menandakan bahwa pada parameter tersebut terjadi proses *differencing*. Semua koefisien ini adalah bagian dari persamaan ARIMA di bawah ini. Semakin tinggi jumlah lag yang digunakan dalam model, semakin panjang persamaannya. Variabel ar.L1, ma.L1, dan ma.L2 mendapatkan nilai “coef” masing – masing dari persamaan (5) dan persamaan (6) karena variabel tersebut bergantung kepada titik lag pada plot ACF dan PACF.
2. Kolom “std err” merupakan perkiraan kesalahan nilai prediksi. Ini memberi tahu seberapa kuat pengaruh kesalahan residual pada parameter estimator.
3. 'z' sama dengan nilai “coef” dibagi dengan “std err”. Dengan demikian koefisien standar.
4. kolom $P > |z|$ adalah nilai-p dari koefisien.

Dari Gambar 4.14 tersebut diperoleh persamaan di bawah ini. Untuk keterangan simbol yang digunakan dalam persamaan ARIMA dapat dilihat pada Sub Bab 2.3.4.

$$\begin{aligned} \phi_1(B)(1 - B)^1 Z_t &= \theta_0 + \theta_1(B) \alpha_t \\ (1 - \phi_1 B)(1 - B) Z_t &= \theta_0 + (1 - \theta_1 B - \theta_2 B^2) \alpha_t \\ (1 - B - \phi_1 B + \phi_1 B^2) Z_t &= \theta_0 + (1 - \theta_1 B - \theta_2 B^2) \alpha_t \\ (1 - (1 + \phi_1)B + \phi_1 B^2) Z_t &= \theta_0 + (1 - \theta_1 B - \theta_2 B^2) \alpha_t \\ Z_t &= (1 + \phi_1) Z_{t-1} - \phi_1 Z_{t-2} + \theta_0 + (1 - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2}) \alpha_t \\ Z_t &= (1 + \phi_1) Z_{t-1} - \phi_1 Z_{t-2} + \theta_0 + \alpha_t - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} \\ Z_t &= (1 - 0.3945) Z_{t-1} + 0.3945 Z_{t-2} + 2.1246 + \alpha_t - 0.1634 \alpha_{t-1} \\ &+ 0.2246 \alpha_{t-2} \end{aligned}$$

Kemudian untuk model ARIMA(1,1,3) diperoleh *output* seperti Gambar 4.15.

	coef	std err	z	P> z
const	2.1058	2.924	0.720	0.471
ar.L1.D.Open	-0.0320	0.803	-0.040	0.968
ma.L1.D.Open	-0.2029	0.800	-0.254	0.800
ma.L2.D.Open	-0.1442	0.191	-0.757	0.449
ma.L3.D.Open	0.0729	0.117	0.626	0.532

Gambar 4.15. Output Tabel Parameter ARIMA(1,1,3) Data TLKM

Tabel *output* dari Gambar 4.15 dapat dibuat persamaan sebagai berikut.

$$\begin{aligned} \phi_1(B)(1 - B)^1 Z_t &= \theta_0 + \theta_1(B) \alpha_t \\ (1 - \phi_1 B)(1 - B) Z_t &= \theta_0 + (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3) \alpha_t \\ (1 - B - \phi_1 B + \phi_1 B^2) Z_t &= \theta_0 + (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3) \alpha_t \\ (1 - (1 + \phi_1)B + \phi_1 B^2) Z_t &= \theta_0 + (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3) \alpha_t \\ Z_t &= (1 + \phi_1) Z_{t-1} - \phi_1 Z_{t-2} + \theta_0 + (1 - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} - \theta_3 \alpha_{t-3}) \alpha_t \\ Z_t &= (1 + \phi_1) Z_{t-1} - \phi_1 Z_{t-2} + \theta_0 + \alpha_t - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} - \theta_3 \alpha_{t-3} \\ Z_t &= (1 - 0.0320) Z_{t-1} + 0.0320 Z_{t-2} + 2.1058 + \alpha_t + 0.2029 \alpha_{t-1} \\ &+ 0.1442 \alpha_{t-2} - 0.0729 \alpha_{t-3} \end{aligned}$$

Selanjutnya akan melihat *output* nilai estimasi data KLBF dengan paramateter model ARIMA(1,1,1) yang terdapat pada Gambar 4.16.

	coef	std err	z	P> z
const	0.5128	1.389	0.369	0.712
ar.L1.D.Open	0.2334	0.196	1.191	0.234
ma.L1.D.Open	-0.4755	0.172	-2.762	0.006

Gambar 4.16. Output Tabel Parameter ARIMA(1,1,1) Data KLBF

Dari *output* tersebut diperoleh persamaan ARIMA sebagai berikut.

$$\begin{aligned} \phi_1(B)(1 - B)^1 Z_t &= \theta_0 + \theta_1(B) \alpha_t \\ (1 - \phi_1 B)(1 - B) Z_t &= \theta_0 + (1 - \theta_1 B) \alpha_t \\ (1 - B - \phi_1 B + \phi_1 B^2) Z_t &= \theta_0 + (1 - \theta_1 B) \alpha_t \\ (1 - (1 + \phi_1)B + \phi_1 B^2) Z_t &= \theta_0 + (1 - \theta_1 B) \alpha_t \\ Z_t &= (1 + \phi_1) Z_{t-1} - \phi_1 Z_{t-2} + \theta_0 + (1 - \theta_1 B) \alpha_t \\ Z_t &= (1 + \phi_1) Z_{t-1} - \phi_1 Z_{t-2} + \theta_0 + \alpha_t - \theta_1 \alpha_{t-1} \end{aligned}$$

Fajar Akbardipura, 2022

PERBANDINGAN PERFORMA PERAMA PERAMALAN HARGA SAHAM 5 PERUSAHAAN PADA INDEKS LQ45 MENGGUNAKAN AUTOREGRESSIVE INTEGRATED MOVING AVERAGE DAN LONG SHORT - TERM MEMORY

UPN Veteran Jakarta, Fakultas Ilmu Komputer, S1 Informatika

[www.upnvj.ac.id - www.library.upnvj.ac.id - www.repository.upnvj.ac.id]

$$Z_t = (1 + 0.2334) Z_{t-1} - 0.2334 Z_{t-2} + 0.5128 + \alpha_t + 0.4755 \alpha_{t-1}$$

Kemudian untuk parameter model ARIMA(0,1,1) diperoleh *output* pada Gambar 4.17.

	coef	std err	z	P> z
const	0.4988	1.504	0.332	0.740
ma.L1.D.Open	-0.2608	0.077	-3.407	0.001

Gambar 4.17. Output Tabel Parameter ARIMA(0,1,1) Data KLBF

Tabel *output* dari Gambar 4.17 dapat dibuat persamaan sebagai berikut.

$$\phi_0(B)(1 - B)^1 Z_t = \theta_0 + \theta_1(B) \alpha_t$$

$$1(1 - B)Z_t = \theta_0 + (1 - \theta_1 B) \alpha_t$$

$$Z_t - Z_t B = \theta_0 + (1 - \theta_1 B) \alpha_t$$

$$Z_t = Z_{t-1} + \theta_0 + \alpha_t - \theta_1 \alpha_{t-1}$$

$$Z_t = Z_{t-1} + 0.4988 + \alpha_t + 0.2608 \alpha_{t-1}$$

Selanjutnya dengan data ICBP yang menghasilkan *output* dari parameter model ARIMA(1,2,1) dapat dilihat pada Gambar 4.18.

	coef	std err	z	P> z
const	0.1916	0.163	1.176	0.240
ar.L1.D2.Open	-0.0165	0.072	-0.228	0.820
ma.L1.D2.Open	-1.0000	0.014	-73.139	0.000

Gambar 4.18. Output Tabel Parameter ARIMA(1,2,1) Data ICBP

Diperoleh persamaan sebagai berikut.

$$\phi_1(B)(1 - B)^2 Z_t = \theta_0 + \theta_1(B) \alpha_t$$

$$(1 - \phi_1 B)(1 - B)^2 Z_t = \theta_0 + (1 - \theta_1 B) \alpha_t$$

$$(1 - \phi_1 B)(1 - 2B + B^2) Z_t = \theta_0 + (1 - \theta_1 B) \alpha_t$$

$$(1 - \phi_1 B)(1 - 2B + B^2) Z_t = \theta_0 + \alpha_t - \theta_1 \alpha_{t-1}$$

$$(1 - 2B + B^2 - \phi_1 B + 2\phi_1 B^2 - \phi_1 B^3) Z_t = \theta_0 + \alpha_t - \theta_1 \alpha_{t-1}$$

$$Z_t - 2B + B^2 - \phi_1 B + 2\phi_1 B^2 - \phi_1 B^3 = \theta_0 + \alpha_t - \theta_1 \alpha_{t-1}$$

$$Z_t - (2 + \phi_1)B + (1 + 2\phi_1)B^2 - \phi_1 B^3 = \theta_0 + \alpha_t - \theta_1 \alpha_{t-1}$$

Fajar Akbardipura, 2022

PERBANDINGAN PERFORMA PERAMA PERAMALAN HARGA SAHAM 5 PERUSAHAAN PADA INDEKS LQ45 MENGGUNAKAN AUTOREGRESSIVE INTEGRATED MOVING AVERAGE DAN LONG SHORT - TERM MEMORY

UPN Veteran Jakarta, Fakultas Ilmu Komputer, S1 Informatika

[www.upnvj.ac.id - www.library.upnvj.ac.id - www.repository.upnvj.ac.id]

$$Z_t = (2 + \phi_1)Z_{t-1} - (1 + 2\phi_1)Z_{t-2} + \phi_1 Z_{t-3} + \theta_0 + \alpha_t - \theta_1 \alpha_{t-1}$$

$$Z_t = (2 - 0.0165)Z_{t-1} - (1 + 2(-0.0165))Z_{t-2} - 0.0165 Z_{t-3} + 0.1916 + \alpha_t + 1 \alpha_{t-1}$$

Untuk data selanjutnya, yaitu data ANTM dengan parameter model ARMA(1,1). Model untuk kasus data ANTM ini berbeda dengan model data yang lainnya, hal ini disebabkan karena data tersebut dinyatakan stasioner, sehingga tidak diperlukan adanya *differencing*. Menjadikan model yang digunakan adalah model ARMA(p,q) dibandingkan ARIMA(p,d,q). Berikut *output* parameter model yang terdapat pada Gambar 4.19.

	coef	std err	z	P> z
const	2432.0099	52.312	46.490	0.000
ar.L1.Open	0.9094	0.033	27.907	0.000
ma.L1.Open	-0.1566	0.082	-1.903	0.057

Gambar 4.19. Output Tabel Parameter ARMA(1,1) Data ANTM

Diperoleh persamaan ARMA seperti di bawah ini. Untuk keterangan simbol yang digunakan dalam persamaan ARMA dapat dilihat pada Sub Bab 2.3.3.

$$\phi_p(B)Z_t = \theta_q(B)\alpha_t$$

$$\phi_1(B)Z_t = \theta_1(B)\alpha_t$$

$$(1 - \phi_1 B) Z_t = \phi_0 + (1 - \theta_1 B) \alpha_t$$

$$Z_t - \phi_1 Z_t B = \phi_0 + (1 - \theta_1 B) \alpha_t$$

$$Z_t = \phi_0 + \phi_1 Z_{t-1} + \alpha_t - \theta_1 \alpha_{t-1}$$

$$Z_t = 2432.0099 + 0.9094 Z_{t-1} + \alpha_t + 0.1566 \alpha_{t-1}$$

Data terakhir yang akan dilihat *output* nya adalah ADRO yang memiliki parameter model ARIMA(1,2,1) yang dapat dilihat pada Gambar 4.20.

	coef	std err	z	P> z
const	0.0386	0.051	0.757	0.449
ar.L1.D2.Open	0.0359	0.072	0.499	0.618
ma.L1.D2.Open	-1.0000	0.014	-71.018	0.000

Gambar 4.20. Output Tabel Parameter ARIMA(1,2,1) Data ADRO

Berdasarkan *output* Gambar 4.20 diperoleh persamaan berikut.

Fajar Akbardipura, 2022

PERBANDINGAN PERFORMA PERAMA PERAMALAN HARGA SAHAM 5 PERUSAHAAN PADA INDEKS LQ45 MENGGUNAKAN AUTOREGRESSIVE INTEGRATED MOVING AVERAGE DAN LONG SHORT – TERM MEMORY

UPN Veteran Jakarta, Fakultas Ilmu Komputer, S1 Informatika

[www.upnvj.ac.id – www.library.upnvj.ac.id – www.repository.upnvj.ac.id]

$$\begin{aligned}
\phi_1(B)(1 - B)^2 Z_t &= \theta_0 + \theta_1(B) \alpha_t \\
(1 - \phi_1 B)(1 - B)^2 Z_t &= \theta_0 + (1 - \theta_1 B) \alpha_t \\
(1 - \phi_1 B)(1 - 2B + B^2) Z_t &= \theta_0 + (1 - \theta_1 B) \alpha_t \\
(1 - \phi_1 B)(1 - 2B + B^2) Z_t &= \theta_0 + \alpha_t - \theta_1 \alpha_{t-1} \\
(1 - 2B + B^2 - \phi_1 B + 2\phi_1 B^2 - \phi_1 B^3) Z_t &= \theta_0 + \alpha_t - \theta_1 \alpha_{t-1} \\
Z_t - 2B + B^2 - \phi_1 B + 2\phi_1 B^2 - \phi_1 B^3 &= \theta_0 + \alpha_t - \theta_1 \alpha_{t-1} \\
Z_t - (2 + \phi_1)B + (1 + 2\phi_1)B^2 - \phi_1 B^3 &= \theta_0 + \alpha_t - \theta_1 \alpha_{t-1} \\
Z_t = (2 + \phi_1)Z_{t-1} - (1 + 2\phi_1)Z_{t-2} + \phi_1 Z_{t-3} + \theta_0 + \alpha_t - \theta_1 \alpha_{t-1} \\
Z_t = (2 + 0.0359)Z_{t-1} - (1 + 2(0.0359)Z_{t-2} + 0.0359 Z_{t-3} + 0.0386 + \\
\alpha_t - 1 \alpha_{t-1}
\end{aligned}$$

Dari pemodelan yang telah dilakukan, terdapat beberapa data yang memiliki 2 parameter model. Model tersebut akan dibandingkan dan dilihat yang memberikan hasil yang paling baik dari prediksi harga saham bulan Januari 2021 sampai dengan bulan Januari 2022.

4.9. Training LSTM

Dalam model LSTM yang digunakan ini akan di *training* tidak hanya sekali saja karena akan dilihat model yang memiliki hasil optimal dalam prediksi harga saham tersebut. Terdapat beberapa parameter yang berada di LSTM yang akan digunakan seperti jumlah_units, fungsi untuk optimasi, *loss function*, *batch size*, dan jumlah dari *epoch*.

```

# Parameter Initialization
jumlah_units = 50
optimasi     = 'adam'
loss_func    = 'mean_squared_error'
ukuran_batch = 32
jumlah_epoch = 200

```

Dari parameter yang sudah ditentukan maka model tersebut siap untuk dilakukan *training* model. Kemudian LSTM akan bekerja dengan cara berikut.

1. Proses LSTM diawali dengan menentukan informasi yang akan dihilangkan dari *cell state* dan yang tidak, melalui *forget gate* dengan menggunakan persamaan (10).
2. Selanjutnya mekanisme *input gate* akan menentukan nilai *input* dengan menghitung persamaan (12) dan (13).
3. Kemudian *cell state* akan di *update* nilai nya dengan menggunakan persamaan (14).
4. Dilanjut dengan tahapan *output gate* untuk menentukan nilai *output* menggunakan persamaan (15) dan (16).
5. Proses yang telah dilakukan dari 1-4 akan terus berulang sebanyak jumlah dari epoch yang ditentukan. Proses akan berhenti jika sudah memenuhi jumlah epoch. Terdapat kemungkinan jika proses masih belum berhenti maka akan dioptimasi dengan optimizer yang ditentukan pada parameter awal yaitu adam untuk hitung setiap nilai loss.
6. Segala proses yang telah dilalui dalam LSTM akan menghasilkan *output* berupa nilai dari prediksi harga saham yang dilakukan dalam penelitian ini. Kemudian dapat dihitung nilai MAPE dari model LSTM tersebut.

4.10. Model LSTM

Tahapan berikutnya yaitu pengujian model menggunakan data *test*. Pengujian model dilakukan dengan proses uji parameter jumlah epoch terlebih dahulu kemudian akan diuji juga dengan *batch size*. Parameter yang memiliki nilai MAPE yang baik akan dipilih pada tiap model.

4.10.1. Uji dengan jumlah epoch

Proses pengujian pertama untuk LSTM akan dilakukan dengan optimasi berdasarkan jumlah epoch. Parameter epoch yang akan diuji untuk mencari hasil terbaik termasuk 50, 100, 150, 200. Berikut hasil dari uji dengan jumlah epoch.

Tabel 4.11. Uji parameter epoch data TLKM

Epoch	Batch Size	MAPE	Avg Loss	Avg Time (Detik)
50	32	0.02597 (+- 2,5 %)	0.007017	18.755
100	32	0.02744 (+- 2,7 %)	0.006109	36.076
150	32	0.02383 (+- 2,3 %)	0.005129	55.613
200	32	0.01795 (+- 1,7 %)	0.004207	70.856

Berdasarkan Tabel 4.10. model yang terbaik berdasarkan jumlah epoch yaitu model dengan jumlah epoch = 200 dengan nilai MAPE sebesar 0,017957 atau sekitar 1,7% yang memiliki *average time* untuk komputasi selama 70,856 detik.

Tabel 4.12. Uji parameter epoch data KLBF

Epoch	Batch Size	MAPE	Avg Loss	Avg Time (Detik)
50	32	0.01422 (+- 1,4 %)	0.013520	9.090
100	32	0.01666 (+- 1,6 %)	0.009693	15.650
150	32	0.01352 (+- 1,3 %)	0.008078	22.069
200	32	0.01216 (+- 1,2 %)	0.007348	28.221

Berdasarkan Tabel 4.12. model yang terbaik berdasarkan jumlah epoch yaitu model dengan jumlah epoch = 200 dengan nilai MAPE sebesar 0.01216 atau sekitar 1,2% yang memiliki *average time* untuk komputasi selama 28,221 detik.

Tabel 4.13. Uji parameter epoch data ICBP

Epoch	Batch Size	MAPE	Avg Loss	Avg Time (Detik)
50	32	0.00812 (+- 0,8 %)	0.016455	10.589
100	32	0.00865 (+- 0,8 %)	0.012432	16.927
150	32	0.00912 (+- 0,9 %)	0.010950	23.099
200	32	0.01082 (+- 1 %)	0.010485	29.891

Berdasarkan Tabel 4.13. model yang terbaik berdasarkan jumlah epoch yaitu model dengan jumlah epoch = 50 dengan nilai MAPE sebesar 0.00812 atau sekitar 0,8 % yang memiliki *average time* untuk komputasi selama 10,589 detik.

Tabel 4.14. Uji parameter epoch data ANTM

Epoch	Batch Size	MAPE	Avg Loss	Avg Time (Detik)
50	32	0.05034 (+- 5 %)	0.016373	10.230
100	32	0.04076 (+- 4 %)	0.009841	17.828
150	32	0.04426 (+- 4,4 %)	0.009442	25.519
200	32	0.04376 (+- 4,3 %)	0.007429	31.286

Berdasarkan Tabel 4.14. model yang terbaik berdasarkan jumlah epoch yaitu model dengan jumlah epoch = 100 dengan nilai MAPE sebesar 0.04076 atau sekitar 4 % yang memiliki *average time* untuk komputasi selama 17,828 detik.

Tabel 4.15. Uji parameter epoch data ADRO

Epoch	Batch Size	MAPE	Avg Loss	Avg Time (Detik)
50	32	0.05268 (+- 5,2 %)	0.005398	10.127
100	32	0.06516 (+- 6,5 %)	0.004239	16.554
150	32	0.04575 (+- 4,5 %)	0.003514	23.777
200	32	0.04926 (+- 4,9 %)	0.003081	30.993

Berdasarkan Tabel 4.15. model yang terbaik berdasarkan jumlah epoch yaitu model dengan jumlah epoch = 150 dengan nilai MAPE sebesar 0.04575 atau sekitar 4,5 % yang memiliki *average time* untuk komputasi selama 17,828 detik.

4.10.2. Uji dengan *batch size*

Proses pengujian yang selanjutnya untuk LSTM dilakukan dengan optimasi berdasarkan *batch size*. Parameter *batch size* yang akan diuji untuk mencari hasil terbaik termasuk 16, 32, 48, 64. Parameter lainnya, yaitu epoch akan menggunakan hasil epoch terbaik yang sebelumnya sudah diuji yaitu 200. Berikut hasil dari uji dengan *batch size*.

Tabel 4.16. Uji parameter *batch size* data TLKM

Epoch	Batch Size	MAPE	Avg Loss	Avg Time (Detik)
200	16	0.01782 (+- 1,7 %)	0.003766	121.406
200	32	0.01813 (+- 1,8 %)	0.004215	75.645
200	48	0.02490 (+- 2,4 %)	0.005577	45.734
200	64	0.02042 (+- 2 %)	0.005004	51.711

Berdasarkan Tabel 4.16. model yang terbaik berdasarkan *batch size* yaitu 16 dengan nilai MAPE sebesar 0.01782 atau sekitar 1,7% yang memiliki *average time* untuk komputasi selama 121.406 detik.

Tabel 4.17. Uji parameter *batch size* data KLBF

Epoch	Batch Size	MAPE	Avg Loss	Avg Time (Detik)
200	16	0.00996 (+- 0,9 %)	0.006214	43.628
200	32	0.01216 (+- 1,2 %)	0.007348	28.221
200	48	0.01169 (+- 1,1 %)	0.008796	21.871
200	64	0.01408 (+- 1,4 %)	0.008877	23.924

Berdasarkan Tabel 4.17. model yang terbaik berdasarkan *batch size* yaitu 16 dengan nilai MAPE sebesar 0.00996 atau sekitar 0,9% yang memiliki *average time* untuk komputasi selama 43,628 detik.

Tabel 4.18. Uji parameter *batch size* data ICBP

Epoch	Batch Size	MAPE	Avg Loss	Avg Time (Detik)
50	16	0.00986 (+- 0,9 %)	0.012233	13.779
50	32	0.00812 (+- 0,8 %)	0.016455	10.589
50	48	0.00968 (+- 0,9 %)	0.019207	7.461
50	64	0.00965 (+- 0,9 %)	0.022308	7.767

Berdasarkan Tabel 4.18. model yang terbaik berdasarkan *batch size* yaitu 32 dengan nilai MAPE sebesar 0.00812 atau sekitar 0,8% yang memiliki *average time* untuk komputasi selama 10,589 detik.

Tabel 4.19. Uji parameter *batch size* data ANTM

Epoch	Batch Size	MAPE	Avg Loss	Avg Time (Detik)
100	16	0.03871 (+- 3,8 %)	0.007818	24.771
100	32	0.02744 (+- 2,7 %)	0.006109	36.076
100	48	0.05241 (+- 5,2 %)	0.012892	13.529
100	64	0.04491 (+- 4,4 %)	0.013742	15.114

Berdasarkan Tabel 4.19. model yang terbaik berdasarkan *batch size* yaitu 32 dengan nilai MAPE sebesar 0.02744 atau sekitar 2,7% yang memiliki *average time* untuk komputasi selama 36,076 detik.

Tabel 4.20. Uji parameter *batch size* data ADRO

Epoch	Batch Size	MAPE	Avg Loss	Avg Time (Detik)
150	16	0.03212 (+- 3,2 %)	0.003013	36.035
150	32	0.04575 (+- 4,5 %)	0.003514	23.777
150	48	0.05290 (+- 5,2 %)	0.003951	18.864
150	64	0.04564 (+- 4,5 %)	0.004123	18.713

Berdasarkan Tabel 4.20. model yang terbaik berdasarkan *batch size* yaitu 16 dengan nilai MAPE sebesar 0.03212 atau sekitar 3,2% yang memiliki *average time* untuk komputasi selama 36,035 detik.

4.11. Denormalisasi

Semua data yang diproses oleh model LSTM sebelumnya merupakan data yang nilai sudah dinormalisasikan, oleh karena itu pada tahapan ini akan dilakukan denormalisasi yang dimana proses ini akan mengembalikan sebuah nilai yang dinormalisasi menjadi nilai dengan bentuk awal sebelum dinormalisasi, yaitu menjadi harga saham itu sendiri. Berikut perhitungan untuk denormalisasi.

$$X_n = 0.64796364$$

$$X_{min} = 3030$$

$$X_{max} = 4300$$

$$x = X_n(X_{max} - X_{min}) + X_{min}$$

$$x = 0.64796364 (4300 - 3030) + 3030$$

$$x = 3852.91382$$

4.12. Evaluasi Hasil Perbandingan Algoritma

Semua proses yang sudah dilalui oleh ARIMA dan LSTM sampai di tahapan terakhir yaitu evaluasi hasil prediksi yang dilakukan masing masing algoritma menggunakan dataset TLKM. Teknik mengukur kesalahan pada penelitian ini

menggunakan MAPE. Nilai MAPE inilah yang akan membuktikan model mana yang terbaik dalam kasus prediksi harga saham pembuka ini. Jika semakin kecil nilai MAPE yang dihasilkan maka semakin baik model tersebut. Berikut perhitungan MAPE dan tabel perbandingan algoritma dapat dilihat pada Tabel 4.11.

$$\text{MAPE} = \frac{\sum_{t=1}^n \left| \left(\frac{A_t - F_t}{A_t} \right) 100 \right|}{n}$$

$$\text{MAPE} = \frac{\left| \left(\left(\frac{8950 - 8893.8144}{8950} \right) + \left(\frac{8825 - 8890.7216}{8825} \right) + \dots + \left(\frac{8675 - 8745.3608}{8675} \right) \right) \times 100 \right|}{50}$$

$$\text{MAPE} = \frac{0,4536}{50} = 0.009072$$

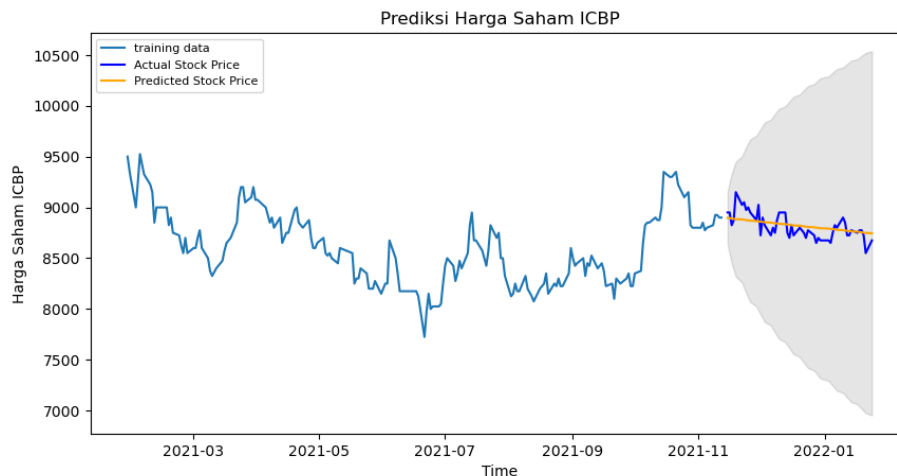
Tabel 4.21. Perbandingan hasil MAPE prediksi harga saham

Algoritma	Data	MAPE
ARIMA(1,1,2)	TLKM	0.08269 (+- 8,2 %)
ARIMA(1,1,3)	TLKM	0.083724 (+- 8,3 %)
ARIMA(0,1,1)	KLBF	0.014408 (+- 1,4 %)
ARIMA(1,1,1)	KLBF	0.014284 (+- 1,4 %)
ARIMA(0,1,0)	ICBP	0.009072 (+- 0,9 %)
ARIMA(1,2,1)	ICBP	0.066159 (+- 6,6 %)
ARMA(1,1)	ANTM	0.086309 (+- 8,6 %)
ARIMA(1,2,1)	ADRO	0.090117 (+- 9%)
LSTM (batch_size = 32, epoch = 200)	TLKM	0.01795 (+- 1,7 %)
LSTM (batch_size = 16, epoch = 200)	TLKM	0.01782 (+- 1,7 %)
LSTM (batch_size = 32, epoch = 200)	KLBF	0.01216 (+- 1,2 %)
LSTM (batch_size = 16, epoch = 200)	KLBF	0.00996 (+- 0,9 %)
LSTM (batch_size = 32, epoch = 50)	ICBP	0.00812 (+- 0,8 %)
LSTM (batch_size = 32, epoch = 100)	ANTM	0.04076 (+- 4 %)

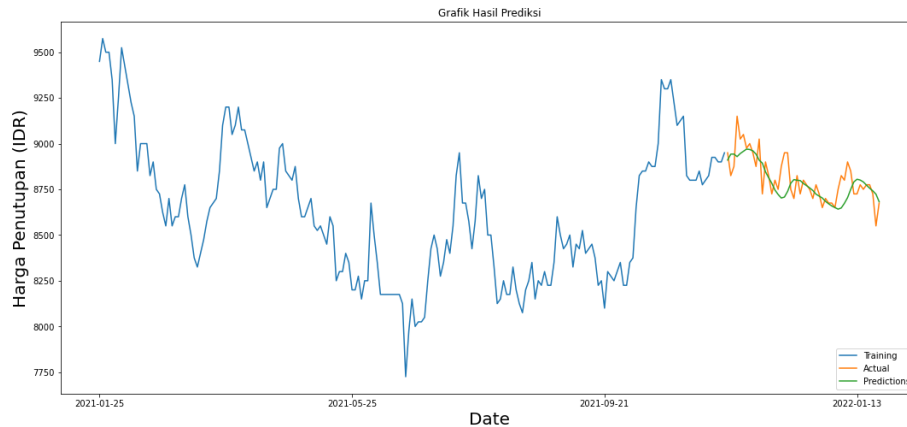
LSTM (batch_size = 16, epoch = 100)	ANTM	0.03871 (+- 3,8 %)
LSTM (batch_size = 32, epoch = 150)	ADRO	0.04575 (+- 4,5 %)
LSTM (batch_size = 16, epoch = 150)	ADRO	0.03212 (+- 3,2 %)

Dari Tabel 4.11 dapat dilihat hasil MAPE yang berbeda-beda pada setiap parameter model algoritma dan juga data yang digunakan. Jika diperhatikan bahwa pada model ARIMA lebih banyak menghasilkan MAPE yang bernilai lebih besar dibandingkan dengan LSTM. Hasil dari percobaan yang telah dilakukan menunjukkan bahwa nilai MAPE terkecil yang dihasilkan, yaitu 0.00812 atau sekitar 0,8 % yang menggunakan model LSTM.

Dalam hal ini pergerakan harga saham pada periode-periode sebelumnya dan juga pemilihan parameter model yang tepat sangat berpengaruh dalam memprediksi kedepannya menggunakan model ARIMA maupun LSTM. Hal ini ditunjukkan dengan nilai MAPE terkecil pada masing-masing algoritma adalah ketika memprediksi harga saham ICBP. Untuk lebih jelasnya dapat dilihat grafik prediksi harga saham ICBP pada Gambar 4.21 dan 4.22.



Gambar 4.21. Grafik Prediksi Harga Saham Model ARIMA Data ICBP



Gambar 4.22. Grafik Prediksi Harga Saham Model LSTM Data ICBP

Pergerakan harga saham ICBP tidak terlalu ekstrem untuk naik turunnya harga sehingga berpengaruh terhadap hasil prediksi oleh kedua algoritma. Oleh karena itu data ICBP mendapat hasil MAPE terkecil dengan parameter model yang tepat.