

**PERAMALAN SAHAM PT. TELKOM INDONESIA (PERSERO) TBK [TLKM]  
MENGUNAKAN ANALISIS RUNTUN WAKTU ARIMA (p,d,q)**



**Disusun oleh:**

Rifqi Hafizuddin	(2106638204)
Aurelio Naufal Effendy	(2106638526)
Kamal Muftie Yafi	(2106725034)

**PROGRAM STUDI S1 STATISTIKA  
DEPARTEMEN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS INDONESIA  
DEPOK  
2023**

Nama	NPM	Kontribusi	Tingkat Kontribusi
Rifqi Hafizuddin	2106638204	Aktif dalam berdiskusi, Mengerjakan bab analisis data. Ikut membantu menulis <i>code</i> dalam analisis data, mengecek kembali tugas, menulis dan melakukan pengecekan ulang terhadap daftar pustaka.	100%
Aurelio Naufal Effendy	2106638526	Aktif dalam berdiskusi, mengerjakan bab pendahuluan dan penutup, membantu mengerjakan bab tinjauan pustaka, ikut membantu menulis <i>code</i> dalam analisis data, mengecek kembali tugas, menulis dan melakukan pengecekan ulang terhadap daftar pustaka.	100%
Kamal Muftie Yafi	2106725034	Aktif dalam berdiskusi, mengerjakan subbab diagnostik model dan <i>forecasting</i> , menulis <i>code</i> dalam analisis data, mengecek kembali tugas, menulis dan melakukan pengecekan ulang terhadap daftar pustaka.	100%

## KATA PENGANTAR

Puji syukur kami panjatkan ke hadirat Tuhan Yang Maha Esa karena dengan rahmat dan hidayah-Nya, Kami dapat menyelesaikan makalah untuk mata kuliah Metode Peramalan (*Time Series Analysis*) dengan judul **“Peramalan PT. Telkom Indonesia (Persero) Tbk [TLKM] dengan Menggunakan Analisis Runtun Waktu ARIMA (p,d,q)”** dengan baik. Terima kasih Kami ucapkan kepada Ibu Fevi Novkaniza selaku dosen mata kuliah Metode Peramalan Departemen Matematika yang telah memberikan ilmu mengenai mata kuliah Metode Peramalan dengan baik. Kami berharap makalah ini dapat berguna dalam rangka menambah wawasan serta pengetahuan Kami mengenai penerapan model ARIMA dalam menyelesaikan masalah di kehidupan sehari-hari dan menerapkannya dalam melakukan peramalan di masa yang akan datang. Kami juga menyadari sepenuhnya bahwa di dalam makalah ini masih terdapat kekurangan dan jauh dari kata sempurna. Oleh sebab itu, Kami berharap adanya kritik dan saran demi perbaikan makalah yang telah Kami buat di masa yang akan datang, mengingat tidak ada sesuatu yang sempurna tanpa saran yang membangun.

Semoga makalah ini dapat berguna bagi tim penulis maupun orang yang membacanya di kemudian hari.

Depok, 18 Juni 2023

Tim Penulis

## DAFTAR ISI

<b>KATA PENGANTAR.....</b>	<b>i</b>
<b>DAFTAR ISI.....</b>	<b>ii</b>
<b>BAB 1 PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	1
1.3 Tujuan Penelitian.....	2
1.4 Manfaat Penelitian.....	2
<b>BAB 2 TINJAUAN PUSTAKA .....</b>	<b>3</b>
2.1 Saham .....	3
2.2 Runtun Waktu .....	3
2.3 Model ARIMA ( <i>Autoregressive Integrated Moving Average</i> ) .....	4
2.3.1 Proses <i>Autoregressive</i> .....	4
2.3.2 Proses <i>Moving Average</i> .....	4
2.3.3 Random Walk .....	4
2.3.4 Stasioneritas .....	4
2.3.5 Model Terintegrasi.....	5
2.4 Proses Analisis Runtun Waktu .....	5
2.4.1 Identifikasi Data.....	5
2.4.2 Spesifikasi Model.....	6
2.4.3 Estimasi Parameter Model .....	8
2.4.4 Diagnosis Model .....	9
2.4.5 Peramalan ( <i>Forecasting</i> ).....	10
<b>BAB 3 ANALISIS DATA .....</b>	<b>12</b>
3.1 Data .....	12
3.2 Tahapan Penelitian .....	13
3.3 Uji Stasioner.....	13
3.4 Spesifikasi Model.....	14
3.4.1 ACF dan PACF .....	15
3.4.2 EACF .....	15
3.4.3 Pemilihan Model Terbaik .....	15
3.5 Estimasi Parameter .....	17
3.6 Model Diagnostik.....	17
3.6.1 Analisis Residual .....	17

3.6.2	<i>Overfitting</i> .....	20
3.7	Peramalan ( <i>Forecasting</i> ) .....	21
3.7.1.	<i>Ex Post</i> .....	21
3.7.2.	<i>Ex Ante</i> .....	22
<b>BAB 4</b>	<b>PENUTUP</b> .....	<b>24</b>
4.2.	Kesimpulan.....	24
4.3.	Saran.....	24
<b>DAFTAR PUSTAKA</b>	.....	<b>25</b>
<b>LAMPIRAN</b>	.....	<b>26</b>

# **BAB 1**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Peramalan merupakan salah satu unsur penting dalam pengambilan keputusan dan sangat di perlukan dalam kehidupan sehari-hari untuk memprediksi nilai suatu data pada peristiwa yang akan terjadi. Salah satu contoh penggunaannya adalah untuk memprediksi nilai harga penutupan saham.

Saham sendiri adalah kertas di mana di dalamnya tertulis dengan jelas jumlah, nama perusahaan serta hak kewajiban yang telah dijelaskan kepada setiap investornya. Saham dapat memberikan tingkat keuntungan yang menarik. Pergerakan harga saham termasuk dalam runtun waktu karena harganya berubah seiring berjalannya waktu. Pergerakan naik dan turunnya harga saham menjadi perhatian para investor. Jika harga saham yang dimiliki naik, maka investor akan membeli saham. Dalam melakukan investasi, harga saham adalah faktor yang harus diperhatikan karena harga saham menunjukkan prestasi emiten. Kinerja emiten bergerak searah dengan harga saham, semakin baik prestasi dari emiten maka keuntungan yang diperoleh dari operasi usaha semakin besar. Harga saham juga menunjukkan nilai suatu perusahaan dan merupakan indeks yang tepat untuk efektivitas perusahaan. Semakin tinggi harga saham, maka semakin tinggi pula nilai perusahaan tersebut dan sebaliknya.

Perusahaan Perseroan PT Telekomunikasi Indonesia Tbk atau biasa disingkat Telkom atau Telkom Indonesia adalah sebuah badan usaha milik negara Indonesia (BUMN) yang bergerak di bidang jasa layanan teknologi informasi dan komunikasi (TIK) dan jaringan telekomunikasi di Indonesia yang sudah terdaftar di Bursa Efek Indonesia (BEI) dengan kode “TLKM”. Perusahaan ini merupakan salah satu perusahaan operator seluler terbesar di Indonesia yang memiliki jangkauan jaringan paling luas hingga ke daerah terpencil. Pemegang saham mayoritas Telkom adalah Pemerintah Republik Indonesia sebesar 52.09%, sedangkan 47.91% sisanya dikuasai oleh publik.

Saham TLKM ini merupakan salah satu saham *blue-chip* atau saham unggulan di BEI, yang sering menjadi pilihan bagi para investor institusional dan ritel karena stabilitasnya dalam jangka panjang, relatif aman, dan likuiditasnya yang tinggi. Saham TLKM juga memiliki rekam jejak yang positif dalam hal memberikan dividen yang menarik bagi para pemegang sahamnya.

Akan tetapi, harga penutupan saham (*Closing Price*) saham TLKM tidak dapat di prediksi kapan terjadi kenaikan atau penurunan. Oleh sebab itu, diperlukan peramalan harga penutupan saham untuk mengatasi masalah tersebut.

### **1.2 Rumusan Masalah**

Berdasarkan uraian yang terdapat pada latar belakang, rumusan masalah yang akan digunakan pada penelitian ini adalah sebagai berikut:

1. Bagaimana gambaran umum data harga penutupan saham TLKM?
2. Bagaimana proses analisis runtun waktu sampai didapatkannya model terbaik yang dapat meramalkan dari data yang ada?
3. Bagaimana model runtun waktu terbaik yang dapat digunakan untuk melakukan peramalan harga penutupan saham TLKM?
4. Bagaimana prediksi harga penutupan saham TLKM dalam beberapa bulan ke depan (4 bulan setelah Juni 2023) berdasarkan prediksi menggunakan model ARIMA runtun waktu?

### **1.3 Tujuan Penelitian**

Adapun tujuan yang ingin dicapai yaitu:

1. Mengetahui gambaran umum data harga penutupan saham TLKM
2. Mengetahui proses dalam menganalisis data runtun waktu sampai mendapatkan model yang sesuai pada data tersebut.
3. Mendapatkan model runtun waktu yang terbaik dalam efisiensi parameter maupun kemampuan prediktif dalam memprediksi harga penutupan saham TLKM per bulan berdasarkan data dari bulan Oktober 2004 sampai Juni 2023 (tanggal 16).
4. Mendapatkan hasil prediksi harga penutupan saham TLKM dalam beberapa bulan ke depan (4 bulan setelah Juni 2023).

### **1.4 Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat yakni menjadi salah satu contoh penerapan peramalan dalam ilmu statistik, menjadi informasi dan referensi dalam meramalkan suatu data, bahan acuan untuk penelitian selanjutnya, dan yang utama adalah sebagai bahan pertimbangan bagi investor dalam memperjual belikan saham terutama saham TLKM.

## **BAB 2**

### **TINJAUAN PUSTAKA**

#### **2.1 Saham**

Saham (*stock*) merupakan salah satu instrumen pasar keuangan yang paling populer. Menerbitkan saham merupakan salah satu pilihan perusahaan ketika memutuskan untuk pendanaan perusahaan. Pada sisi yang lain, saham merupakan instrumen investasi yang banyak dipilih para investor karena saham mampu memberikan tingkat keuntungan yang menarik.

Saham dapat didefinisikan sebagai tanda penyertaan modal seseorang atau pihak (badan usaha) dalam suatu perusahaan atau perseroan terbatas. Dengan menyertakan modal tersebut, maka pihak tersebut memiliki klaim atas pendapatan perusahaan, klaim atas aset perusahaan, dan berhak hadir dalam Rapat Umum Pemegang Saham (RUPS).

Pada dasarnya terdapat dua keuntungan yang diperoleh investor dengan memiliki saham, yakni mendapatkan dividen yaitu pembagian keuntungan yang diberikan perusahaan dan berasal dari keuntungan yang dihasilkan perusahaan, serta memperoleh *capital gain* yaitu selisih antara harga beli dan harga jual saham. Akan tetapi, sebagai instrumen investasi, tentunya saham memiliki risiko. Salah satunya adalah *capital loss* yakni kebalikan dari *capital gain*, yaitu suatu kondisi di mana investor menjual saham lebih rendah dari harga beli. Selain itu terdapat risiko likuidasi, yakni perusahaan yang sahamnya dimiliki, dinyatakan bangkrut oleh pengadilan, atau perusahaan tersebut dibubarkan. Dalam hal ini hak klaim dari pemegang saham mendapat prioritas terakhir setelah seluruh kewajiban perusahaan dapat dilunasi. Jika masih terdapat sisa dari hasil penjualan kekayaan perusahaan tersebut, maka sisa tersebut dibagi secara proporsional kepada seluruh pemegang saham. Namun jika tidak terdapat sisa kekayaan perusahaan, maka pemegang saham tidak akan memperoleh hasil dari likuidasi tersebut. Kondisi ini merupakan risiko yang terberat dari pemegang saham. Untuk itu seorang pemegang saham dituntut untuk secara terus menerus mengikuti perkembangan perusahaan.

#### **2.2 Runtun Waktu**

Runtun waktu merupakan serangkaian data yang diobservasi berdasarkan waktu kemunculannya, satuan waktu yang digunakan bisa beragam mulai dari milidetik, detik, menit, jam, hari, bulan, tahun, dan masih banyak satuan lainnya. Data runtun waktu yang didapatkan akan digunakan untuk dianalisis agar dapat suatu model yang bisa berguna dalam memprediksi data runtun waktu tersebut dalam beberapa waktu ke depan dari waktu terakhir yang dimiliki.



## 2.3 Model ARIMA (*Autoregressive Integrated Moving Average*)

### 2.3.1 Proses *Autoregressive*

*Autoregressive Process*  $AR(p)$  merupakan metode di mana dijalankannya penaksiran regresi terhadap data itu sendiri, proses ini memiliki order sampai  $p$  untuk suatu model  $Y_t$ .

Pada umumnya, model  $AR(p)$  di notasikan dengan model berikut:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \dots + \phi_p Y_{t-p} + a_t$$

dimodel yang telah dituliskan di atas,  $a_t$  diasumsikan independen terhadap  $Y_{t-1}, Y_{t-2}, \dots \ni a_t$  disebut sebagai White Noise. Selain itu, model  $AR(p)$  hanya akan stasioner jika dan hanya jika nilai  $|\phi_p| < 1$  dan  $\phi_1 + \phi_2 + \dots + \phi_p < 1$ .

### 2.3.2 Proses *Moving Average*

*Moving Average*  $MA(q)$  adalah salah satu model yang dibentuk dengan mengambil rata-rata sebanyak  $q$  data kebelakang yang pada umumnya dapat dibentuk modelnya sebagai berikut:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Terkhusus untuk  $MA(1)$ , Kovariansi antar lag-nya adalah 0 untuk lag yang lebih besar sama dengan 2.

### 2.3.3 Random Walk

Random Walk didefinisikan sebagai proses di mana nilai variabel saat ini terdiri dari nilai masa lalu ditambah dengan White Noise (berdistribusi normal dengan mean nol dan varians  $\sigma_e^2$ ). Secara umum, modelnya adalah sebagai berikut

$$Y_t = Y_{t-1} + e_t$$

di mana

$$Y_1 = e_1$$

$$Y_2 = e_1 + e_2$$

$$\vdots$$

$$Y_t = e_1 + e_2 + \dots + e_t$$

### 2.3.4 Stasioneritas

Stasioneritas adalah suatu keadaan di mana perilaku dari suatu proses (umumnya pada kasus stokastik, namun dapat berlaku juga pada kasus non stokastik) tidak bergantung terhadap berjalannya waktu. Stasioneritas sendiri terbagi menjadi 2 yaitu:

1. *Strictly Stationary*
2. *Weakly Stationary*

Kasus *strictly stationary* adalah suatu keadaan di mana pdf bersama  $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$  akan memiliki pdf yang sama dengan pdf bersama  $Y_{t_1-k}, Y_{t_2-k}, \dots, Y_{t_n-k}$ . Pada kasus ini akan lebih banyak dibahas mengenai kasus *weakly stationary*, yaitu keadaan di mana baik mean, variansi, dan kovariansi antar lag-nya terbebas dari waktu. Untuk notasi matematikanya adalah sebagai berikut:

1.  $E[Y_t] = C_1$ , dimana  $C_1$  adalah suatu konstanta.
2.  $Var[Y_t] = C_2$ , dimana  $C_2$  adalah suatu konstanta.
3.  $Cov(Y_t, Y_{t-k})$  akan merupakan suatu fungsi  $t$ .

### 2.3.5 Model Terintegrasi

Model terintegrasi adalah model ARMA yang didiferensiasi sebanyak  $d$  kali, dimana dinotasikan sebagai  $W = \nabla^d Y_t$  yang sifatnya stasioner dan bermodelkan  $ARMA(p, q) \ni Y_t$  bermodelkan  $ARIMA(p, d, q)$ . Contohnya adalah untuk suatu proses  $ARIMA(p, 1, q)$  akan didapati  $W_t = Y_t - Y_{t-1}$  dimana

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \dots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

$$Y_t - Y_{t-1} = \phi_1 (Y_{t-1} - Y_{t-2}) + \dots + \phi_p (Y_{t-p} - Y_{t-p-1}) + \dots + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

## 2.4 Proses Analisis Runtun Waktu

### 2.4.1 Identifikasi Data

Data runtun waktu yang dimiliki, tentunya akan diidentifikasi terlebih dahulu sebelum berlanjut dalam penentuan model, lantaran dalam penetapan model diperlukan *dataset* yang sudah bersifat stasioner. Maka dari itu, pada tahap ini akan dicari cara sedemikian sehingga data menjadi stasioner dan dapat diproses lebih lanjut untuk penetapan model.

#### 1. *Plot Data Runtun*

Langkah pertama adalah membuat *plot* dari *dataset* yang dimiliki, *plot* ini akan memberikan gambaran awal dari sifat yang dimiliki dari data. Terkadang dari *plot* dapat secara langsung ditarik kesimpulan mengenai stasioneritas suatu data. Selain itu terkadang dapat terlihat juga apakah data sedang mengalami *trend* ataupun bersifat *seasonal* berdasarkan *plot* yang diberikan.

#### 2. Uji Stasioner

Terkadang penetapan stasioneritas data bisa ditetapkan secara langsung dengan melihat *plot*, hanya saja terkadang cara tersebut tidak terstandarisasi dan

cenderung subjektif. Maka dari itu terdapat dua metode yang pada umumnya digunakan dalam menguji stasioneritas suatu data.

#### a. Ljung Box

Pada Uji Ljung Box akan diuji hipotesis  $H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0$  dengan suatu  $\alpha$  di mana  $H_0$  akan ditolak apabila statistik uji

$$Q = n(n+2) \sum_{k=1}^m \left( \frac{r_k}{n-k} \right)^2 > \chi_{\alpha, m}^2$$

Apabila  $H_0$  ditolak, maka disimpulkan data tidak stasioner.

#### b. Augmented Dickey-Fuller (ADF)

Pada Uji ADF akan diuji hipotesis  $H_0: \pi = 0$  melawan  $H_1: \pi < 0$  dengan suatu  $\alpha$  dimana  $H_0$  akan ditolak apabila statistik uji

$$|\tau| = \left| \frac{\hat{\theta} - 1}{se(\hat{\theta})} \right| = \left| \frac{\hat{\pi}}{se(\hat{\pi})} \right| < |DF_{tabel}|$$

Apabila  $H_0$  ditolak, maka disimpulkan data stasioner.

Dalam melakukan pengujian stasioneritas, pada beberapa *software* tertentu juga bisa ditarik kesimpulan mengenai  $H_0$  berdasarkan  $p - value$  yang didapatkan pada masing-masing uji. yaitu  $H_0$  ditolak untuk  $p - value < \alpha$ .

### 2.4.2 Spesifikasi Model

Pada tahap ini, akan ditentukan model yang dianggap cocok dengan *dataset* yang dimiliki dengan mengambil kesimpulan berdasarkan dari visualisasi yang disajikan oleh ACF, PACF, dan EACF.

#### 1. ACF

*Autocorrelation Function* adalah fungsi yang bekerja dengan melihat pola dari  $r_k$  dan membandingkannya dengan  $\rho_k$  dari model *ARMA*, di mana pada ACF  $r_k$  didefinisikan sebagai berikut:

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

ACF memiliki kemampuan untuk mengidentifikasi model  $MA(q)$ , namun kurang mampu dalam mengidentifikasi  $AR(p)$ . Grafik yang disajikan oleh ACF merupakan visualisasi dari uji hipotesis  $H_0: \rho_k = 0$  dimana apabila ada tiang pancang yang melewati garis putus putus menandakan adanya  $\rho_k$  yang signifikan.

Berikut merupakan ciri-ciri tertentu yang akan muncul pada grafik ACF apabila ada suatu model yang sepertinya cocok untuk digunakan pada *dataset*.

- a. Apabila terbentuk pola menyusut secara eksponensial atau adanya gelombang sinus yang kurang jelas, maka  $AR(p)$  bisa menjadi model yang bisa dipertimbangkan
- b. Apabila terdapat suatu tiang pancang yang sangat jelas pada lag- $q$ , maka model  $MA(q)$  dapat dipertimbangkan.

## 2. PACF

PACF bekerja dengan melihat korelasi antara  $Y_t$  dan  $Y_{t-k}$  dimana efek dari  $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$  diabaikan. Berbeda dengan ACF yang kurang mampu mengidentifikasi model  $AR(p)$ , PACF justru mampu mengidentifikasi model  $AR(p)$  dan kurang mampu dalam mengidentifikasi model  $MA(q)$ . Bentuk dari PACF sendiri didefinisikan sebagai berikut:

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \dots + \phi_{kk}\rho_{j-k}$$

Berkat kemampuan PACF dalam mengidentifikasi  $AR(p)$  lebih baik dibanding ACF, maka berikut adalah interpretasi identifikasi model berdasarkan grafik yang disajikan oleh grafik PACF.

- a. Apabila terbentuk pola menyusut secara eksponensial atau adanya gelombang sinus yang kurang jelas, maka  $MA(q)$  bisa menjadi model yang bisa dipertimbangkan.
- b. Apabila terdapat suatu tiang pancang yang sangat jelas pada lag- $p$ , maka model  $AR(p)$  dapat dipertimbangkan.

## 3. Extended Autocorrelation Function (EACF)

Baik PACF maupun ACF memiliki kemampuan yang baik dalam mengidentifikasi model  $AR(p)$  atau  $MA(q)$ , hanya saja apabila diminta untuk mengidentifikasi model  $ARMA(p, q)$  akan terlalu banyak kombinasi dari nilai-nilai yang tidak nol dan ini tentunya menjadi masalah dalam mengidentifikasi model  $ARMA(p, q)$  yang bersesuaian. Dengan metode EACF komponen  $AR$  akan dihilangkan dari data sehingga hanya tersisa komponen  $MA$  sehingga dapat diidentifikasi oleh ACF.

Untuk definisinya tertulis sebagai berikut:

$$W_{t,k,j} = Y_t - \tilde{\phi}_1 Y_{t-1} - \dots - \tilde{\phi}_k Y_{t-k}$$

## 4. Memilih Model Terbaik

Setelah didapatkan sekiranya pertimbangan mengenai model-model yang ingin digunakan melalui metode ACF, PACF, dan EACF berikutnya akan

ditentukan model yang terbaik dengan memperhatikan kriteria-kriteria yang memaksimalkan nilai dari estimasi *Likelihood*, di mana terbagi menjadi dua yaitu:

**a. AIC (*Akaike Information Criterion*)**

Untuk suatu  $k = p + q + 1$  akan didefinisikan

$$AIC = -2 \log(\text{maximum Likelihood}) + 2k$$

Khusus untuk model yang mengandung konstanta, maka  $k$  akan didefinisikan sebagai  $k = p + q$ .

Berikutnya akan didefinisikan pula AICc sebagai berikut:

$$AICc = AIC + \frac{2(k+1)(k+2)}{n-k-2}$$

Dalam aplikasinya, model akan semakin direkomendasikan apabila skor AIC-nya semakin kecil, dan AICc akan menjadi bahan pertimbangan akhir semisal didapatkan model dengan kriteria yang terlalu berdekatan antara satu sama lain.

**b. BIC (*Bayesian Information Criterion*)**

Untuk  $k$  regressor akan didefinisikan

$$BIC = -2 \log(\text{maximum Likelihood}) + k \log(n)$$

Semakin kecil nilai dari BIC, semakin direkomendasikan model tersebut.

### 2.4.3 Estimasi Parameter Model

Setelah melalui pemilihan model dengan memperhatikan nilai AIC dan BIC, tentunya akan didapatkan suatu model  $ARIMA(p, d, q)$  yang ingin digunakan sebagai model. Teruntuk nilai  $d$  sudah didapatkan lantaran  $d$  didefinisikan sebagai banyaknya *differencing* yang dijalankan pada data. Berikutnya akan dijelaskan mengenai tiga metode yang digunakan untuk menaksir  $ARMA(p, q)$ .

**1. Metode Momen**

Metode momen bekerja dengan menyelesaikan persamaan dari momen sampel dengan momen teoritis. Metode ini cukup bagus untuk model *AR*. Contoh penggunaannya adalah menaksir *mean* proses stasioner dengan *mean* sampel.

**2. Metode *Least-Square***

Metode ini meminimumkan jumlah kuadrat dari fungsi di bawah ini.

$$SS = \sum_{t=2}^n (Y_t - \mu - \phi(Y_{t-1} - \mu))^2$$

di mana selisih kuadrat dari  $Y_t - \mu$  dengan  $\phi(Y_{t-1} - \mu)$  akan diminimumkan untuk mencari  $\hat{\phi}$ .

### 3. Metode Maximum Likelihood

Metode ini bekerja dengan mencari nilai maksimum dari fungsi *Likelihood*. Metode ini bekerja dengan baik untuk model yang memiliki runtun waktu yang pendek dan model yang bersifat musiman. Hanya saja, dikarenakan metode ini memaksimumkan fungsi *Likelihood*, perlu diketahui *joint pdf* dari fungsi *Likelihood* yang akan dimaksimumkan, terkadang hal ini kurang efisien dan sulit untuk dicari.

Metode estimasi parameter yang digunakan pada penelitian ini adalah *Conditional Sum Square - Maximum Likelihood Estimation* (MLE). Metode ini dipilih karena memiliki konvergensi yang lebih cepat dibandingkan dengan metode MLE dengan *Unconditional Sum Square*. Sebagai contoh, untuk kasus *AR*(1), CSS-MLE akan memaksimumkan fungsi likelihood  $L(\phi, \mu, \sigma_e^2)$ , di mana

$$L(\phi, \mu, \sigma_e^2) = (2\pi\sigma_e^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma_e^2} S(\phi, \mu)\right]$$

di mana

$$S(\phi, \mu) = \sum_{t=2}^n [(Y_t - \mu) - \phi(Y_{t-1} - \mu)]^2$$

Sedangkan untuk *Unconditional Sum Square - MLE*,

$$S(\phi, \mu) = \sum_{t=2}^n [(Y_t - \mu) - \phi(Y_{t-1} - \mu)]^2 + (1 - \phi^2)(Y_1 - \mu)^2$$

Perhatikan bahwa  $(1 - \phi^2)(Y_1 - \mu)^2$  mengakibatkan persamaan  $\frac{\partial S(\phi, \mu)}{\partial \phi} = 0$  dan  $\frac{\partial S(\phi, \mu)}{\partial \mu} = 0$  menjadi non-linier terhadap  $\phi$  dan  $\mu$ , hal ini tentu saja mempengaruhi konvergensi dan nilai parameter.

#### 2.4.4 Diagnosis Model

Langkah selanjutnya adalah melakukan diagnosis model. Diagnosis model ini terbagi menjadi analisis residual dan *overfitting*.

##### 1. Analisis Residual

Dikarenakan model yang baik adalah model yang memiliki residual yang berdistribusi normal, independen, dan stasioner. Untuk itu, akan dilakukan uji normalitas residual dan uji Ljung-Box untuk menguji apakah terdapat korelasi antar residual atau tidak.

Pada tahap ini akan dilihat apakah residual dari data memenuhi sifat-sifat berikut:

- Berdistribusi Normal
- Residual stasioner
- Residual independen

Sehingga akan dijalankan pengujian pada residual meliputi Uji Normalitas dan Uji Autokorelasi pada residual.

Untuk Uji Normalitas, akan dijalankan Uji Normalitas untuk  $H_0$ : residual berdistribusi Normal untuk suatu  $\alpha$ , di mana  $H_0$  akan ditolak apabila didapatkan  $p - value < \alpha$ . Pengujian yang digunakan adalah Uji Jarque-Bera untuk menguji normalitas residual.

Untuk Uji Autokorelasi, akan dijalankan Uji Autokorelasi untuk  $H_0$ : tidak adanya multikolinearitas untuk suatu  $\alpha$ , di mana  $H_0$  akan ditolak apabila didapatkan  $p - value < \alpha$ . Pengujian yang digunakan adalah Uji Ljung-Box untuk menguji autokorelasi residual.

## 2. *Overfitting*

Model yang dipilih akan diuji dengan melakukan *fitting* dengan model yang lebih umum dari model yang terpilih untuk melihat bagaimana model tersebut bereaksi. Contohnya adalah semisal memilih model  $ARMA(1,1)$  maka akan di *fit* dengan model  $ARMA(1,2)$  untuk melihat apakah terdapat perubahan parameter yang sangat signifikan atau tidak, apabila didapatkan adanya perubahan yang signifikan, maka model masih kurang baik dan mengalami *overfitting*.

### 2.4.5 Peramalan (*Forecasting*)

*Forecasting* atau peramalan adalah tujuan akhir dari keseluruhan rangkaian analisis runtun waktu yang telah dijalankan, tentunya untuk mengetahui akurasi dari model yang telah ditetapkan, perlu diketahui variansi *error* dari *forecast* yang telah dijalankan. Maka dari itu akan didefinisikan ekspektasi dari *forecast* sepanjang  $l$  waktu kedepan sebagai berikut:

$$\hat{Y}_t(l) = E(Y_{t+l} | Y_1, Y_2, \dots, Y_t)$$

Terdapat dua model yang dapat digunakan dalam *forecasting*, yaitu:

#### 1. *Trend Deterministik*

Untuk suatu  $Y_t \sim (0, \gamma_0)$  dengan  $X_t$  sebagai komponen stokastik, akan didefinisikan model tren deterministiknya dengan taksiran *forecast*-nya sebagai berikut:

$$Y_t = \mu_t + X_t$$

$$\hat{Y}_t(l) = \mu_{t+l}$$

#### 2. *ARIMA Forecasting*

Berikut merupakan daftar dari definisi  $Y_t$  serta taksiran *forecasting* secara berurutan.

##### a. *AR(1)*

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + e_t$$

$$\hat{Y}_t(l) = \mu + \phi^l(Y_t - \mu)$$

**b.  $MA(1)$**

$$Y_t = \mu + e_t - \theta e_{t-1}$$

$$\hat{Y}_t(l) = \begin{cases} \mu - \theta e_t & ; l = 1 \\ \mu & ; l > 1 \end{cases}$$

**c.  $ARMA(p, q)$**

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

$$\begin{aligned} \hat{Y}_t(l) = & \phi_1 \hat{Y}_t(l-1) + \dots + \phi_p \hat{Y}_t(l-p) + \theta_0 - \theta_1 E(e_{t+l-1} | Y_1, Y_2, \dots, Y_t) \\ & - \theta_2 E(e_{t+l-2} | Y_1, Y_2, \dots, Y_t) - \dots - \theta_q E(e_{t+l-q} | Y_1, Y_2, \dots, Y_t) \end{aligned}$$

**d.  $ARIMA(0, 1, 0)$ : Random Walk dengan Drift**

$$\hat{Y}_t = Y_{t-1} + \theta_0 + e_t$$

$$\hat{Y}_t(l) = Y_t + \theta_0 l$$



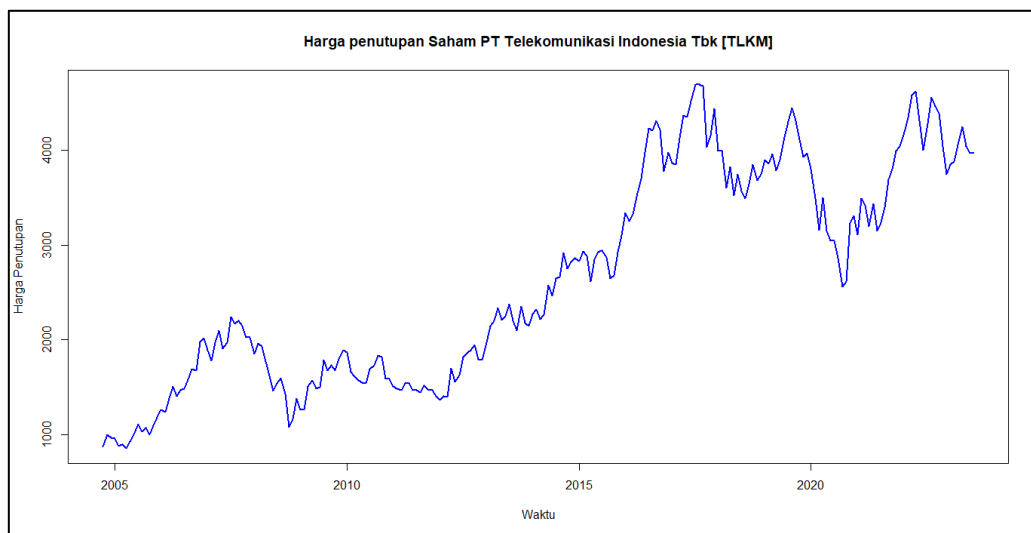
## BAB 3

### ANALISIS DATA

#### 3.1 Data

Pada penelitian ini akan dilakukan analisis untuk data nilai saham yaitu harga penutupan saham. Pada penelitian ini digunakan satu variabel yaitu data harga penutupan saham Perusahaan Perseroan PT Telekomunikasi Indonesia Tbk (TLKM). Sampel dalam penelitian ini adalah data harga penutupan saham bulanan selama 226 bulan sejak 1 Oktober 2004 hingga Juni 2023 (tanggal 16 Juni). Adapun data nilai saham dapat diakses melalui tautan berikut:

<https://finance.yahoo.com/quote/TLKM.JK/history?period1=1096329600&period2=1686960000&interval=1mo&filter=history&frequency=1mo&includeAdjustedClose=true>



*Gambar 1 Plot Data Harga Penutupan Nilai Saham TLKM*

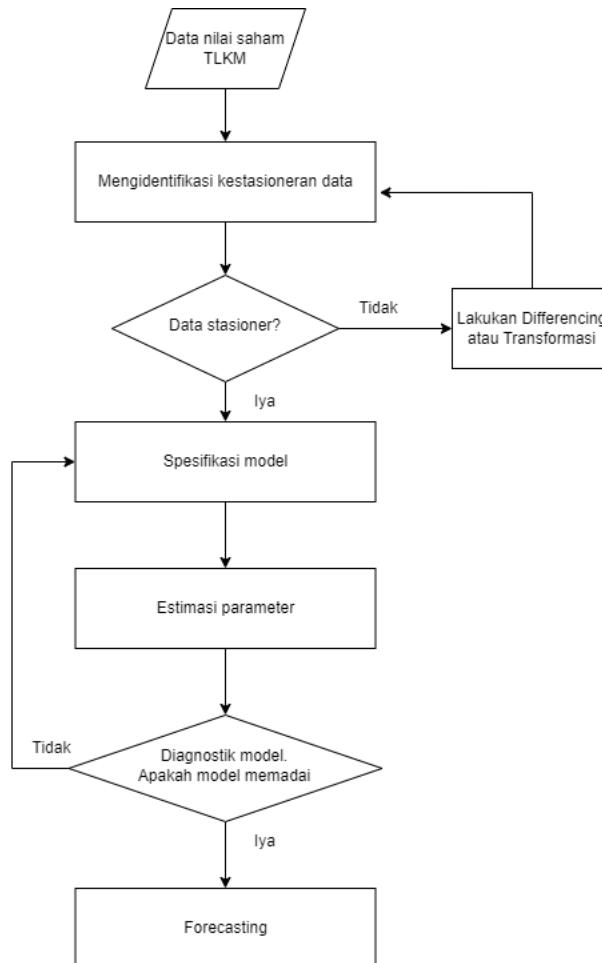
Waktu	Harga. Penutupan
Min. :2004-10-01	Min. : 855
1st Qu.:2009-06-08	1st Qu.:1602
Median :2014-02-15	Median :2345
Mean :2014-02-14	Mean :2636
3rd Qu.:2018-10-24	3rd Qu.:3772
Max. :2023-06-16	Max. :4690

*Gambar 2 Statistika Deskriptif*

Dari data yang dimiliki, terlihat bahwa harga penutupan tertinggi yaitu Rp4.690,00 pada bulan Juli dan Agustus tahun 2017, sedangkan harga penutupan terendah yaitu Rp855,00 pada bulan April 2005. Berdasarkan *plot* data harga penutupan tertinggi terlihat adanya kenaikan dan penurunan serta kenaikan dengan berjalannya waktu.

### 3.2 Tahapan Penelitian

Alur penelitian ini dapat dilihat pada diagram alur di bawah ini.



Gambar 3 Alur Tahapan Penelitian

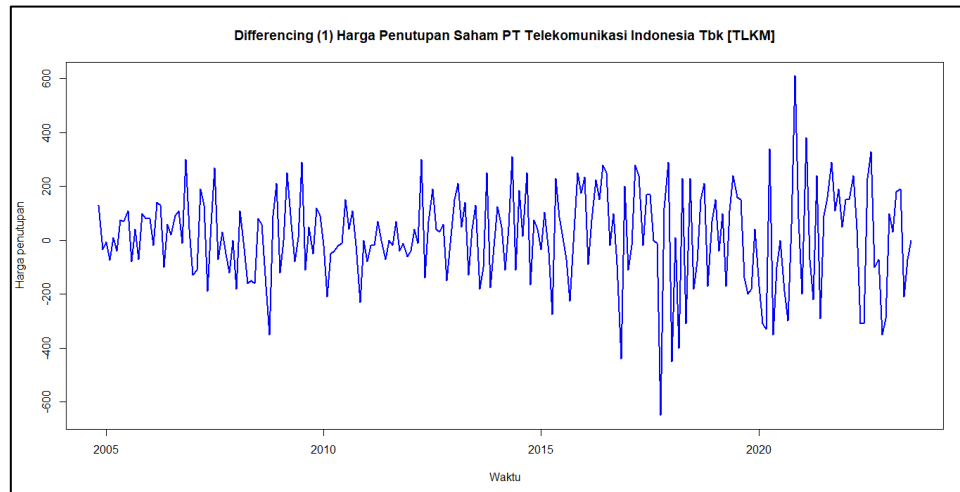
### 3.3 Uji Stasioner

Uji stasioner digunakan untuk melihat apakah data observasi yang digunakan bersifat stasioner. Data yang stasioner adalah data dengan *mean* dan variansi yang konstan sepanjang waktu. Dari **Gambar 1** dapat dilihat bahwa secara subjektif data harga penutupan saham yang dimiliki belum bersifat stasioner karena masih bisa dilihat terdapat kecenderungan naik. Untuk memastikan, akan dilakukan pengujian *Augmented Dickey-Fuller Test* atau uji ADF.

```
Augmented Dickey-Fuller Test
data: ts
Dickey-Fuller = -2.5793, Lag order = 6, p-value = 0.3327
alternative hypothesis: stationary
```

Gambar 4 Uji ADF Data Penutupan Harga Saham TLKM

Dari hasil uji ADF terlihat bahwa  $p - value = 0.3327 > 0.05$  sehingga  $H_0$  tidak ditolak dan data belum stasioner sehingga perlu dilakukan transformasi pada data dengan melakukan *differencing* terhadap data. *Differencing* adalah proses menentukan dan menggunakan selisih antara nilai observasi dengan observasi sebelumnya. *Plot* dari runtun waktu setelah *differencing* dapat dilihat pada **Gambar 5**.



Gambar 5 Plot Data Harga Penutupan TLKM Setelah Differencing ( $d = 1$ )

Dengan melihat plot pada Gambar 5, dapat dilihat bahwa data sudah bersifat stasioner. Kemudian, akan dilakukan pengujian *Augmented Dickey-Fuller Test* untuk data setelah *differencing*.

```
Augmented Dickey-Fuller Test
data: diff_ts
Dickey-Fuller = -5.1035, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary
```

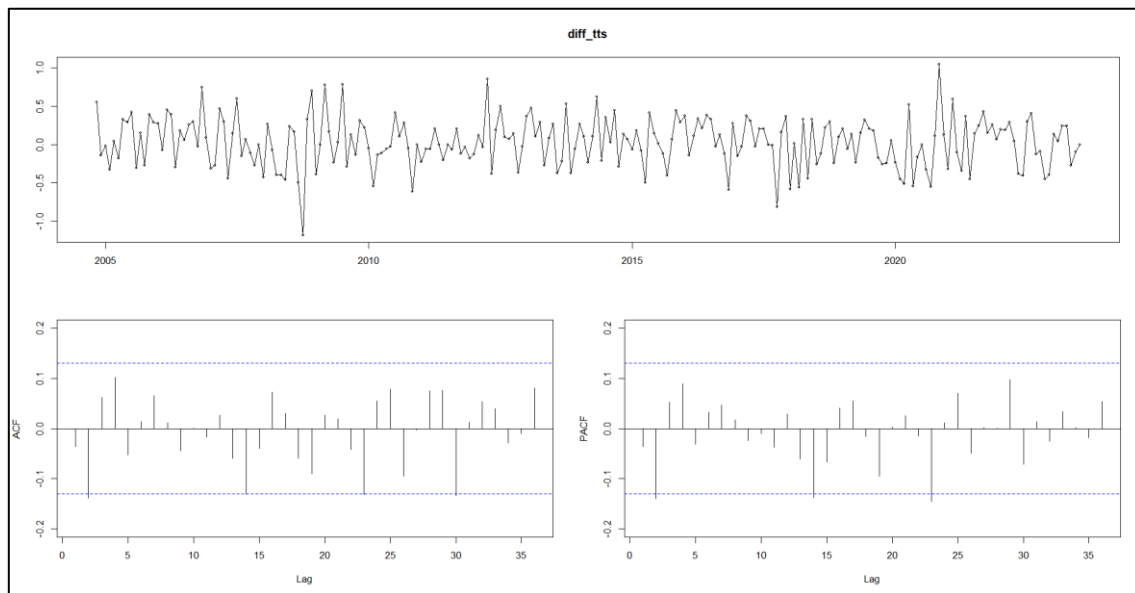
Gambar 6 Uji ADF Data Harga Penutupan TLKM Setelah Differencing

Berdasarkan **Gambar 6**, diperoleh  $p - value = 0.01 < 0.05$ , sehingga keputusan yang diambil adalah menolak  $H_0$  dan disimpulkan bahwa data observasi hasil *differencing* bersifat stasioner.

### 3.4 Spesifikasi Model

Setelah data observasi stasioner, langkah selanjutnya adalah menentukan model yang tepat. Pemilihan model ini akan dilakukan berdasarkan grafik ACF, PACF, dan EACF dari data *differencing*. Garis biru putus-putus pada grafik ACF dan PACF adalah batas dimana nilai ACF dan PACF berbeda dengan nol secara signifikan, dan dapat diperoleh dari perhitungan  $\frac{z_\alpha}{\sqrt{n}}$ .

### 3.4.1 ACF dan PACF



Gambar 4 Plot ACF dan PACF Data Harga Penutupan TLKM Setelah Differencing ( $d = 1$ )

Berdasarkan **Gambar 7**, pada *plot* ACF terlihat bahwa terdapat beberapa tiang pancang. Sementara itu, pada *plot* PACF juga terdapat beberapa tiang pancang. Pola yang didapat dari ACF dan PACF tidak menggambarkan pola AR atau MA. Maka, model runtun waktu yang diduga adalah ARIMA yang ordenya akan dicari melalui tabel EACF.

### 3.4.2 EACF

Berdasarkan tabel EACF pada **Gambar 8**, terdapat tiga model yang akan menjadi kandidat, yaitu  $ARIMA(0,1,0)$ ,  $ARIMA(1,1,1)$ , dan  $ARIMA(0,1,2)$ . Selanjutnya, ketiga model tersebut akan diseleksi untuk memperoleh model terbaik.

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	o	x	o	o	o	o	o	o	o	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	o	o	o	o	o	o	o	o	o	o	o	o
3	x	x	x	o	o	o	o	o	o	o	o	o	o	o
4	x	x	x	o	o	o	o	o	o	o	o	o	o	o
5	x	o	x	o	x	o	o	o	o	o	o	o	o	o
6	x	x	o	o	x	o	o	o	o	o	o	o	o	o
7	x	x	x	o	o	o	o	o	o	o	o	o	o	o

Gambar 5 Tabel EACF Data Harga Penutupan TLKM Setelah Differencing ( $d = 1$ )

### 3.4.3 Pemilihan Model Terbaik

Pemilihan model terbaik didasari atas nilai *Log-Likelihood* terbesar, nilai *Akaike Information Criterion* (AIC) terkecil dan *Bayesian Information Criterion* (BIC) terkecil. Metode pemilihan dilakukan dengan mendefinisikan semua model *ARIMA*

yang menjadi kandidat model terbaik, kemudian membandingkan nilai *Log-Likelihood*, AIC, dan BIC.

```
model1 <- Arima(tf_ts, order = c(0,1,0), include.constant = FALSE)
model2 <- Arima(tf_ts, order = c(1,1,1), include.constant = TRUE)
model3 <- Arima(tf_ts, order = c(0,1,2), include.constant = TRUE)
```

*Gambar 6 Code Pembentukan dan Perbandingan Kriteria Model Dugaan*

	model1	model2	model3
coef	numeric,0	numeric,3	numeric,3
sigma2	0.105439	0.1053358	0.1040974
var.coef	numeric,0	numeric,9	numeric,9
mask	logical,0	logical,3	logical,3
loglik	-66.17767	-64.56055	-63.24037
aic	134.3553	137.1211	134.4807
arma	integer,7	integer,7	integer,7
residuals	ts,226	ts,226	ts,226
call	expression	expression	expression
series	"tf_ts"	"tf_ts"	"tf_ts"
code	0	0	0
n.cond	0	0	0
nobs	225	225	225
model	list,10	list,10	list,10
aicc	134.3733	137.3029	134.6626
bic	137.7714	150.7855	148.1451
xreg	ts,226	integer,226	integer,226
x	ts,226	ts,226	ts,226
fitted	numeric,0	ts,226	ts,226

*Gambar 7 Tabel Perbandingan Nilai Kriteria Model Dugaan*

Berdasarkan hasil pada **Gambar 10**, model 1 merupakan model yang memiliki nilai AIC, BIC, dan *Log-Likelihood* terkecil, sehingga model *ARIMA*(0,1,0) adalah model yang terbaik yang akan digunakan.

Selain itu, dapat digunakan juga fungsi *auto.arima* pada *R* untuk mengetahui model yang paling baik.

```

> auto.arima(ts, trace = TRUE)

Fitting models using approximations to speed things up...

ARIMA(2,1,2)(1,0,1)[12] with drift      : Inf
ARIMA(0,1,0) with drift      : 2951.215
ARIMA(1,1,0)(1,0,0)[12] with drift      : 2965.642
ARIMA(0,1,1)(0,0,1)[12] with drift      : 2954.523
ARIMA(0,1,0) with drift      : 2950.602
ARIMA(0,1,0)(1,0,0)[12] with drift      : 2962.919
ARIMA(0,1,0)(0,0,1)[12] with drift      : 2952.629
ARIMA(0,1,0)(1,0,1)[12] with drift      : 2964.915
ARIMA(1,1,0) with drift      : 2953.677
ARIMA(0,1,1) with drift      : 2953.098
ARIMA(1,1,1) with drift      : 2955.61

Now re-fitting the best model(s) without approximations...

ARIMA(0,1,0) : 2960.915

Best model: ARIMA(0,1,0)

Series: ts
ARIMA(0,1,0)

sigma^2 = 30115: log likelihood = -1479.45
AIC=2960.9 AICc=2960.91 BIC=2964.31

```

Gambar 11 Hasil *auto.arima* pada Data untuk Menentukan Model Terbaik

Sehingga, berdasarkan hasil pengecekan yang diperoleh pada **Gambar 11**, telah dipastikan model terbaik yang akan digunakan adalah *ARIMA(0,1,0)* tanpa *drift*, yang mana merupakan *random walk*.

### 3.5 Estimasi Parameter

Metode estimasi parameter yang digunakan untuk data Penutupan Harga Saham adalah penggunaan *conditional sum of square* untuk menentukan nilai awal kemudian diestimasi dengan memanfaatkan metode *maximum-likelihood* (CSS-ML).

```

Series: tf_ts
ARIMA(0,1,0)

sigma^2 = 0.1054: log likelihood = -66.18
AIC=134.36 AICc=134.37 BIC=137.77

```

Gambar 12 Estimasi Parameter Model *ARIMA(0,1,0)* untuk Data Harga Penutupan TLKM

Diperoleh estimasi yang dapat diekspresikan sebagai berikut:

$$\sigma_e^2 = 0.1054$$

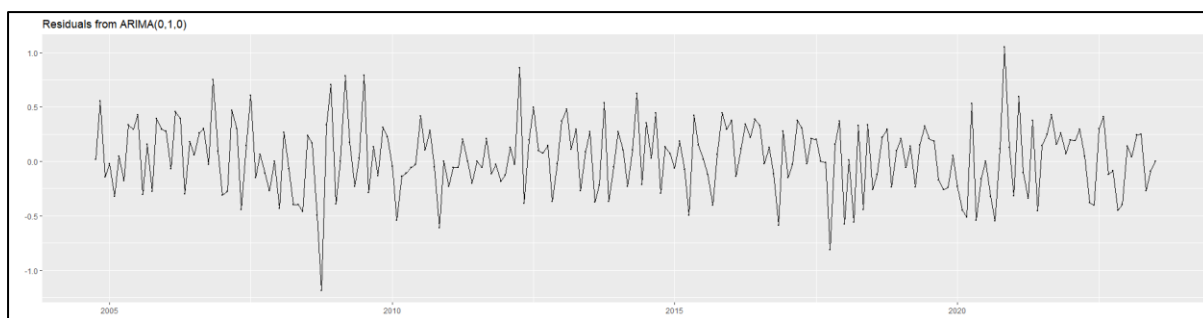
$$\nabla Y_t = e_t$$

### 3.6 Model Diagnostik

#### 3.6.1 Analisis Residual

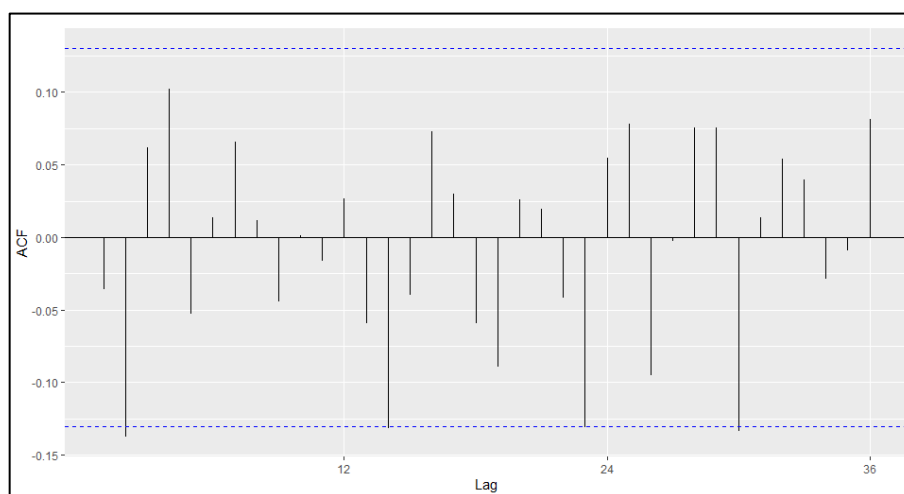
##### 1. Uji Independensi/Autokorelasi Residual

Pada tahapan analisis residual, dilakukan pemeriksaan terhadap stasioneritas residual model beserta dengan independensi residual model. Gambar berikut merupakan plot dari residual.



*Gambar 13 Plot Runtun Waktu Residual Model ARIMA(0,1,0) Harga Penutupan Saham TLKM*

Secara subjektif, dari **Gambar 13**, dapat dilihat bahwa *plot* yang dihasilkan stasioner karena tidak terlihat kecenderungan rata-rata yang meningkat atau menurun seiring waktu. Selain itu, juga tidak terlihat adanya bentuk deterministik dari *plot* residual yang terbentuk. Hasil ini juga didukung dari bentuk plot ACF yang diperoleh seperti yang terlihat pada **Gambar 14**.



*Gambar 14 Plot ACF Residual Model ARIMA(0,1,0) Saham TLKM*

Berdasarkan **Gambar 14**, autokorelasi lag- $k$  dikatakan tidak berbeda signifikan dari 0 apabila tidak melewati garis putus-putus. Dapat dilihat bahwa seluruh nilai autokorelasi hingga lag-36 tidak melewati garis putus-putus tersebut. Sehingga dari hasil ini dapat dikatakan bahwa autokorelasi dari residual model ARIMA (0,1,0) untuk data harga penutupan saham tidak berbeda signifikan dari 0 dan memenuhi salah satu syarat kestasioneran data runtun waktu.

```

Ljung-Box test

data:  Residuals from ARIMA(0,1,0)
Q* = 26.099, df = 24, p-value = 0.3481

Model df: 0.    Total lags used: 24

```

*Gambar 15 Uji Ljung-Box Analisis Residual Model ARIMA(0,1,0) Harga Penutupan TLKM*

Dari **Gambar 15**, dapat dilihat bahwa  $p - value = 0.3481 > 0.05$ . Diambil keputusan untuk tidak menolak  $H_0$ , sehingga disimpulkan data residual  $\forall \rho_k = 0$  (residual saling independen, dikarenakan nilai autokorelasinya yang tidak berbeda signifikan dengan 0).

```

Augmented Dickey-Fuller Test

data:  fit$residuals
Dickey-Fuller = -5.0943, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary

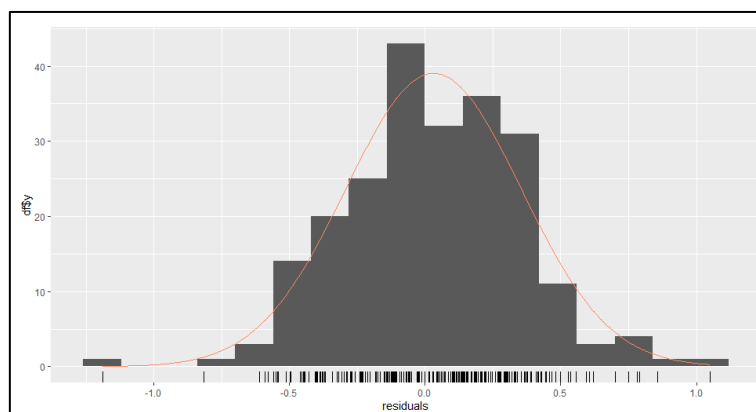
```

*Gambar 16 Uji ADF Analisis Residual Model ARIMA(0,1,0) Saham TLKM*

Dari **Gambar 16**, dapat dilihat bahwa  $p - value = 0.01 < 0.05$ . Diambil keputusan untuk menolak  $H_0$ , sehingga disimpulkan data residual stasioner. Dari analisis ini dapat disimpulkan bahwa residual model runtun waktu *ARIMA* (0,1,0) untuk data Saham TLKM stasioner dan independen.

## 2. Uji Kenormalan Residual

Selanjutnya dilakukan uji normalitas yaitu uji untuk menentukan apakah residual model yang dibentuk mengikuti distribusi Normal atau tidak. Berdasarkan *plot* yang diperoleh pada **Gambar 17**, dapat dilihat bahwa residual model menyerupai distribusi Normal.

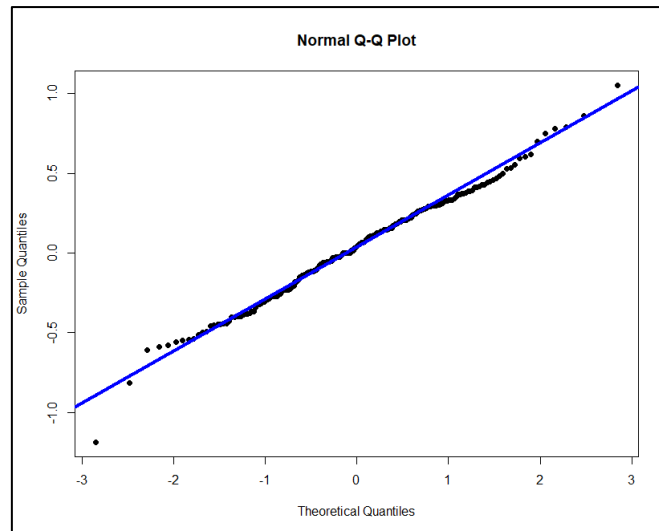


*Gambar 17 Histogram Frekuensi Residual Data Saham TLKM*

Kemudian, secara subjektif dapat dilihat dari **Gambar 18** bahwa *Q-Q plot* dari residual model menunjukkan persebaran data tidak sepenuhnya mengikuti garis

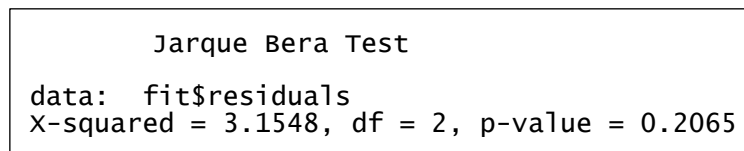


diagonal, di mana garis diagonal tersebut merepresentasikan data yang berdistribusi Normal. Dengan kata lain, secara subjektif, melalui persamaan Q-Q *plot* dapat disimpulkan residual model tidak berdistribusi Normal.



Gambar 18 Q-Q Plot Residual Data Saham TLKM

Terdapat perbedaan hasil interpretasi antara histogram dan Q-Q *plot*, sehingga perlu dilakukan uji statistik untuk normalitas data. Akan digunakan uji Jarque-Bera (*JB Test*) untuk memastikan kenormalan residual model. Diperoleh hasil pengujian seperti pada **Gambar 19**.



Gambar 19 Uji Jarque Bera Residual Data Saham TLKM

Dengan menggunakan taraf signifikansi sebesar 5%, dapat dilihat bahwa  $p - value = 0.2065 > 0.05$  maka  $H_0$  gagal ditolak dan dapat disimpulkan dari aturan keputusan bahwa data residual berdistribusi Normal.

### 3.6.2 Overfitting

*Overfitting* bertujuan untuk melihat apakah model yang diajukan apakah model yang paling baik, atau masih ada model lain yang lebih baik dari model yang diajukan. Pada kasus ini, *overfitting* akan dilakukan dengan mengajukan model *ARIMA* (0,1,1) dan *ARIMA* (1,1,0). Model *ARIMA* (0,1,1) didefinisikan sebagai overfit1, Model *ARIMA* (1,1,0) didefinisikan sebagai overfit2. Akan dibandingkan ketiga model sebagai berikut.

	fit	overfit1	overfit2
coef	numeric,0	-0.03681815	-0.02733988
sigma2	0.04063222	0.04077336	0.0407841
var.coef	numeric,0	0.005986149	0.00450805
mask	logical,0	TRUE	TRUE
loglik	40.91707	41.02914	40.99994
aic	-79.83414	-78.05828	-77.99989
arma	integer,7	integer,7	integer,7
residuals	ts,225	ts,225	ts,225
call	expression	expression	expression
series	"ts"	"tf_ts"	"tf_ts"
code	0	0	0
n.cond	0	0	0
nobs	224	224	224
model	list,10	list,10	list,10
aicc	-79.81612	-78.00399	-77.94559
bic	-76.4225	-71.23499	-71.1766
lambda	0.1402781	ts,225	ts,225
x	ts,225	ts,225	ts,225
fitted	ts,225	-0.03681815	-0.02733988

Gambar 20 Perbandingan Model Fit dengan Model Overfit

Kriteria *overfitting* akan digunakan dengan membandingkan nilai *log-likelihood*, AIC, dan BIC. Berdasarkan **Gambar 20**, dapat dilihat bahwa model *fit* ARIMA(0,1,0) memiliki nilai AIC dan BIC yang paling kecil jika dibandingkan dengan model *overfit*. Sementara itu, nilai *log-likelihood* yang didapat dari model *fit* merupakan yang paling kecil, namun nilainya tidak berbeda terlalu signifikan dengan model *overfit*.

Sehingga dapat disimpulkan model *fit* ARIMA(0,1,0) merupakan model terbaik yang akan digunakan untuk memprediksi Saham TLKM. Untuk memperkuat pernyataan ini, dapat dilakukan uji koefisien antara model *fit* dan *overfit*. Namun, kami tidak dapat dilakukan uji koefisien karena model *random walk* tidak memuat koefisien.

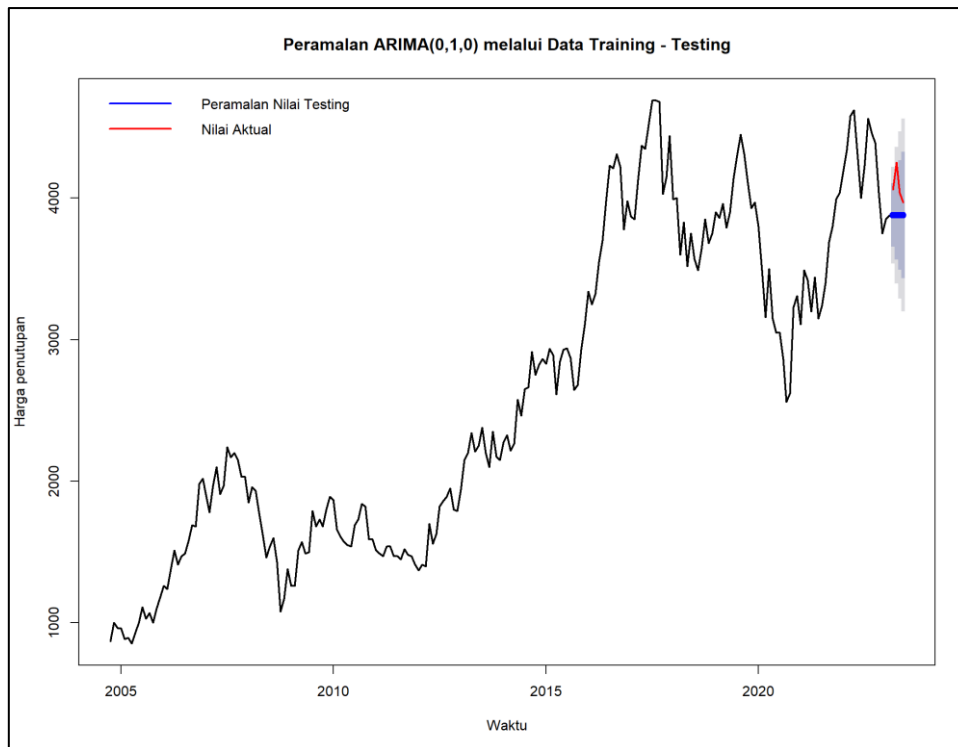
### 3.7 Peramalan (*Forecasting*)

#### 3.7.1. *Ex Post*

Akan dilihat akurasi *forecast* dengan metode *cross validation*. Metode *cross validation* membagi data menjadi dua, yaitu data *train* dan data *test*. Data *train* (Oktober 2004 – Februari 2023) adalah data yang digunakan untuk membentuk model, sedangkan data *test* (Maret 2023 – Juni 2023) adalah data yang digunakan sebagai pembandingan dengan hasil *forecast* yang dimiliki.

	test	fc_train.Point	Forecast	fc_train.Lo 80	fc_train.Hi 80	fc_train.Lo 95	fc_train.Hi 95
Mar 2023	4060		3880	3657.050	4102.950	3539.027	4220.973
Apr 2023	4250		3880	3564.701	4195.299	3397.791	4362.209
May 2023	4040		3880	3493.839	4266.161	3289.417	4470.583
Jun 2023	3970		3880	3434.099	4325.901	3198.054	4561.946

Gambar 21 Perbandingan Nilai Aktual dan Nilai Forecast Harga Penutupan Saham TLKM



Gambar 22 Plot Nilai Aktual dan Nilai Forecast Harga Penutupan Saham TLKM

Berdasarkan **Gambar 21** dan **Gambar 22**, terlihat bahwa seluruh nilai aktual berada pada interval kepercayaan 95%. Hal ini dapat disimpulkan bahwa model yang diajukan dapat memprediksi data Harga Penutupan Saham TLKM cukup baik.

<i>i</i>	<i>Absolute Error</i>	<i>Relative Error (%)</i>	<i>Mean Absolute Error</i>
223	180	4.433498	200
224	370	8.705882	
225	160	3.960396	
226	90	2.267003	

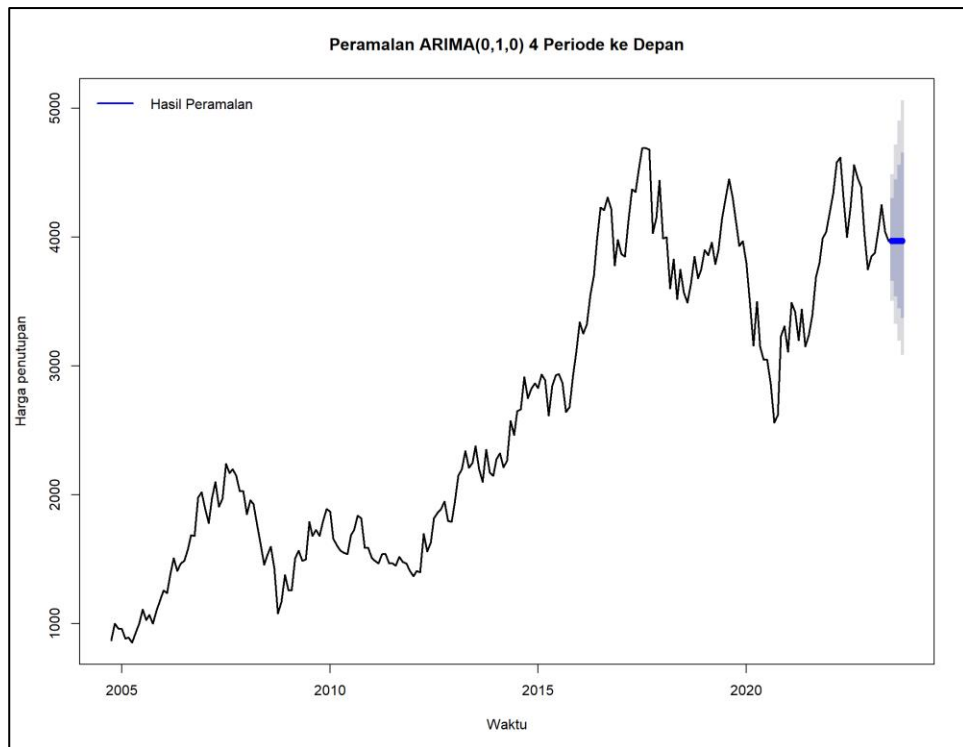
Tabel 1 Presisi antara Nilai Aktual dan Nilai Forecast Harga Penutupan Saham TLKM

### 3.7.2. Ex Ante

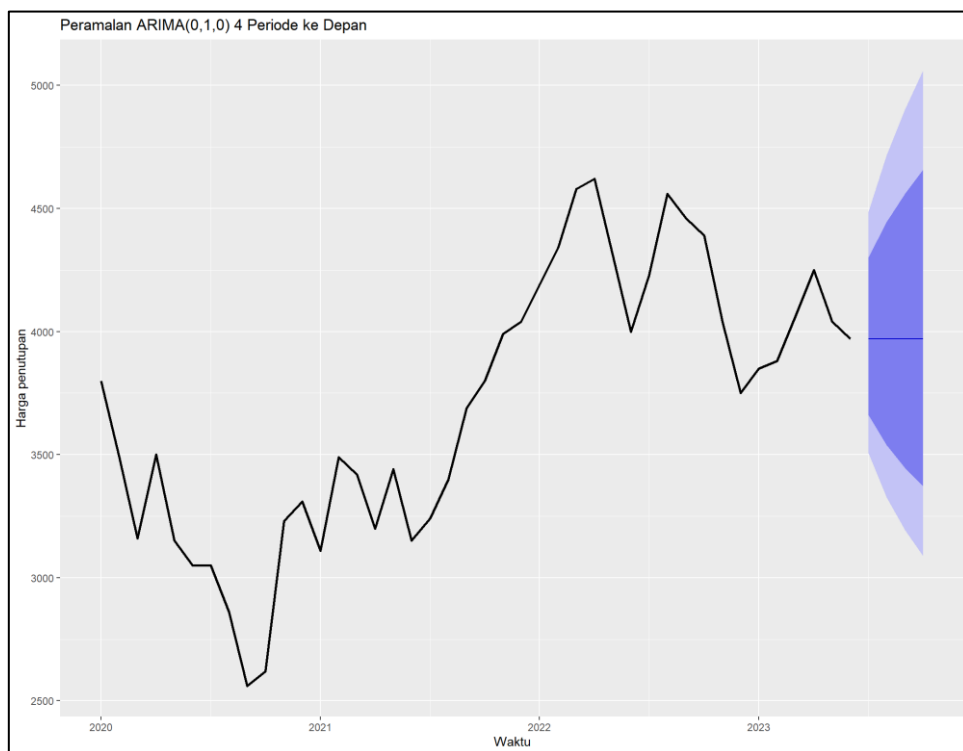
Selanjutnya dilakukan peramalan untuk meramal harga penutupan saham TLKM untuk 4 bulan ke depan yaitu bulan Juli 2023 – Oktober 2023. Berikut ini adalah hasil *forecasting*.

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jul 2023	3970	3660.201	4302.079	3504.789	4487.340
Aug 2023	3970	3538.104	4446.467	3326.296	4718.018
Sep 2023	3970	3446.806	4560.066	3194.686	4901.867
Oct 2023	3970	3371.418	4657.764	3087.237	5061.612

Gambar 23 Hasil Forecasting



*Gambar 24 Plot Hasil Forecasting*



*Gambar 25 Plot Hasil Forecasting Dimulai dari Tahun 2020*

## BAB 4 PENUTUP

### 4.2. Kesimpulan

Kesimpulan yang kami peroleh dari penelitian ini di antaranya:

1. Gambaran umum pada data harga penutupan saham PT. Telkom Indonesia (Persero) Tbk [TLKM] pada periode bulan Oktober 2004 sampai Juni 2023 (tanggal 16) yakni harga penutupan tertinggi yaitu Rp4.690,00 pada bulan Juli dan Agustus tahun 2017, sedangkan harga penutupan terendahnya yaitu Rp855,00 pada bulan April 2005, dengan nilai rata-ratanya sebesar Rp2.636,00. Terjadi tren menurun yang terlihat signifikan pada awal tahun 2020, hal ini disebabkan oleh adanya pandemi *COVID-19* yang hampir membuat semua harga saham mengalami penurunan.
2. Proses analisis data runtun waktu dilakukan pada data yang stasioner, yang mana memerlukan suatu transformasi variabel (*differencing*), selanjutnya dilakukan langkah spesifikasi model dan diagnostik model untuk mendapatkan model terbaik untuk data harga penutupan saham PT. Telkom Indonesia (Persero) Tbk [TLKM].
3. Diperoleh Model runtun waktu terbaik untuk memprediksi harga penutupan saham Telkom Indonesia (Persero) Tbk [TLKM] per bulan adalah  $\nabla Y_t = \nabla Y_{t-1} + e_t$  di mana  $\nabla Y_t$  adalah *first difference* harga penutupan saham TLKM pada bulan ke- $t$  dan bulan ke- $(t - 1)$ .
4. Hasil peramalan yang diperoleh pada data harga penutupan saham (*Closing Price*) PT. Telkom Indonesia (Persero) Tbk [TLKM] dapat diperoleh dengan model ARIMA(0,1,0) yang telah didapatkan. Sehingga hasil peramalan yang diperoleh untuk 4 bulan kedepan adalah konstan yakni sebesar Rp3.970,00.

### 4.3. Saran

Meskipun telah didapatkan model ARIMA yang cukup baik akan tetapi masih terdapat beberapa penetapan dan keputusan yang sifatnya subjektif dan masih bisa diperbaiki lagi, misalnya terdapat beberapa model *seasonal* yang tidak dipertimbangkan dalam pemilihan model. Selain itu masih banyak faktor eksternal lain yang seharusnya ikut dipertimbangkan dalam pembentukan model lantaran subjek yang dibahas menyangkut harga saham sangat berhubungan dengan kondisi fundamental perusahaan, isu yang beredar di pasar, aksi korporasi perusahaan, adanya kepentingan dari investor tertentu, kondisi ekonomi seperti krisis, resesi atau pandemi serta kebijakan pemerintah.

## DAFTAR PUSTAKA

- Awal, S. (2022, Desember 31). *6 Faktor Yang Mempengaruhi Harga Saham*. Diambil kembali dari Stockbit Snips: <https://snips.stockbit.com/investasi/faktor-yang-mempengaruhi-harga-saham>
- Marzaen, M. Y. (2019). *PENERAPAN FUZZY TIME SERIES-CHEN DALAM DATA HARGA PENUTUPAN SAHAM PT BANK NEGARA INDONESIA (Persero) Tbk. (IDX : BBNI)*. Yogyakarta: Universitas Islam Indonesia.
- PT Bursa Efek Indonesia. (t.thn.). *Saham*. Diambil kembali dari PT Bursa Efek Indonesia: <https://www.idx.co.id/id/produk/saham>
- Santosa, U. A. (7, Mei 2023). *Mengenal Saham Telkomsel dan Rekam Jejak Perusahaan Induknya*. Diambil kembali dari BMmoey: <https://bmoney.id/blog/saham-tlkm-121449>
- Telkom Indonesia. (t.thn.). *Tentang Telkomgroup*. Diambil kembali dari Telkom: [https://www.telkom.co.id/sites/about-telkom/id\\_ID/page/profil-dan-riwayat-singkat-22](https://www.telkom.co.id/sites/about-telkom/id_ID/page/profil-dan-riwayat-singkat-22)

## LAMPIRAN

### Script R

```
#### Package ####
library(TSA)
library(forecast)
library(tseries)
library(ggplot2)
library(lmtest)

#### Import Data ####
tlkm <- read.csv("E:\\Documents\\OneDrive - UNIVERSITAS INDONESIA
\\Semester 4\\[SCST602201] Metode Peramalan\\Dataset\\TLKM.JK.csv")
tlkm$Date <- as.Date(tlkm$Date)
View(tlkm)

#### Identifikasi Data ####
summary(tlkm)
ts <- ts(tlkm[,5],
        start = c(2004,10),
        end = c(2023,6),
        frequency = 12)
ts

#### Plot Data ####
tsdisplay(ts)
plot(ts,
      xlab = "waktu",
      ylab = "Harga Penutupan",
      main = "Harga penutupan Saham PT Telekomunikasi Indonesia Tbk [TLKM]",
      lwd = 2,
      col = "blue")

#### Uji Stasioner ####
# ADF
adf.test(ts)

#### Differencing Model ####
# Differencing
diff_ts <- diff(ts,
                differences = 1)
diff_ts

# Plot
tsdisplay(diff_ts)
BoxCox.ar(ts)
plot(diff_ts,
      xlab = "waktu",
      ylab = "Harga penutupan",
      main = "Differencing (1) Harga Penutupan Saham PT Telekomunikasi
Indonesia Tbk [TLKM]",
      lwd = 2,
      col = "blue")

# Uji Stasioner (ADF test)
adf.test(diff_ts)

# Transformasi setelah melihat visualisasi grafik
tf_ts <- BoxCox(ts, BoxCox.lambda(ts))
```

```

diff_tts <- diff(tf_ts, 1)
tsdisplay(diff_tts)

adf.test(diff_tts)

#### Model ARIMA ####
# ACF dan PACF
tsdisplay(diff_tts)

#EACF
eacf(diff_tts)

#Calon model: ARIMA(0,1,0), ARIMA(1,1,1), dan ARIMA(0,1,2)
model1 <- Arima(tf_ts, order = c(0,1,0), include.constant = FALSE)
model2 <- Arima(tf_ts, order = c(1,1,1), include.constant = TRUE)
model3 <- Arima(tf_ts, order = c(0,1,2), include.constant = TRUE)
cbind(model1, model2, model3)

fit <- Arima(tf_ts, order = c(0,1,0), include.constant = FALSE)
fit

auto.arima(ts, trace = TRUE)

#### Diagnosis Model ####
# Uji Stasioner
qqnorm(fit$residuals, pch = 16)
qqline(fit$residuals, lwd = 4, col = "blue")

adf.test(fit$residuals)

# Jarque-Bera
jarque.bera.test(fit$residuals)

# Independensi (Hanya berlaku asumsi kenormalan)
checkresiduals(fit)
adf.test(fit$residuals)

#### Overfitting ####
overfit1 <- Arima(tf_ts, order = c(0,1,1), include.constant = FALSE)
overfit2 <- Arima(tf_ts, order = c(1,1,0), include.constant = FALSE)
cbind(fit, overfit1, overfit2)

#### Forecasting ####
train <- window(ts, end = c(2023,2))
test <- window(ts, start = c(2023,3))

train_fit <- Arima(train, order = c(0,1,0), include.constant = FALSE)
fc_train <- forecast(train_fit, 4)
fc_train
cbind(test, fc_train)

plot(fc_train,
     fcol = "blue",
     lwd = 2,
     main = "Peramalan ARIMA(0,1,0) melalui Data Training - Testing",
     xlab = "waktu",
     ylab = "Harga penutupan")
lines(test,
     col = "red",
     lwd = 2)

```



```

legend("topleft",
      col = c("blue", "red"),
      legend = c("Peramalan Nilai Testing","Nilai Aktual"),
      lwd = 2,
      bty = "n")

fit <- Arima(ts, order = c(0,1,0), include.constant = FALSE, lambda =
BoxCox.lambda(ts))

forecast_final <- forecast(fit, h = 4)
forecast_final
plot(fcol = "blue",
      forecast_final,
      main = "Peramalan ARIMA(0,1,0) 4 Bulan ke Depan",
      xlab = "waktu",
      ylab = "Harga penutupan",
      lwd = 2)
legend("topleft",
      col = "blue",
      legend = "Hasil Peramalan",
      lwd = 2,
      bty = "n")

win <- window(ts, start = c(2020,1))
autoplot(win, main = "Peramalan ARIMA(0,1,0) 4 Periode ke Depan",
          xlab = "waktu",
          ylab = "Harga penutupan",
          lwd = 1) +
  autolayer(forecast_final)

```