**Peramalan dengan Metode Regresi *Time Series*, ARIMA, *Neural Network*, dan *Deep Neural Network* Terhadap Minat Pencarian Grab dan Gojek di Indonesia dan Malaysia Menggunakan Data *Google Trend***

**1Rahma Metrikasi, 2Dr. Suhartono, S.Si., M.Sc., 3Ulil Azmi, S.Si., M.Sc**

1,2 Departemen Statistika, Fakultas Matematika, Komputasi dan Sains Data

3Departemen Aktuaria, Fakultas Matematika, Komputasi dan Sains Data

1,2,3, Institut Teknologi Sepuluh Nopember

Jl. Arief Rahman Hakim, Surabaya, Indonesia 60111

e-mail: 1rahmametrikasari@gmail.com [2suhartono@](mailto:2suhartono@)statistika.its.ac.id, [3ulilazmi0211@gmail.com](mailto:3ulilazmi0211@gmail.com).

**Abstrak**

Berkembangnya teknologi membuat segala hal semakin mudah. Salah satu kemudahan yang ditawarkan yaitu dalam bidang transportasi dengan bermunculannya layanan transportasi berbasis *online* . Di Indonesia, terdapat 2 *Start Up* yang merajai layanan transportasi yaitu Grab dan Gojek. Oleh karena itu, pada penelitian ini akan dilakukan *forecast* minat pengguna layanan Grab dan Gojek dengan data pencarian Grab dan Gojek di *Google Trend* di negara Indonesia dan Malaysia. Adapun metode yang digunakan untuk *forecasting* yaitu regresi time series, ARIMA*, Neural Network*, dan *Deep Neural Network*. Dimana hasil akurasi dari keempat metode tersebut akan dibandingkan untuk mengetahui metode apa yang paling baik untuk digunakan pada tiap data. Dari hasil analisis yang dilakukan, metode yang memberikan hasil akurasi *forecasting* terbaik untuk data Gojek Indonesia adalah metode *Neural Network*, untuk data Grab Indonesia metode yang memberikan hasil akurasi *forecasting* terbaik adalah metode regresi *time series*. Sedangkan untuk data Grab Malaysia metode yang memberikan hasil akurasi terbaik adalah metode regresi *time series.*

1. **Pendahuluan**

Perkembangan teknologi yang pesat dalam beberapa tahun terakhir, membuat berbagai peluang bisnis melalui teknologi semakin menjanjikan. Hal tersebut juga didukung dengan jumlah pengguna internet yang semakin banyak tiap tahun nya. Bahkan pada akhir tahun 2018, komisi telekomunikasi internasional (international telecommunication union) memprediksi jumlah pengguna internet mencapai 3,9 miliar atau sekitar 51,2 persen dari total populasi dunia. Penerapan teknologi dan informasi telah menciptakan perubahan yang cukup signifikan dalam bidang bisnis. Salah satu contoh dampak penerapan teknologi dalam bidang bisnis yaitu jumlah *offline store* yang semakin berkurang karena maraknya *online store* yang lebih praktis dan disukai oleh masyarakat. Tidak hanya *online store,* bisnis yang bergerak dalam bidang jasa lain seperti layanan transportasi berbasis *online* juga makin marak dan di minati masyarakat. Salah satu layanan transportasi *online* terbesar di Asia Tenggara adalah Grab.

Grab merupakan aplikasi layanan transportasi online yang didirikan oleh Anthony Tan dan Hooi Ling Tan yang merupakan warga negara Malaysia, sebagai sebuah solusi dari tidak efisiennya system transportasi yang ada di Malaysia pada saat itu. Pada awal kemunculannya, Grab hanya memiliki satu jenis layanan yaitu Grab Taxi namun kini jenis layanan yang ada di Grab semakin banyak yaitu Grab Bike, Grab Express, Grab Car, Grab Food, dan fitur terbaru nya yaitu Belanjaan. Dengan kemudahan yang ditawarkan dan harga yang bersaing, membuat jumlah pengguna Grab khususnya di Indonesia pun meningkat. Dengan inovasi dan kualitas layanan yang semakin baik, Grab menjadi salah satu *start up* di Asia Tenggara yang berhasil mencapai status *Decacorn* pada pertengahan 2019 lalu. Keberhasilan tersebut tentunya tidak terlepas dari minat konsumen dalam menggunakan layanan aplikasi Grab. Sedangkan GoJek adalah sebuah layanan transportasi online yang didirikan oleh Nadiem Makarim warga negara Indonesia. Pada awalnya, layanan utama yang ditawarkan GoJek yaitu Go-Ride namun seiring berjalannya waktu, layanan GoJek semakin bertambah. Kini, GoJek menawarkan empat jasa layanan yang bisa dimanfaatkan oleh para pelanggannya diantaranya yaitu Pengantaran Barang, Transportasi Jasa Angkutan, *Shopping* dan *Corporate*. Berdasarkan data comScore, jumlah pengguna Grab pada bulan Desember 2017 mencapai 9,6 juta pengguna sedangkan jumlah pengguna GoJek mencapai 9,7 juta pengguna. Di Indonesia sendiri, persaingan antara Grab dan GoJek dalam memperebutkan penumpang sangat sengit mulai dari pereng harga sampai perbaikan kualitas layanan yang semakin baik. Bahkan pada tahun yang sama yaitu pada tahun 2019, kedua *start up* dalam bidang transportasi tersebut menyandang status *Decacorn*.

Berdasarkan hasil survei *Consumers Awareness*, konsumen lebih banyak menggunakan Grab setidaknya hingga kuartal IV-2018. Sebanyak 34% pengguna GrabCar, salah satu layanan dari Grab, menyebutkan bahwa mereka menggunakan layanan sebanyak 3-4 kali per minggu. Sementara itu, 25% pengguna Go-Car cenderung hanya menggunakan layanan 1-2 kali dalam seminggu. Namun pada kategori roda dua, Go-Ride masih menjadi pilihan utama pengguna transportasi *online*. Dari total responden yang memilih Go-Ride, sebanyak 64% menggunakannya hingga 1-2 kali sehari, sedangkan pemilih Grab Bike yang menggunakan 1-2 kali daIam sehari ada 58%.

Menurut alexa.com, sebuah situs milik amazon.com yang mengumpulkan data berupa kebiasaan berselancar penggunanya di internet, dalam lingkup nasional persentase pengunjung website Grab di Indonesia sebesar 31,8 persen. Sedangkan persentase pengunjung website GoJek sebesar 78,2 persen. Persentase tersebut mengalami penurunan setiap bulannya. Hal tersebut dikarenakan banyak konsumen yang merasa tidak puas akibat kenaikan tarif yang sangat tinggi atau tidak mendapatkan *driver* karena semua *driver* sedang sibuk pada saat terjadi *high demand*. Oleh karena itu, pada penelitian ini akan dilakukan analisis tentang peramalan minat pencarian Grab sehingga dapat diketahui seperti apa perilaku dan pola minat masyarakat di Indonesia terhadap layanan Grab dan Go Jek dan membandingkan perilaku dan pola minat masyarakat terhadap layanan Grab di Indonesia dan Malaysia sehingga dapat diketahui pada kapan *high demand* akan terjadi sehingga ketidaknyamanan konsumen akibat *high demand* dapat diminimalisir.

Pada penelitian ini, data yang digunakan adalah data Google Trend tentang minat pencarian Grab di dua negara yaitu Malaysia dan Indonesia sehingga dapat dibandingkan seperti apa perilaku dan pola minat masyarakat terhadap Grab di kedua negara tersebut. Metode peramalan yang digunakan pada penelitian ini yaitu ARIMA, *Neural Network* , Regresi Time Series, dan *Deep Neural Network* untuk kemudian dibandingkan hasil akurasi nya dan diketahui motode peramalan yang memberikan hasil paling baik untuk meramalkan minat masyarakat di Indonesia dan Malaysia terhadap Grab.

1. **Metodologi Penelitian** 
   1. **Sumber Data**

Data yang digunakan dalam penelitian ini adalah data sekunder yang diunduh dari laman web *Google Trend* tentang minat pencarian Gojek Indonesia dan Grab di Indonesia dan Malaysia. Periode data adalah harian dengan selang waktu satu minggu antara 19 April 2019 hingga 26 April 2019. Adapun untuk analisis maka sebanyak 144 data pada periode 19 April 2019 pukul 8 pagi hingga 25 April 2019 pukul 6 pagi digunakan sebagai data *in sample* dan sebanyak 24 data pada periode 25 April 2019 pukul 7 pagi hingga 26 April 2019 pukul 7 pagi digunakan sebagai *out sample*.

* 1. **Langkah-langkah Penelitian**

Adapun langkah-langkah penelitian yang dilakukan pada penelitian ini yaitu sebagai berikut.

1. Melakukan analisis statistika deskriptif
2. Analisis ARIMA

* Membagi data menjadi data *training* dan data *testing*.
* Membuat plot time series data
* Mengidentifikasi kestasioneran data dengan melihat ACF dan PACF
* Menebak model ARIMA
* Melakukan pengujian estimasi parameter model
* Melakukan pengujian asumsi residual yaitu asumsi *white noise* dengan menggunakan statistik uji *Ljung-Box* dan asumsi distribusi normal dengan menggunakan statistik uji *Kolmogorov Smirnov*

1. Analisis Regresi *Time Series*

* Membagi data menjadi data *training* dan data *testing*.
* Membuat variabel *dummy* jam
* Meregresikan variabel *dummy* dengan variabel y untuk data *training*
* Mendapatkan model regresi *time series* dari data *training*
* Mendapatkan hasil peramalan untuk data *training* dan *testing*
* Mengecek apakah residual sudah *white noise* atau belum
* Menghitung akurasi hasil *forecast* untuk data *training* dan data *testing*

1. Analisis *Neural Network*

* Menghitung akurasi hasil peramalan dengan *Neural Network*

1. Analisis *Deep Neural Network*

* Menghitung akurasi hasil peramalan dengan *Deep Neural Network*

1. Menentukan metode peramalan terbaik dengan membandingkan hasil akurasi dari keempat metode tersebut.
2. **Analisis dan Pembahasan**
   1. **Statistika Deskriptif**

Analisis statistika deskriptif dilakukan untuk mengetahui karakteristik dari data yang digunakan. Tabel 1 berikut berupakan hasil perhitungan *mean*, *variance*, maksimum dan minimum dari data Grab Indonesia, Grab Malaysia, dan Gojek Indonesia.

**Tabel 1** Statistika Deskriptif

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variabel | *Mean* | *Variance* | *Minimum* | *Maximum* |
| Grab: (Indonesia) | 73.23 | 247.53 | 38.00 | 100.00 |
| Gojek: (Indonesia) | 48.869 | 51.049 | 32.000 | 63.000 |
| Grab: (Malaysia) | 62.23 | 197.89 | 31.00 | 100.00 |

Berdasarkan Tabel 1 diatas, dapat dilihat bahwa rata-rata pencarian grab di Indonesia lebih tinggi daripada di Malaysia dan Gojekartinya, minat masyarakat Indonesia terhadap grab lebih tinggi dari pada gojek dan lebih tinggi daripada minat grab di Malaysia.

**3.2 Regresi *Time Series***

Regresi *time series* pada umumnya sama seperti regresi linier, yang membedakan yaitu pada regresi *time series* variabel dependen nya merupakan data deret waktu. Pada penelitian ini, variabel independen yang digunakan adalah variabel *dummy* jam.

1. **Regresi *Time Series* Untuk Data Gojek Indonesia**

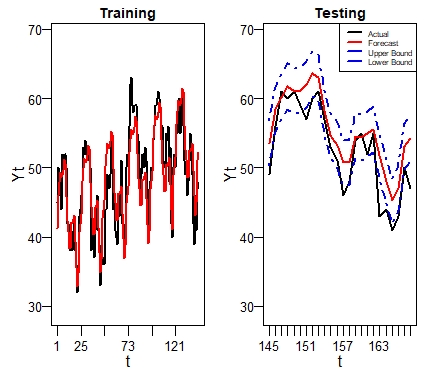
Dari hasil analisis regresi *time series* yang dilakukan menggunakan data *training*, diperoleh model regresi *time series* untuk pencarian kata Gojek di Indonesia sebagai berikut.







Untuk mengetahui apakah parameter yang didapat pada model tersebut signifikan, maka dilakukan pengujian signifikansi parameter secara parsial. Dari hasil pengujian yang dilakukan, di peroleh *p-value* pada masing-masing parameter yaitu 0,000 sehingga dari hasil pengujian yang telah dilakukan dapat disimpulkan bahwa parameter yang didapatkan signifikan terhadap model, artinya dalam satu hari banyak minat pencarian kata Gojek di Indonesia dipengaruhi oleh jam sehingga jumlah pencarian pada tiap jam berbeda dan cenderung membuat pola di jam yang sama pada hari berikutnya. Setelah mendapatkan model dari data *training*, model tersebut digunakan untuk *forecast* dengan data *testing*. Dari hasil *forecast* yang telah dilakukan dengan model regresi teresbut, diperoleh plot *time series* hasil *forecast*  seperti pada Gambar 1 berikut



**Gambar 1** *Time Series* Plot Hasil *Forecast*

Dan dari hasil *forecast* tersebut diperoleh hasil akurasi untuk data *training* dan *testing* pada Tabel 2 berikut,

**Tabel 2** Akurasi Hasil *Forecast* Data Gojek Indonesia

|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 3.432921 | 3.623343 |
| *MAE* | 2.789311 | 2.961806 |
| *MAPE* | 5.901459 | 6.046107 |

Setelah mendapatkan akurasi, langkah selanjutnya adalah mengecek apakah residual yang didapatkan telah memenuhi asumsi *white noise* dan berdistribusi normal atau tidak dengan melakukan uji *Ljung Box* dan *Kolmogorov-Smirnov*. Dari hasil pengujian *Ljung Box*, diperoleh *p-value* < 0,05 sehingga dapat disimpulkan yaitu residual yang diperoleh dari model regresi tidak memenuhi asumsi *white noise* sehingga residual nya masih bisa dimodelkan. Sedangkan dari hasil pengujian *Kolmogorov-Smirnov*, diperoleh *p-value* sebesar 0,6383. Karena *p-value* yang diperoleh > 0,05 maka kesimpulan dari hasil pengujian yaitu residual mengikuti distribusi normal.

1. **Regresi Time Series Untuk Data Grab Indonesia**

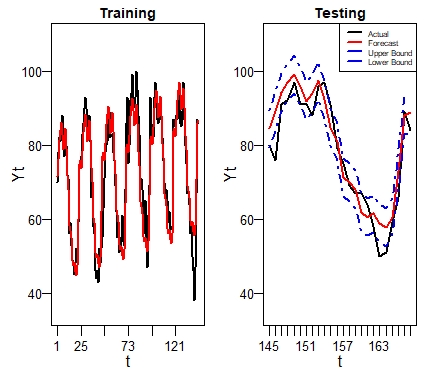
Dari hasil analisis regresi *time series* yang dilakukan menggunakan data *training*, diperoleh model regresi *time series* untuk pencarian kata Grab di Indonesia sebagai berikut.







Untuk mengetahui apakah parameter yang didapat pada model tersebut signifikan, maka dilakukan pengujian signifikansi parameter secara parsial. Dari hasil pengujian yang dilakukan, di peroleh *p-value* pada masing-masing parameter yaitu 0,000 sehingga dari hasil pengujian yang telah dilakukan dapat disimpulkan bahwa parameter yang didapatkan signifikan terhadap model artinya dalam satu hari banyak minat pencarian kata Grab di Indonesia dipengaruhi oleh jam sehingga jumlah pencarian pada tiap jam berbeda dan cenderung membuat pola di jam yang sama pada hari berikutnya. Setelah mendapatkan model dari data *training*, model tersebut digunakan untuk *forecast* dengan data *testing*. Dari hasil *forecast* yang telah dilakukan dengan model regresi teresbut, diperoleh plot *time series* hasil *forecast*  seperti pada Gambar 2 berikut



**Gambar 2** *Time Series* Plot Hasil *Forecast*

Dan dari hasil *forecast* tersebut diperoleh hasil akurasi untuk data *training* dan *testing* pada Tabel 3 berikut,

**Tabel 3** Akurasi Hasil *Forecast* Data Grab Indonesia

|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 5.334851 | 5.102407 |
| *MAE* | 4.035995 | 4.240278 |
| *MAPE* | 6.087288 | 5.865351 |

Setelah mendapatkan akurasi, langkah selanjutnya adalah mengecek apakah residual yang didapatkan telah memenuhi asumsi *white noise* dan berdistribusi normal atau tidak dengan melakukan uji *Ljung Box* dan *Kolmogorov-Smirnov*. Dari hasil pengujian *Ljung Box*, diperoleh *p-value* < 0,05. Sehingga dapat disimpulkan yaitu residual yang diperoleh dari model regresi tidak memenuhi asumsi *white noise* sehingga residual nya masih bisa dimodelkan. Sedangkan dari hasil pengujian *Kolmogorov-Smirnov*, diperoleh *p-value* sebesar 0,5759. Karena *p-value* yang diperoleh > 0,05 maka kesimpulan dari hasil pengujian yaitu residual mengikuti distribusi normal.

1. **Regresi Time Series Untuk Data Grab Malaysia**

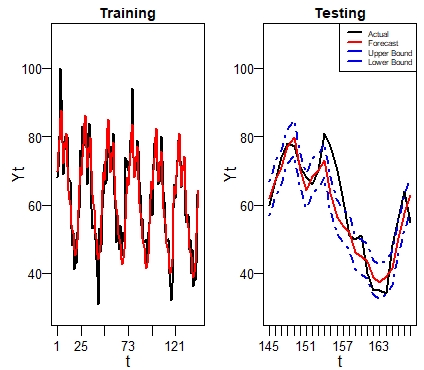
Dari hasil analisis regresi *time series* yang dilakukan menggunakan data *training*, diperoleh model regresi *time series* untuk pencarian kata Grab di Malaysia sebagai berikut.







Untuk mengetahui apakah parameter yang didapat pada model tersebut signifikan, maka dilakukan pengujian signifikansi parameter secara parsial. Dari hasil pengujian yang dilakukan, di peroleh *p-value* pada masing-masing parameter yaitu 0,000 sehingga dari hasil pengujian yang telah dilakukan dapat disimpulkan bahwa parameter yang didapatkan signifikan terhadap model artinya dalam satu hari banyak minat pencarian kata Grab di Malaysia dipengaruhi oleh jam sehingga jumlah pencarian pada tiap jam berbeda dan cenderung membuat pola di jam yang sama pada hari berikutnya. Setelah mendapatkan model dari data *training*, model tersebut digunakan untuk *forecast* dengan data *testing*. Dari hasil *forecast* yang telah dilakukan dengan model regresi teresbut, diperoleh plot *time series* hasil *forecast*  seperti pada Gambar 3 berikut



**Gambar 3** *Time Series* Plot Hasil *Forecast*

Dan dari hasil *forecast* tersebut diperoleh hasil akurasi untuk data *training* dan *testing* pada Tabel 4 berikut,

**Tabel 4** Akurasi Hasil *Forecast* Data Grab Malaysia

|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 5.260777 | 5.854737 |
| *MAE* | 4.182821 | 4.613194 |
| *MAPE* | 7.210051 | 8.028911 |

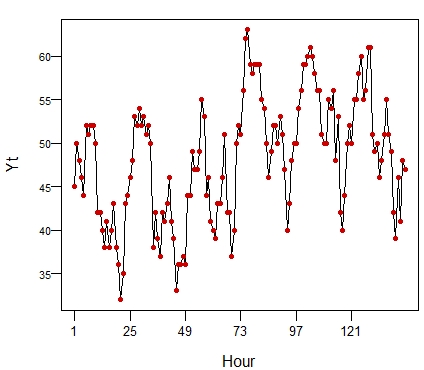
Setelah mendapatkan akurasi, langkah selanjutnya adalah mengecek apakah residual yang didapatkan telah memenuhi asumsi *white noise* dan berdistribusi normal atau tidak dengan melakukan uji *Ljung Box* dan *Kolmogorov-Smirnov*. Dari hasil pengujian *Ljung Box*, diperoleh *p-value* < 0,05. Sehingga dapat disimpulkan yaitu residual yang diperoleh dari model regresi tidak memenuhi asumsi *white noise* sehingga residual nya masih bisa dimodelkan. Sedangkan dari hasil pengujian *Kolmogorov-Smirnov*, diperoleh *p-value* sebesar 0,9744. Karena *p-value* yang diperoleh > 0,05 maka kesimpulan dari hasil pengujian yaitu residual mengikuti distribusi normal.

* 1. **ARIMA**

ARIMA adalah salah satu metode dalam *forecasting*. ARIMA dapat digunakan untuk meramalkan data yang stasioner, trend, maupun seasonal. Adapun orde dari ARIMA yaitu p, d, q untuk data non seasonal dan P, D, Q untuk data seasonal dimana p dan P adalah orde untuk AR, q dan Q adalah orde untuk MA, dan d dan D adalah order untuk I . Orde dari ARIMA dapat ditentukan dengan melihat plot ACF dan PACF dari data, namun data harus stasioner terlebih dahulu. Selain itu, terdapat asumsi-asumsi yang harus dipenuhi metode *forecast* dengan ARIMA yaitu residual telah *white noise* dan berdistribusi normal.

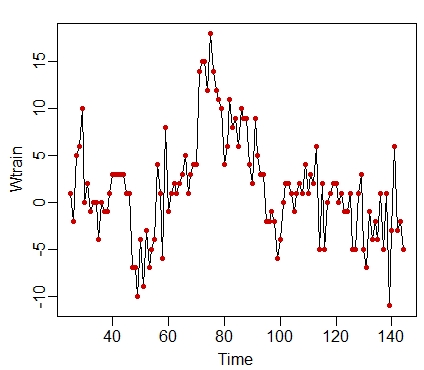
1. **ARIMA Data Gojek Indonesia**

Sebelum melakukan analisis lebih lanjut, tahapan yang perlu dilakukan adalah membuat plot data untuk mengetahui apakah data mengikuti pola yang trend, stasioner, atau seasonal. Gambar 4 berikut merupakan plot dari data Gojek Indonesia.



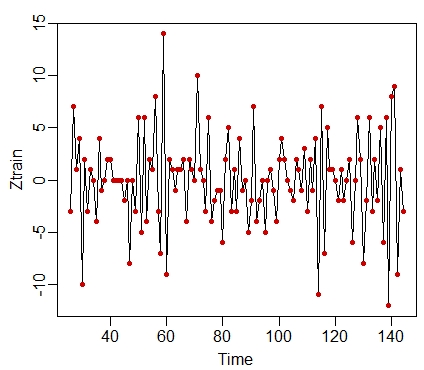
**Gambar 4** Plot *Time Series* Data Gojek Indonesia

Berdasarkan Gambar 1 diatas, dapat dilihat bahwa plot *time series* data Gojek Indonesia membentuk pola seasonal dengan periode musiman 24. Oleh karena itu, perlu dilakukan *differencing* dan diperoleh hasil pada Gambar 5 berikut.



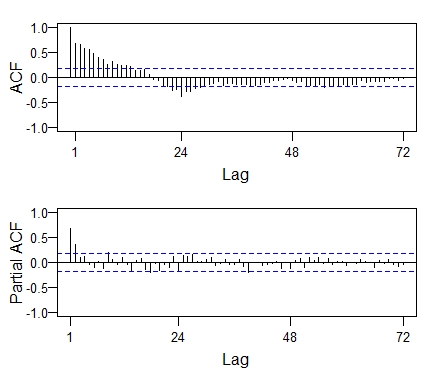
**Gambar 5** Plot *Time Series* Hasil *Differencing* Musiman 24

Bedasarkan Gambar 5 diatas, dapat dilihat bahwa plot *time series* hasil *differencing* tampak sudah stasioner, untuk memastikan hal tersebut, maka dilakukan pengujian *Augmented Dicky Fuller* untuk mengetahui apakah data hasil *differencing* musiman sudah stasioner atau belum. Dari hasil pengujian yang telah dilakukan, diperoleh *p-value* sebesar 0,644 . Karena *p-value* yang diperoleh dari hasil pengujian > 0,05, maka disimpulkan bahwa data belum stasioner sehingga perlu dilakukan *differencing*. Dari hasil *differencing*  lag 1 yang dilakukan, diperoleh hasil seperti pada Gambar 6 berikut.



**Gambar 6** Plot *Time Series* Hasil *Differencing* Lag 1

Bedasarkan Gambar 6 diatas, dapat dilihat bahwa plot *time series* hasil *differencing* sudah stasioner, namun untuk memastikan hal tersebut, maka dilakukan pengujian *Augmented Dicky Fuller* untuk mengetahui apakah data hasil *differencing* tersebut sudah stasioner atau belum. Dari hasil pengujian yang telah dilakukan, diperoleh *p-value* sebesar 0,01 . Karena *p-value* yang diperoleh dari hasil pengujian < 0,05, maka disimpulkan bahwa data sudah stasioner. Karena data sudah stasioner, untuk menebak orde ARIMA dari data, dibuat plot ACF dan PACF seperti pada Gambar 7 berikut.



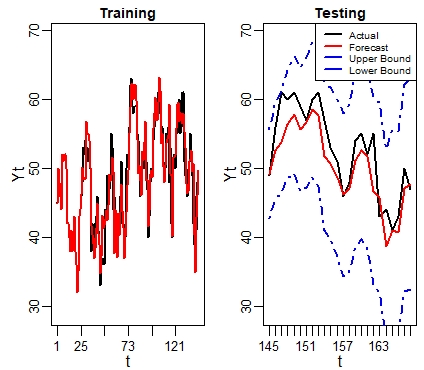
**Gambar 7** Plot ACF dan PACF Dari Data Stasioner

Dari hasil analisis yang dilakukan, didapatkan model ARIMA untuk data Gojek Indonesia adalah ARIMA (3,1,0)(1,1,0)24 dengan koefisien yang signifikan terhadap model seperti pada Tabel 5 berikut.

**Tabel 5** Koefisien Model ARIMA (,1,0)(1,1,0)24

|  |  |
| --- | --- |
| Kofisien | |
| AR1 | -0,59449 |
| AR2 | -0,37885 |
| AR3 | -0,22861 |
| SAR1 | -0,51331 |

Setelah mendapatkan model ARIMA untuk data Gojek Indonesia, tahapan selanjutnya adalah melakukan *forecast* dengan model yang didapatkan untuk data *testing*. Gambar 8 berikut merupakan hasil *forecast* yang dilakukan menggunakan data *training* dan *testing* Gojek Indonesia.



**Gambar 8** Hasil *Forecast* Data *Training* dan *Testing* Gojek Indonesia

Dari hasil *forecast* yang dilakukan, diperoleh akurasi untuk data *training* dan data *testing* seperti pada Tabel 6 berikut.

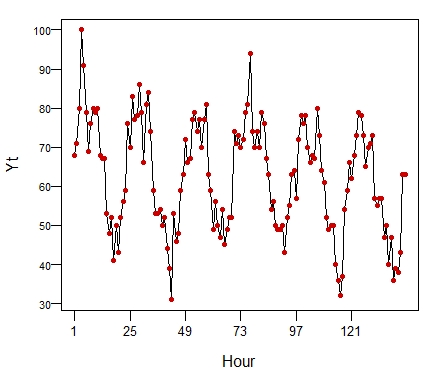
**Tabel 6** Akurasi Hasil *Forecast* Data Gojek Indonesia

|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 2.983570 | 3.455753 |
| *MAE* | 2.045135 | 2.710768 |
| *MAPE* | 4.232732 | 5.036889 |

Setelah melakukan forecast, tahap selanjutnya adalah melakukan pengujian untuk mengetahui apakah residual dari model ARIMA tersebut telah memenuhi asumsi *white noise* dan distribusi normal atau tidak. Dengan uji *LJung Box*, diperoleh *p-value* > 0,05 artinya residual telah memenuhi asumsi *white noise*. Untuk asumsi distribusi normal, dilakukan pengujian dengan *Kolmogorov- Smirnov* dan diperoleh *p-value* sebesar 0,09437. Karena *p-value* yang diperoleh > 0,05 maka dapat disimpulkan bahwa residual telah memenuhi asumsi distribusi normal.

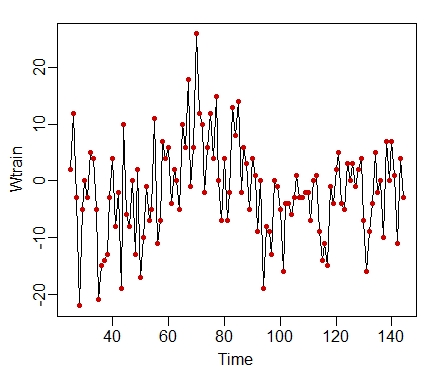
1. **ARIMA Data Grab Malaysia**

Sebelum melakukan analisis lebih lanjut, tahapan yang perlu dilakukan adalah membuat plot data untuk mengetahui apakah data mengikuti pola yang trend, stasioner, atau seasonal. Gambar 9 berikut merupakan plot dari data Grab Malaysia.



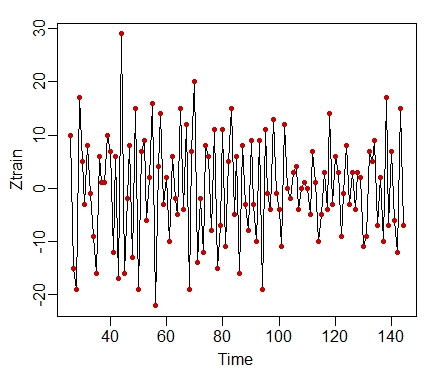
**Gambar 9** Plot *Time Series* Data Grab Malaysia

Berdasarkan Gambar 9 diatas, dapat dilihat bahwa plot *time series* data Grab Malaysia membentuk pola seasonal dengan periode musiman 24. Oleh karena itu, perlu dilakukan *differencing* dan diperoleh hasil pada Gambar 10 berikut.



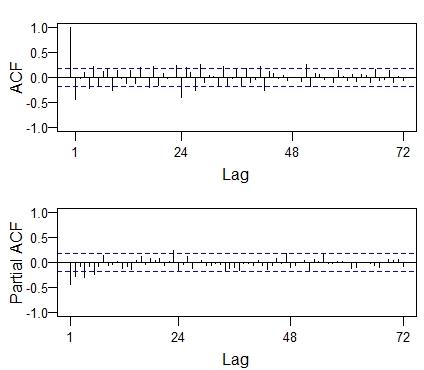
**Gambar 10** Plot *Time Series* Hasil *Differencing* Musiman 24

Bedasarkan Gambar 10 diatas, dapat dilihat bahwa plot *time series* hasil *differencing* tampak belum stasioner, untuk memastikan hal tersebut, maka dilakukan pengujian *Augmented Dicky Fuller* untuk mengetahui apakah data hasil *differencing* musiman sudah stasioner atau belum. Dari hasil pengujian yang telah dilakukan, diperoleh *p-value* sebesar 0,493 . Karena *p-value* yang diperoleh dari hasil pengujian > 0,05, maka disimpulkan bahwa data belum stasioner sehingga perlu dilakukan *differencing*. Dari hasil *differencing*  lag 1 yang dilakukan, diperoleh hasil seperti pada Gambar 11 berikut.



**Gambar 11** Plot *Time Series* Hasil *Differencing* Lag 1

Bedasarkan Gambar 11 diatas, dapat dilihat bahwa plot *time series* hasil *differencing* sudah stasioner, namun untuk memastikan hal tersebut, maka dilakukan pengujian *Augmented Dicky Fuller* untuk mengetahui apakah data hasil *differencing* tersebut sudah stasioner atau belum. Dari hasil pengujian yang telah dilakukan, diperoleh *p-value* sebesar 0,01 . Karena *p-value* yang diperoleh dari hasil pengujian < 0,05, maka disimpulkan bahwa data sudah stasioner. Karena data sudah stasioner, untuk menebak orde ARIMA dari data, dibuat plot ACF dan PACF seperti pada Gambar 12 berikut.



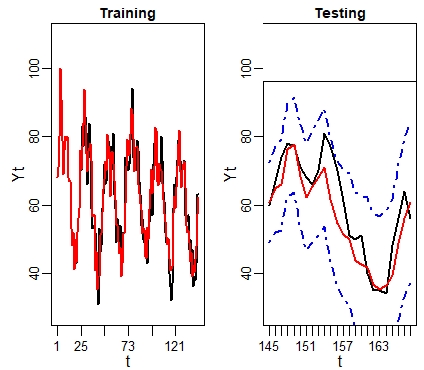
**Gambar 12** Plot ACF dan PACF Dari Data Stasioner

Dari hasil analisis yang dilakukan, didapatkan model ARIMA untuk data Grab Malaysia adalah ARIMA (4,1,[48])(1,1,0)24 dengan koefisien yang signifikan terhadap model seperti pada Tabel 7 berikut.

**Tabel 7** Koefisien Model ARIMA (4,1,[48])(1,1,0)24

|  |  |
| --- | --- |
| Kofisien | |
| AR1 | -0,601495 |
| AR2 | -0,419843 |
| AR3 | -0,359384 |
| AR4 | -0,980701 |
| MA48 | -0.996589 |
| SAR1 | -0.996589 |

Setelah mendapatkan model ARIMA untuk data Grab Malaysia, tahapan selanjutnya adalah melakukan *forecast* dengan model yang didapatkan untuk data *testing*. Gambar 13 berikut adalah hasil *forecast* data training dan testing Grab Malaysia.



**Gambar 13** Hasil *Forecast* Data *Training* dan *Testing* Grab Malaysia

Dari hasil *forecast* yang dilakukan, diperoleh akurasi untuk data *training* dan data *testing* seperti pada Tabel 8 berikut.

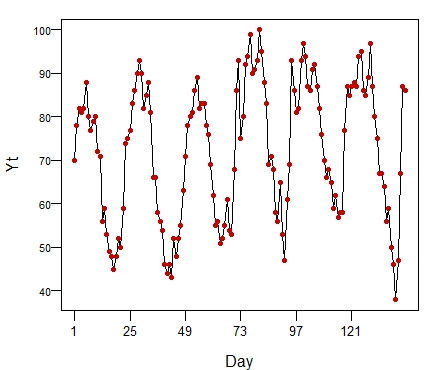
**Tabel 8** Akurasi Hasil *Forecast* Data Grab Indonesia

|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 5.020136 | 6.862155 |
| *MAE* | 3.710568 | 5.232673 |
| *MAPE* | 6.409760 | 8.537716 |

Setelah melakukan *forecast*, tahap selanjutnya adalah melakukan pengujian untuk mengetahui apakah residual dari model ARIMA tersebut telah memenuhi asumsi *white noise* dan distribusi normal atau tidak. Dengan uji *LJung Box*, diperoleh *p-value* > 0,05 artinya residual telah memenuhi asumsi *white noise*. Untuk asumsi distribusi normal, dilakukan pengujian dengan *Kolmogorov- Smirnov* dan diperoleh *p-value* sebesar 0,1245. Karena *p-value* yang diperoleh > 0,05 maka dapat disimpulkan bahwa residual telah memenuhi asumsi distribusi normal.

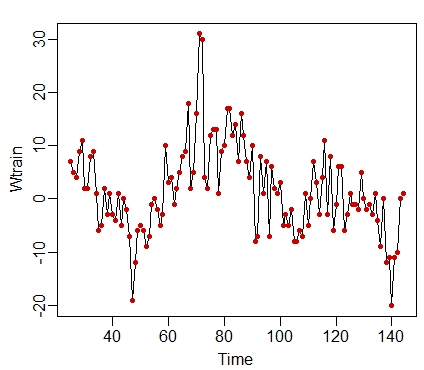
1. **ARIMA Data Grab Indonesia**

Sebelum melakukan analisis lebih lanjut, tahapan yang perlu dilakukan adalah membuat plot data untuk mengetahui apakah data mengikuti pola yang trend, stasioner, atau seasonal. Gambar 14 berikut merupakan plot dari data Grab Indonesia.



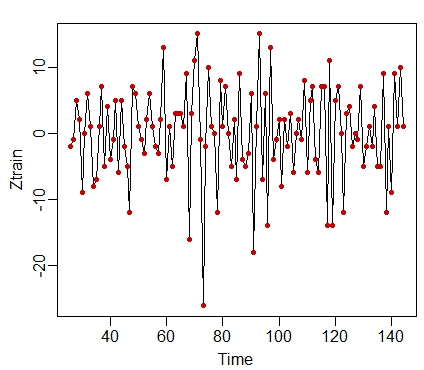
**Gambar 14** Plot *Time Series* Data Grab Indonesia

Berdasarkan Gambar 14 diatas, dapat dilihat bahwa plot *time series* data Grab Indonesia membentuk pola seasonal dengan periode musiman 24. Oleh karena itu, perlu dilakukan *differencing* dan diperoleh hasil pada Gambar 15 berikut.



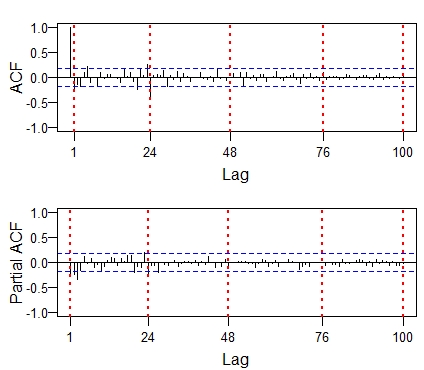
**Gambar 15** Plot *Time Series* Hasil *Differencing* Musiman 24

Bedasarkan Gambar 15 diatas, dapat dilihat bahwa plot *time series* hasil *differencing* tampak belum stasioner, untuk memastikan hal tersebut, maka dilakukan pengujian *Augmented Dicky Fuller* untuk mengetahui apakah data hasil *differencing* musiman sudah stasioner atau belum. Dari hasil pengujian yang telah dilakukan, diperoleh *p-value* sebesar 0,6932 . Karena *p-value* yang diperoleh dari hasil pengujian > 0,05, maka disimpulkan bahwa data belum stasioner sehingga perlu dilakukan *differencing*. Dari hasil *differencing*  lag 1 yang dilakukan, diperoleh hasil seperti pada Gambar 16 berikut.



**Gambar 16** Plot *Time Series* Hasil *Differencing* Lag 1

Bedasarkan Gambar 16 diatas, dapat dilihat bahwa plot *time series* hasil *differencing* sudah stasioner, namun untuk memastikan hal tersebut, maka dilakukan pengujian *Augmented Dicky Fuller* untuk mengetahui apakah data hasil *differencing* tersebut sudah stasioner atau belum. Dari hasil pengujian yang telah dilakukan, diperoleh *p-value* sebesar 0,01 . Karena *p-value* yang diperoleh dari hasil pengujian < 0,05, maka disimpulkan bahwa data sudah stasioner. Karena data sudah stasioner, untuk menebak orde ARIMA dari data, dibuat plot ACF dan PACF seperti pada Gambar 17 berikut.



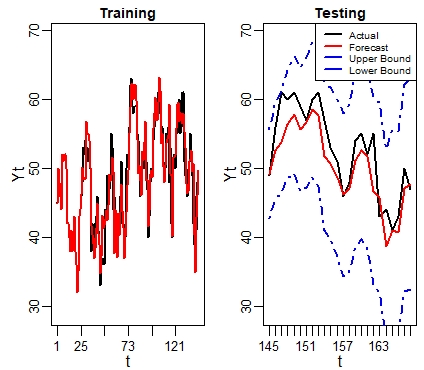
**Gambar 17** Plot ACF dan PACF Dari Data Stasioner

Dari hasil analisis yang dilakukan, didapatkan model ARIMA untuk data Grab Indonesia adalah ARIMA (4,1,[20])(1,1,0)24 dengan koefisien yang signifikan terhadap model seperti pada Tabel 9 berikut.

**Tabel 9** Koefisien Model ARIMA (4,1,[20])(1,1,0)24

|  |  |
| --- | --- |
| Kofisien | |
| AR1 | -0,534881 |
| AR2 | -0,467859 |
| AR3 | -0,442782 |
| AR4 | -0,232806 |
| MA20 | -0.304401 |
| SAR1 | -0.472780 |

Setelah mendapatkan model ARIMA untuk data Grab Indonesia, tahapan selanjutnya adalah melakukan *forecast* dengan model yang didapatkan untuk data *testing*. Gambar 18 berikut adalah hasil *forecast* data training dan testing Grab Indonesia.



**Gambar 18** Hasil *Forecast* Data *Training* dan *Testing* Grab Indonesia

Dari hasil *forecast* yang dilakukan, diperoleh akurasi untuk data *training* dan data *testing* seperti pada Tabel 10 berikut.

**Tabel 10** Akurasi Hasil *Forecast* Data Grab Indonesia

|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 4.504295 | 6.874076 |
| *MAE* | 3.177660 | 5.841041 |
| *MAPE* | 4.608001 | 7.635773 |

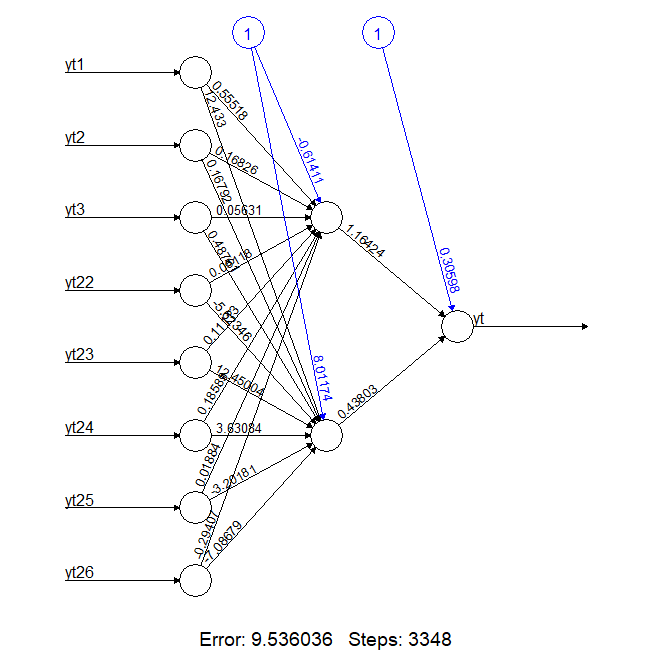
Setelah melakukan *forecast*, tahap selanjutnya adalah melakukan pengujian untuk mengetahui apakah residual dari model ARIMA tersebut telah memenuhi asumsi *white noise* dan distribusi normal atau tidak. Dengan uji *LJung Box*, diperoleh *p-value* > 0,05 artinya residual telah memenuhi asumsi *white noise*. Untuk asumsi distribusi normal, dilakukan pengujian dengan *Kolmogorov- Smirnov* dan diperoleh *p-value* sebesar 0,01249. Karena *p-value* yang diperoleh < 0,05 maka dapat disimpulkan bahwa residual belum memenuhi asumsi distribusi normal.

* 1. ***Neural Network***

Pada penelitian ini, akan dilakukan *forecast* dengan *Neural Network* dilakukan sebanyak 10 kali dan dipilih *Neural Network* dengan hasil akurasi yang memberikan hasil terbaik baik untuk data *training* maupun data *testing*.

1. ***Neural Network* Data Gojek Indonesia**

Dari hasil *forecast* dengan *Neural Network*, dipilih *Neural Network* kedua dengan struktur seperti pada Gambar 19 berikut.



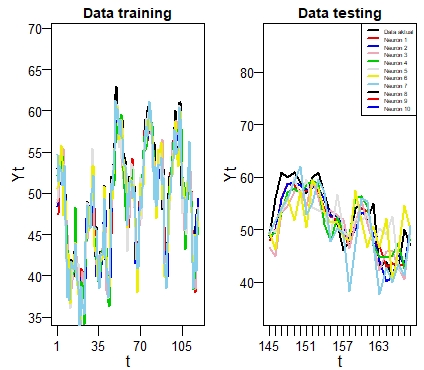
**Gambar 19** Bagan *Neural Network* untuk Data Gojek Indonesia

*Neural Network* kedua dipilh karena memberikan hasil akurasi yang terbaik dimana hasil akurasi yang diperoleh untuk data *training* dan *testing* seperti pada Tabel 11 berikut.

**Tabel 11** Ukuran Akurasi Hasil *Forecast* data Gojek Indonesia

|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 2.940941 | 3.210672 |
| *MAE* | 2.252531 | 2.490530 |
| *MAPE* | 4.679973 | 4.758335 |

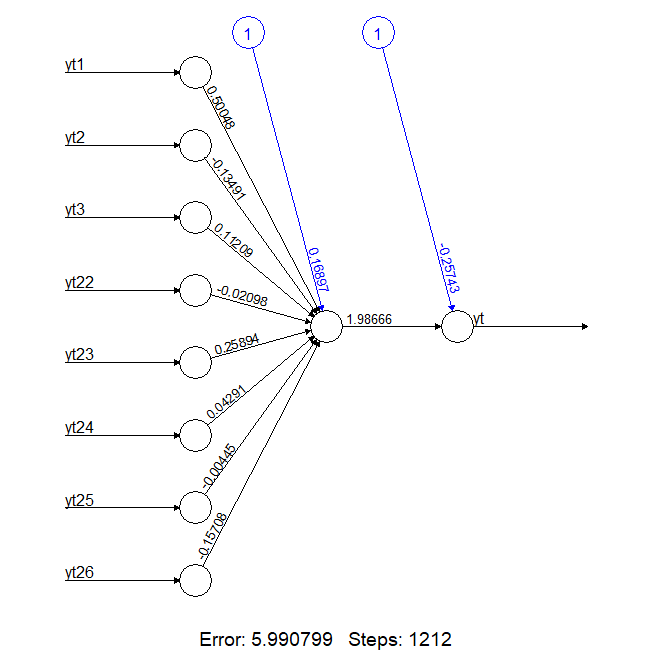
Dari Hasil *forecast* dengan *Neural Network*, diperoleh time series plot untuk data training dan data testing seperti pada Gambar 20 berikut.



**Gambar 20** *Time Series* Plot Hasil *Forecast*

1. ***Neural Network* Data Grab Indonesia**

Dari hasil *forecast* dengan *Neural Network*, dipilih *Neural Network* pertama dengan struktur seperti pada Gambar 21 berikut.



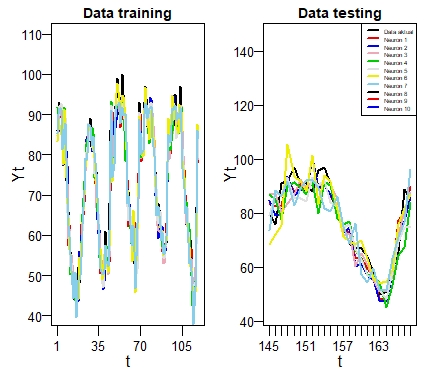
**Gambar 21** Bagan *Neural Network* untuk Data Grab Indonesia

*Neural Network* pertama dipilh karena memberikan hasil akurasi yang terbaik dimana hasil akurasi yang diperoleh untuk data *training* dan *testing* seperti pada Tabel 12 berikut.

**Tabel 12** Ukuran Akurasi Hasil *Forecast* data Grab Indonesia

|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 5.072834 | 5.277946 |
| *MAE* | 3.926651 | 4.441255 |
| *MAPE* | 5.720283 | 5.670660 |

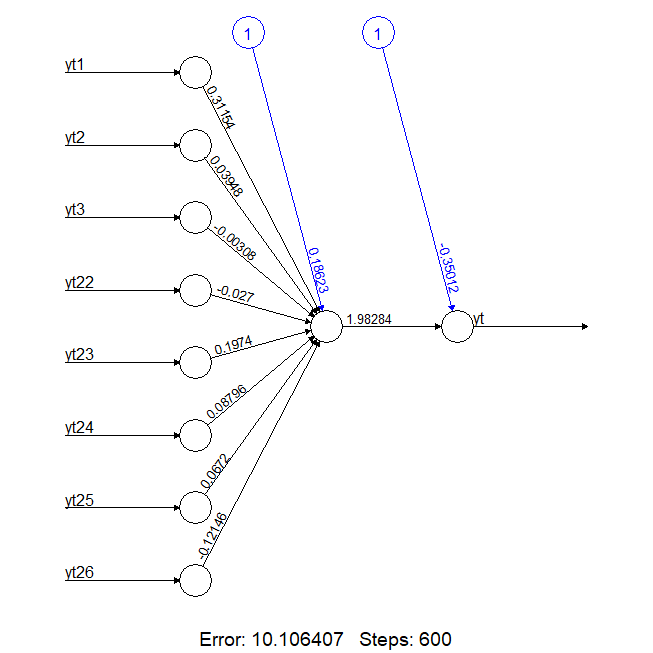
Dari Hasil *forecast* dengan *Neural Network*, diperoleh time series plot untuk data training dan data testing seperti pada Gambar 22 berikut.



**Gambar 22** *Time Series* Plot Hasil *Forecast*

1. ***Neural Network* Data Grab Malaysia**

Dari hasil *forecast* dengan *Neural Network*, dipilih *Neural Network* pertama dengan struktur seperti pada Gambar 23 berikut.



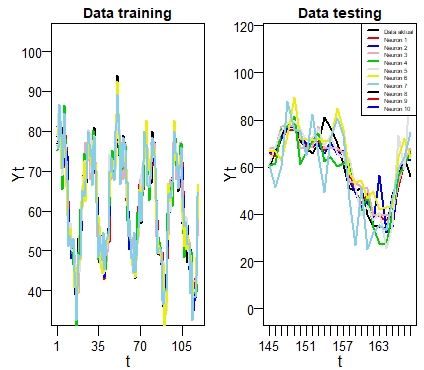
**Gambar 23** Bagan *Neural Network* untuk Data Grab Malaysia

*Neural Network* pertama dipilh karena memberikan hasil akurasi yang terbaik dimana hasil akurasi yang diperoleh untuk data *training* dan *testing* seperti pada Tabel 13 berikut.

**Tabel 13** Ukuran Akurasi Hasil *Forecast* data Grab Malaysia

|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 5.834280 | 5.728950 |
| *MAE* | 4.580732 | 4.576286 |
| *MAPE* | 7.952934 | 8.425166 |

Dari Hasil *forecast* dengan *Neural Network*, diperoleh time series plot untuk data training dan data testing seperti pada Gambar 24 berikut.



**Gambar 24** *Time Series* Plot Hasil *Forecast*

* 1. ***Deep Learning Neural Network ( Deep NN )***

Pada penelitian kali ini, *forecast* dengan *Deep NN* akan dilakukan sebanyak 10 kali dan dipilih *Deep NN* yang memberikan hasil akurasi yang terbaik untuk data *training* dan *testing*.

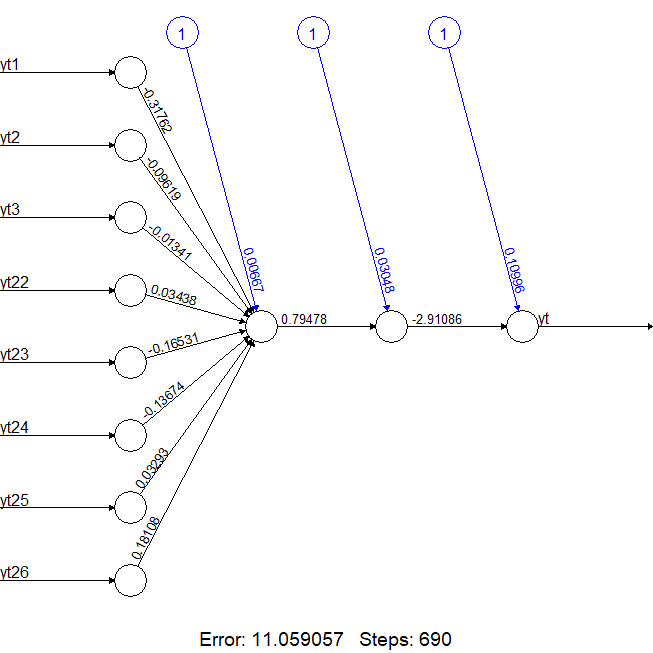
1. ***Deep NN* Data Gojek Indonesia**

Dari hasil *forecast* dengan *Deep NN* yang dilakukan sebanyak 10 kali, *Deep NN* yang memberikan hasil akurasi terbaik untuk data training dan testing yaitu *Deep NN* yang pertama dengan hasil akurasi seperti pada Tabel 14 berikut.

**Tabel 14** Ukuran Hasil *Forecast* Data Gojek Indonesia

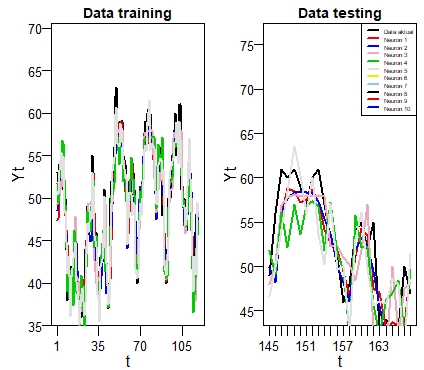
|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 3.156953 | 3.255612 |
| *MAE* | 2.485650 | 2.586802 |
| *MAPE* | 5.179572 | 4.943229 |

Adapun struktur *Deep NN* nya dapat dilihat oada Gambar 25 berikut.



**Gambar 25** *Deep NN* Data Gojek Indonesia

Data Gojek Indonesia setelah di *forecast* dengan Deep NN diperoleh hasil plot *time series* seperti pada Gambar 26 berikut.



**Gambar 26** *Time Series* Plot Hasil *Forecast* dengan *Deep NN*

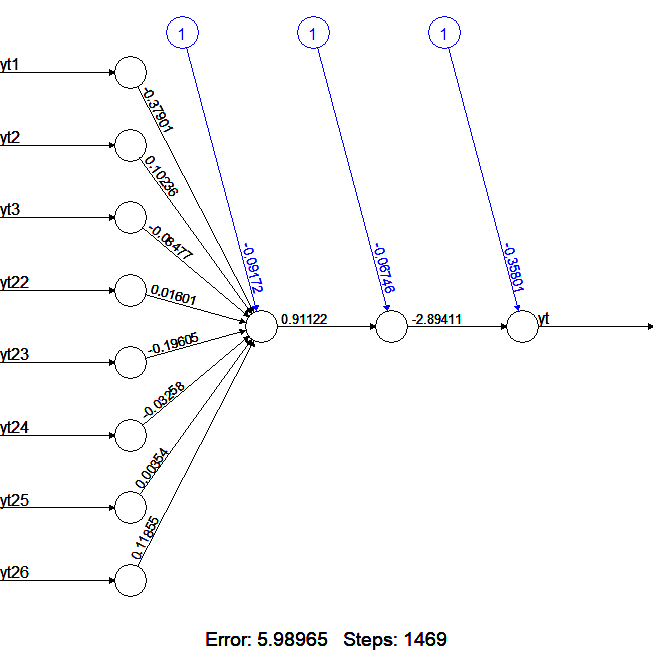
1. ***Deep NN* Data Grab Indonesia**

Dari hasil *forecast* dengan *Deep NN* yang dilakukan sebanyak 10 kali, *Deep NN* yang memberikan hasil akurasi terbaik untuk data training dan testing yaitu *Deep NN* yang pertama dengan hasil akurasi seperti pada Tabel 15 berikut.

**Tabel 15** Ukuran Hasil *Forecast* Data Grab Indonesia

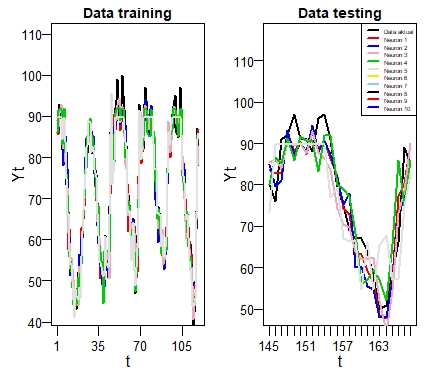
|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 5.072325 | 5.281075 |
| *MAE* | 3.925764 | 4.446597 |
| *MAPE* | 5.719272 | 5.676841 |

Adapun struktur *Deep NN* nya dapat dilihat oada Gambar 27 berikut.



**Gambar 27** *Deep NN* Data Grab Indonesia

Data Grab Indonesia setelah di *forecast* dengan Deep NN diperoleh hasil plot *time series* seperti pada Gambar 28 berikut.



**Gambar 28** *Time Series* Plot Hasil *Forecast* dengan *Deep NN*

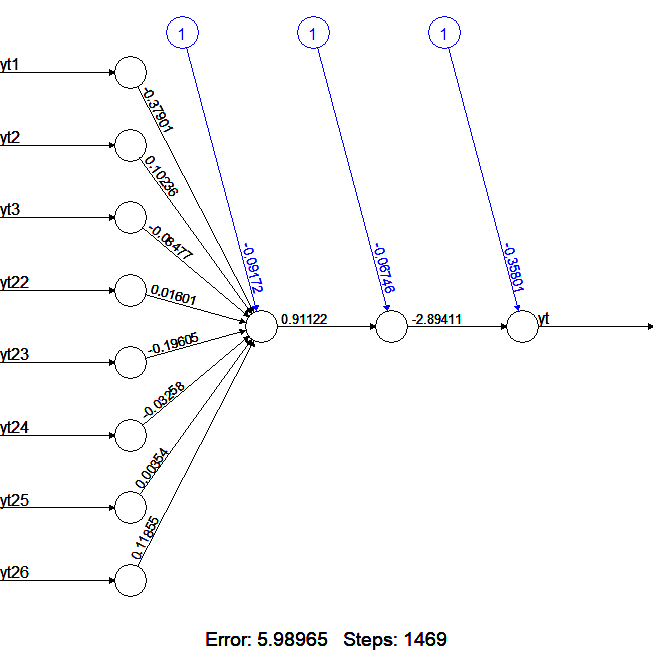
1. ***Deep NN* Data Grab Malaysia**

Dari hasil *forecast* dengan *Deep NN* yang dilakukan sebanyak 10 kali, *Deep NN* yang memberikan hasil akurasi terbaik untuk data training dan testing yaitu *Deep NN* yang pertama dengan hasil akurasi seperti pada Tabel 16 berikut.

**Tabel 16** Ukuran Hasil *Forecast* Data Grab Malaysia

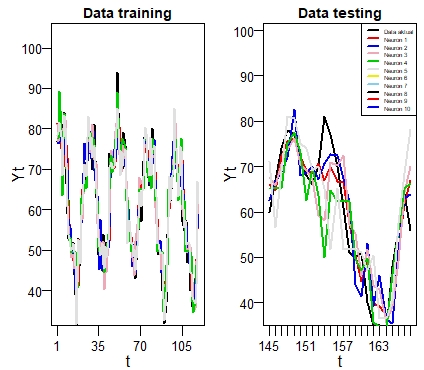
|  |  |  |
| --- | --- | --- |
| Ukuran | *Training* | *Testing* |
| *RMSE* | 5.836449 | 5.729419 |
| *MAE* | 4.582959 | 4.576906 |
| *MAPE* | 7.956233 | 8.420049 |

Adapun struktur *Deep NN* nya dapat dilihat pada Gambar 29 berikut.



**Gambar 29** *Deep NN* Data Grab Malaysia

Data Grab Malaysia setelah di *forecast* dengan Deep NN diperoleh hasil plot *time series* seperti pada Gambar 30 berikut



**Gambar 30** *Time Series* Plot Hasil *Forecast* dengan *Deep NN*

* 1. **Perbandingan Akurasi Keempat Metode**

Setelah mendapatkan akurasi dari keempat metode tersebut, tahap selanjutnya adalah membandingkan akurasi dari keempat metode untuk mengetahui metode yang baik digunakan untuk masing-masing data.

1. **Perbandingan Metode Untuk Data Gojek Indonesia**

Dari hasil *forecasting* menggunakan metode *Neural Network, ARIMA, Deep Neural Network*, dan Regresi *Time Series*, diperoleh perbandingan hasil akurasi untuk RMSE, MAE, dan MAPE untuk data *training* dan *testing* seperti pada Gambar 31 berikut.

|  |  |
| --- | --- |
|  |  |
|  | |

**Gambar 31** Perbandingan Metode Data Gojek Indonesia

Berdasarkan Gambar 31 diatas, dapat dilihat bahwa untuk data Gojek Indonesia, metode yang menghasilkan akurasi terbaik adalah *Neural Network* . Dari Gambar 31 diatas, diketahui bahwa *Neural Network* memberikan nilai eror terkecil untuk data *training* dan data *testing*.

1. **Perbandingan Metode Untuk Data Grab Indonesia**

Dari hasil *forecasting* menggunakan metode *Neural Network, ARIMA, Deep Neural Network*, dan Regresi *Time Series*, diperoleh perbandingan hasil akurasi untuk RMSE, MAE, dan MAPE untuk data *training* dan *testing* seperti pada Gambar 32 berikut.

|  |  |
| --- | --- |
|  |  |
|  | |

**Gambar 32** Perbandingan Metode Data Grab Indonesia

Berdasarkan Gambar 32 diatas, dapat dilihat bahwa untuk data Grab Indonesia, metode yang menghasilkan akurasi terbaik adalah regresi *time series* . Dari Gambar 32 diatas, diketahui bahwa *Neural Network* memberikan nilai eror terkecil untuk data *training* dan data *testing*.

1. **Perbandingan Metode Untuk Data Grab Malaysia**

Dari hasil *forecasting* menggunakan metode *Neural Network, ARIMA, Deep Neural Network*, dan Regresi *Time Series*, diperoleh perbandingan hasil akurasi untuk RMSE, MAE, dan MAPE untuk data *training* dan *testing* seperti pada Gambar 33 berikut.

|  |  |
| --- | --- |
|  |  |
|  | |

**Gambar 14** Perbandingan Metode Data Grab Malaysia

Berdasarkan Gambar 33 diatas, dapat dilihat bahwa untuk data Grab Malaysia, metode yang menghasilkan akurasi terbaik adalah regresi *Time series* . Dari Gambar 33 diatas, diketahui bahwa *Neural Network* memberikan nilai eror terkecil untuk data *training* dan data *testing*.

**Kesimpulan**

Dari hasil analisis yang telah dilakukan, diperoleh kesmpulan yaitu untuk data Gojek Indonesia, metode yang memberikan hasil akurasi *forecasting* terbaik adalah metode *Neural Network*, untuk data Grab Indonesia meode yang memberikan hasil akurasi *forecasting* terbaik adalah metode regresi *time series*. Sedangkan untuk data Grab Malaysia metode yang memberikan hasil akurasi terbaik adalah metode regresi *time series.*

**Saran**

Saran yang dapat diberikan untuk penelitian selanjutnya yaitu sebaiknya error yang masih belum memenuhi asumsi *white noise* pada regresi time series diatasi dengan memodelkan residual tersebut dan untuk residual yang tidak memenuhi asumsi distribusi normal, sebaiknya dilakukan deteksi *outlier*.

**Ucapan Terimakasih**

Terimakasih banyak kepada dosen dan asisten dosen mata kuliah Analisis Data karena telah membimbing dalam pengolahan data sehingga memudahkan dalam mengerjakan laporan ini. Tak lupa saya ucapkan terimakasih untuk teman-teman yang telah membantu ketika terdapat kesulitan.