

Peramalan Jumlah Penumpang Kereta Api di Jabodetabek Menggunakan Metode *Hybrid Autoregressive Integrated Moving Average-Radial Basis Function Neural Network (ARIMA-RBFNN)*

Oleh: Sekar Jati Cahyaning Wulan¹, Tiani Wahyu Utami², M.AI Haris³

¹²³Program Studi Statistika, Univeristas Muhammadiyah Semarang

e-mail: sekarjati29@gmail.com

Article history	Abstract
Submission :	The development of rail transportation continues to be carried out to reduce congestion in the Jabodetabek area. The growth in the number of train passengers which fluctuates every month needs to be taken into account by policy makers, namely PT. KAI and the government, namely by doing forecasting. In this study, the Hybrid ARIMA-RBFNN method will be used because the data on the number of train passengers in Jabodetabek apart from a linear pattern, it is suspected that there is a nonlinear pattern. The combination of these methods is expected to produce a better level of accuracy than the single method, besides that there are assumptions that say that a single method cannot totally identify the characteristics of the time series. From the research, it is known that in the ARIMA (1,1,0) (1,1,1) 12 model the Mean Square Percentage Error (MAPE) value is 182.3232, while the ARIMA-RBFNN Hybrid model has a MAPE value of 182.2542684. Thus, the best model to use is the Hybrid ARIMA-RBFNN model. The best model is a combination of the ARIMA (1,1,0) (1,1,1) 12 model and its residuals are reused to be modeled using the Radial Basis Function Neural Network method. Modeling using RBFNN obtained the best network architecture, namely 6-6-1 (6 input neurons, 6 hidden neurons and 1 output neurons).
Revised :	
Accepted :	
Keyword:	
Peramalan, Hybrid ARIMA-RBFNN, Jumlah Penumpang Kereta Api	

PENDAHULUAN

Salah satu sarana penting yang digunakan sebagai penunjang kegiatan perekonomian adalah transportasi. Transportasi merupakan alat yang digunakan untuk berpindahnya manusia atau barang dari suatu tempat menuju ketempat lain. Sebagai upaya pemerataan ekonomi pemerintah melakukan upaya pembangunan sarana transportasi di seluruh wilayah baik pedesaan maupun perkotaan. Selain pembangunan jalan tol pemerintah juga membangun perkembangan transportasi berbasis rel yaitu kereta api.

Kereta api adalah salah satu sarana transportasi darat yang diminati dan populer di kalangan masyarakat, khususnya di wilayah Jabodetabek. Transportasi kereta api dipilih masyarakat di wilayah Jabodetabek sebagai alternative untuk menghindari kemacetan. Oleh karena itu

pengembangan transportasi terus dilakukan di wilayah Jabodetabek untuk mengatasi kemacetan sehingga diharapkan mampu meningkatkan produktivitas masyarakatnya.

PT Kereta Api Indonesia (PT. KAI) merupakan Badan Usaha Milik Negara Indonesia yang menawarkan jasa angkutan kereta api ke beberapa wilayah di Indonesia. PT. KA Indonesia memiliki tanggung jawab untuk memenuhi kebutuhan fasilitas yang memadai untuk masyarakat pengguna transportasi kereta api.

Pertumbuhan Jumlah penumpang kereta api di Jabodetabek dari tahun 2014 hingga 2020 mengalami fluktuasi. Dengan fluktuasi jumlah penumpang kereta api setiap tahunnya yang tidak bisa ditebak, diperlukannya suatu peramalan untuk memprediksi jumlah penumpang pada periode kedepannya. Peramalan jumlah penumpang kereta api di

Jabodetabek diharapkan memberikan informasi bagi PT KAI sehingga dapat meningkatkan kualitas pelayanan publik dan kapasitas layanan publik. Seiring dengan perkembangan metode peramalan *time series*, terdapat permasalahan yang ditemui pada data time series adalah jarang yang linier atau non linier saja, namun seringkali mengandung keduanya. Permasalahan lain yang sering ditemukan adalah kesulitan dalam menerapkan penggunaan model linier atau model nonlinier pada permasalahan data time series. Selain itu, dalam beberapa literatur peramalan menyatakan bahwa tidak ada model tunggal yang terbaik dalam pada setiap situasi (Zhang, 2003 dalam Al Hikmah, 2017). Banyak studi literatur yang menyarankan bahwa dengan menggabungkan beberapa model yang berbeda dapat meningkatkan akurasi peramalan dibandingkan hanya menggunakan model tunggal. Penelitian menggunakan *Hybrid ARIMA-ANN* pernah dilakukan oleh Virginia Sari et al. (2020) tentang prediksi kecepatan angin dalam mendeteksi gelombang air laut terhadap skala beaufort. Hasil penelitian menunjukkan bahwa model *Hybrid ARIMA-ANN* memiliki nilai MSE *training* dan *testing* terkecil. Penelitian selanjutnya tentang dilakukan oleh DT Wiyanti, et al. (2012) menggabungkan model Radial Basis Function dan Autoregressive Integrated Moving Average pada peramalan data Indeks Harga Perdagangan Besar dan Inflasi Komoditi Indonesia. Hasil penelitian menunjukkan bahwa dengan metode penggabungan ARIMA dan RBF memberikan hasil yang lebih akurat dibandingkan dengan penggunaan salah satu metode saja.

Peramalan yang cocok dengan data jumlah penumpang kereta api di Jabodetabek yaitu menggunakan metode ARIMA. Metode ARIMA ini merupakan pengolahan data secara linier sedangkan RBFNN saja tidak dapat menangkap pola linier dan nonlinier sekaligus. Sehingga penelitian ini digunakan model *Hybrid Autoregressive Integrated Moving Average-Radial Basis Function Neural Network* (*Hybrid ARIMA-RBFNN*) untuk meramalkan jumlah penumpang kereta api di Jabodetabek. Penggabungan model ini diharapkan menghasilkan ramalan yang semakin baik atau *error* yang dihasilkan semakin kecil.

LANDASAN TEORI

Model *Autoregressive Integrated Moving Average* (ARIMA)

Model ARIMA merupakan model gabungan antara *Autoregressive* (AR) orde p dan *Moving Average* (MA) orde q serta proses *differencing* orde d . Model ARIMA meliputi model ARIMA non-musiman dan musiman. Model AR pada orde p merupakan pengamatan pada waktu ke- t berhubungan linier dengan pengamatan masa lampau $X_{t-1}, X_{t-2}, \dots, X_{t-p}$. Bentuk persamaan model AR pada orde p adalah sebagai berikut:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \alpha \quad (1)$$

Model MA digunakan untuk menjelaskan suatu pengamatan pada waktu t yang dinyatakan sebagai kombinasi linier dari sejumlah nilai *residual*. Bentuk persamaan model MA dengan orde q adalah sebagai berikut:

$$X_t = \alpha_t - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} - \dots - \theta_q \alpha_{t-q} \quad (2)$$

Model ARMA merupakan gabungan antara model AR dan MA yang dinotasikan *ARMA*(p, q). Bentuk persamaan model ARMA pada orde p dan q adalah sebagai berikut:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \alpha - \theta_1 \alpha_{t-1} - \dots - \theta_q \alpha_{t-q} \quad (3)$$

Model *ARIMA* (p, d, q) digunakan untuk data *time series* yang stasioner setelah dilakukan *differencing* sebanyak d kali. Bentuk persamaan model ARIMA adalah sebagai berikut:

$$\phi_p(B)(1-B)^d X_t = \theta_0 + \theta_q(B)\alpha_t \quad (4)$$

Pengembangan dari model ARIMA yang mempunyai pola musiman (*seasonal*) maka persamaan model yang terbentuk adalah sebagai berikut:

$$\phi_p(B^s) \phi_p(B)(1-B)^d(1-B^s)^d X_t = \theta_q(B)\theta_q(B^s)\alpha_t \quad (5)$$

Artificial Neural Network

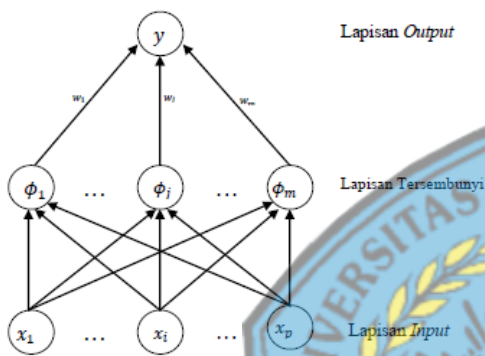
Artificial Neural Network atau yang lebih dikenal dengan sebutan *Neural Network* (NN) yang dalam bahasa Indonesia disebut Jaringan Saraf Tiruan (JST) ialah sistem pengelolaan informasi yang karakteristik menyerupai jaringan saraf biologis (Fauset, 1994:3 dalam Novelia Purba, 2016). Dalam pengoperasian NN terdapat dasar unit pengelolaan informasi yang disebut dengan *neuron*.

Beberapa jenis algoritma pembelajaran dalam *Artificial Neural Network* yang dikembangkan untuk peramalan data deret waktu antara lain yaitu *Backpropagation*, *recurrent network*, *self*

organizing map, radial basis function neural network dan sebagainya.

Radial Basis Function Neural Network

Radial Basis Function neural network (RBFNN) merupakan salah satu metode dari *Neural Network* dengan metode pelatihan hibrida yaitu menggabungkan metode pelatihan terbimbing dan metode pelatihan tak terbimbing. RBFNN memiliki 3 lapisan layer yang meliputi *input layer*, *hidden layer* dan *output layer*. Pada *hidden layer* digunakan fungsi aktivasi yaitu fungsi *Gaussian*. Arsitektur jaringan RBFNN adalah sebagai berikut:



Gambar 1 Arsitektur Jaringan RBFNN

Berdasarkan arsitektur diatas, berikut merupakan algoritma pelatihan Radial Basis function Neural Network (Wulandari, Ayu, 2017):

1. Menentukan banyak input jaringan
2. Menentukan banyak lapisan tersembunyi
3. Menentukan hasil aktivasi jaringan RBFNN dengan menggunakan fungsi Gaussian pada persamaan berikut ini:

$$\phi_j(x) = \exp\left(-\sum_{i=1}^k \left(\frac{|x_i - c_j|^2}{2\sigma_j^2}\right)\right) \quad (6)$$

4. Menghitung bobot pelatihan menggunakan metode least square
5. Menghitung nilai keseluruhan output jaringan RBFNN berdasarkan persamaan $f(x) = \sum_{j=1}^m w_j \phi_j(x)$ (7)

K-Means Clustering

Salah satu ciri model RBFNN ialah terletak pada fungsi aktivasi yang perhitungannya membutuhkan nilai pusat dan varians *neuron* tersembunyi. Metode *K-Means* ini

mengelompokkan data *input* menjadi beberapa kelompok atau *cluster* untuk menghitung nilai pusat dan varians. Pusat cluster adalah rata-rata (*mean*) *cluster* tersebut. Algoritma *K-Means Clustering* adalah sebagai berikut :

1. Tentukan nilai k kluster dan nilai k nilai pusat
2. Tempatkan setiap data/objek ke *cluster* terdekat. Kedekatan dua objek berdasarkan jarak kedua objek tersebut. Jarak dihitung dengan menggunakan jarak *Euclidean*. Persamaan jarak *Euclidean* antara dua titik sebarang P dan Q dengan P (x_1, x_2, \dots, x_p) dan Q (y_1, y_2, \dots, y_p) adalah sebagai berikut:

$$d(P, q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (8)$$

Hitung ulang nilai pusat *cluster* yang menerima data baru dan *cluster* yang kehilangan data.

3. Ulangi langkah ke-2 sampai nilai pusat lama sama dengan nilai pusat baru (stabil)

Metode Kuadrat Terkecil (Least Square)

Desain dari RBFNN akan membentuk pemetaan secara nonlinier dari variabel *input* ke lapisan *hidden* dan pemetaan linier dari lapisan *hidden* ke lapisan *output*. Oleh karena itu, model RBFNN memerlukan optimasi pada lapisan *output* yang dilakukan dengan menggunakan metode kuadrat kecil (*least square*). Metode kuadrat terkecil pada penelitian ini akan digunakan untuk menentukan nilai bobot dengan nilai *error* yang minimum. Model linier yang digunakan yaitu $y = \sum_{j=1}^m w_j \phi_j(x)$ dan *training set* $\{(x_i, \hat{y}_i)\}_{i=1}^n$ maka prinsip dari kuadrat terkecil yaitu meminimalkan jumlah kuadrat kesalahan (*Sum Square Error*) dengan rumus sebagai berikut :

$$SSE = \sum_{i=1}^n (\hat{y}_i - y)^2 \quad (9)$$

Metode Global Ridge-Regression

Metode *global ridge regression* mengestimasi bobot dengan cara menambahkan nilai parameter regulasi yang bernilai positif pada SSE sehingga diperoleh fungsi sebagai berikut (Orr,1996:21 dalam Fajarani Juliaristi,2014):

$$C = \sum_{i=1}^n (y - \hat{y}_i)^2 + \lambda \sum_{j=1}^m w_j^2 \quad (10)$$

Sehingga diperoleh nilai bobot dengan persamaan sebagai berikut:

$$\hat{w} = (\Phi^T \Phi + \lambda I_m)^{-1} \Phi^T \hat{y} \quad (11)$$

Model Hybrid ARIMA-RBFNN

Hybrid adalah kombinasi dari satu atau lebih metode atau model dalam satu sistem. Dalam hal ini adalah kombinasi antara ARIMA dan RBFNN. Penggunaan model *hybrid* diharapkan akan mendapatkan model yang lebih akurat. Di karenakan, pada dasarnya dalam dunia nyata sulit ditemukan kejadian deret waktu yang murni linier ataupun murni non linear. Secara umum, rumus kombinasi deret waktu yaitu sebagai berikut (Zhang,2003):

$$Y_t = \hat{N}_t + \hat{L}_t \quad (12)$$

Dimana \hat{N}_t merupakan komponen non linier dimana model yang digunakan adalah RBFNN sedangkan \hat{L}_t adalah komponen linier dimana model yang digunakan adalah model ARIMA.

Uji Linieritas Data

Uji linieritas dapat digunakan dengan menggunakan Uji Terasvirta. Uji Terasvirta merupakan pengujian linieritas dengan tipe LM (*Lagrange Multiplier*).

Hipotesis :

H_0 : Model linier

H_1 : model non linier

Statistika uji :

$$F_{hit} = \frac{\frac{SSR_0 - SSR_1}{m}}{\frac{SSR_1}{(n-p-1-m)}} \quad (13)$$

Kriteria pengujian H_0 ditolak jika $F_{hitung} > F_{tabel}$ dengan derajat bebas m dan $N-r-l-m$ atau nilai $p-value < \alpha(0,05)$

Ketepatan Model Terbaik

Proses analisis runtun waktu atau time series, ada beberapa model yang dapat menunjukkan suatu keadaan data. Kriteria yang dapat digunakan untuk pemilihan Model ARIMA terbaik setelah dilakukan proses identifikasi model yaitu nilai Akaike's Information Criterion (AIC). Model terbaik adalah model yang memiliki nilai AIC terkecil.

Ketepatan Model Peramalan

Dalam membuat pemodelan peramalan dilakukan validasi untuk mengetahui kinerja metode peramalan yang digunakan. Pengujian tersebut digunakan untuk mengetahui *error* yang ada dalam model peramalan yang dibuat, salah satunya dengan menggunakan *Mean Absolute Percentage Error* (MAPE). *Mean Absolute Percentage Error* (MAPE) digunakan untuk mengukur kesalahan nilai dugaan dalam bentuk persentase rata-rata *absolute* kesalahan. Rumus menentukan nilai MAPE dapat dinyatakan dalam persamaan berikut ini:

$$MAPE = \frac{\sum_{t=1}^n \frac{|x_t - \hat{x}_t|}{x_t}}{n} \times 100\% \quad (14)$$

METODE PENELITIAN

Sumber Data

Dalam penelitian ini, data yang digunakan adalah data sekunder dari data bulanan Jumlah Penumpang Kereta Api Jabodetabek periode Januari 2014 hingga Desember 2020 dengan jumlah data sebanyak 84 data yang diperoleh dari Badan Pusat Statistik Indonesia. Data dibagi menjadi data *training* sebanyak 72 data dari bulan Januari 2014 hingga Desember 2019 dan data *testing* sebanyak 12 data dari bulan Januari 2020 hingga Desember 2020. Data *training* digunakan untuk membentuk suatu model, sedangkan data *testing* digunakan untuk memeriksa kebaikan model yang terbentuk dari data *training*.

Variabel dan Struktur Data

Variabel yang digunakan dalam penelitian ini yaitu Jumlah Penumpang Kereta Api di Jabodetabek periode Januari 2014 hingga Desember 2020. Struktur data dalam penelitian adalah sebagai berikut:

Tabel 1. Struktur Data

Periode	Jumlah Penumpang Kereta Api di Jabodetabek
Januari	Z_1
Februari	Z_2
Maret	Z_3
⋮	⋮

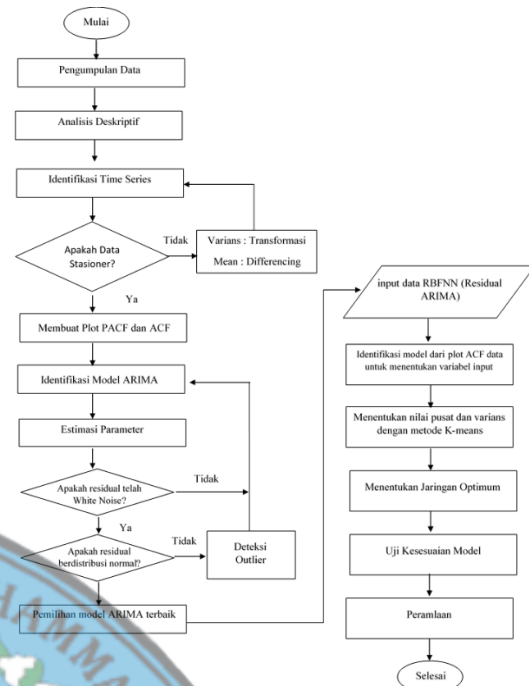
Langkah Penelitian

Langkah-langkah dalam penelitian ini yaitu :

1. Pembagian data training dan testing dengan perbandingan(72:12)
2. Melakukan Identifikasi *time series plot* untuk mengidentifikasi jenis pola data dan kestasioneran data.
3. Identifikasi model dengan membuat plot ACF dan PACF untuk mengidentifikasi orde pada plot ACF dan PACF sehingga diperoleh model sementara.
4. Melakukan estimasi parameter dari model sementara untuk menentukan parameter yang signifikan.
5. Melakukan *Diagnosting Checking* dengan melakukan pemeriksaan terhadap asumsi residual *white noise* dan normalitas pada model ARIMA yang signifikan.
6. Pemilihan model ARIMA terbaik dengan melihat nilai MAPE dan AIC terkecil.
7. Setelah mendapat model ARIMA terbaik, selanjutnya menghitung nilai residual dari model ARIMA untuk digunakan pada metode selanjutnya.
8. Input data residual untuk metode RBFNN.
9. Identifikasi model dengan melihat plot ACF data untuk menentukan jumlah *neuron input*. Penentuan *input* dilihat pada lag-lag yang signifikan pada plot ACF.
10. Menentukan nilai pusat dan varians dengan menggunakan metode *K-Means Clustering*. Banyaknya *cluster* yang terbentuk merupakan banyaknya *neuron* pada lapisan tersembunyi.
11. Menentukan jaringan optimum dengan menggunakan metode *trial and error* terhadap beberapa arsitektur dengan fungsi aktivasi Gaussian yang terbentuk. Pada proses ini digunakan metode kuadrat terkecil dan global ridge regression untuk memperoleh nilai-nilai bobot. Model terbaik dipilih dengan membandingkan nilai MAPE terkecil.
12. Melakukan uji kesesuaian model untuk memvalidasi hasil pengujian dengan melihat nilai MAPE terkecil dari data *testing*.

13. Melakukan peramalan periode kedepan.

Berikut ini gambaran langkah-langkah metode *Hybrid ARIMA-RBFNN*:

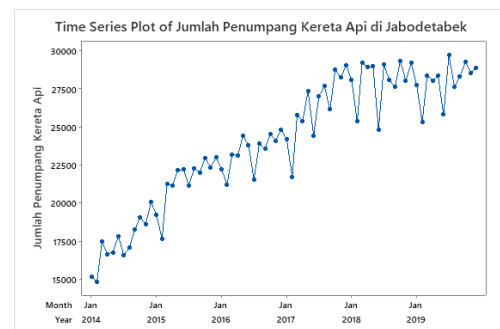


Gambar 2. Diagram Alir Penelitian

HASIL PENELITIAN dan PEMBAHASAN Statistika Deskriptif

Rata-rata jumlah penumpang kereta api di Jabodetabek selama periode 2014 hingga 2020 adalah sekitar 22498,8690 dengan standar deviasi sekitar 6085,03151. Hal ini menunjukkan bahwa varians jumlah penumpang kereta api di Jabodetabek cukup tinggi. Jumlah penumpang paling banyak terjadi pada bulan Juli 2019 sebanyak 29714 penumpang dan paling sedikit terjadi pada bulan Mei 2020 sebanyak 5077 penumpang

Identifikasi *Time Series Plot*

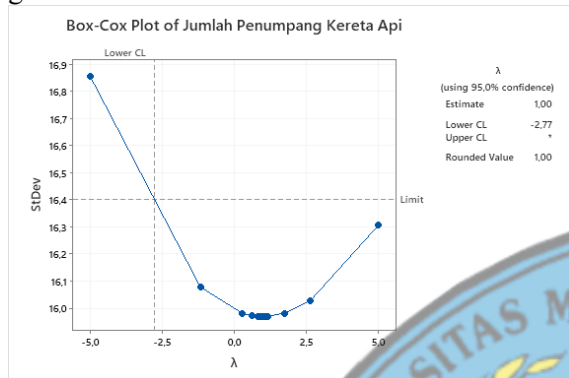


Gambar 3. *Time Series Plot*

Pada gambar diatas, menunjukkan bahwa data dipengaruhi pola trend yaitu gerakan naik dari kiri bawah ke kanan atas pada grafik time series dan terdapat pola musiman pada setiap bulannya. Grafik time series plot tersebut juga menunjukkan bahwa data belum stasioner terhadap rataaan maupun varian. Oleh karena itu, diperlukannya pemeriksaan lebih lanjut.

Uji Stasioneritas

Hasil Uji Box-Cox dapat dilihat pada gambar 4 berikut:



Gambar 4. Box-Cox Transformasi

Pada gambar 4 menunjukkan bahwa data jumlah penumpang kereta api di Jabodetabek yang sudah di transformasi sebanyak 2 kali memiliki nilai *rounded value* sebesar 1, sehingga dapat dikatakan data sudah stasioner terhadap varians.

Hasil uji ADF dijelaskan pada Tabel 3. berikut :

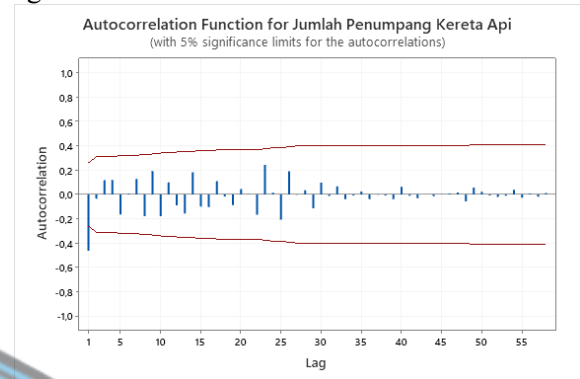
Tabel 2. Uji ADF

Uji ADF	
Nilai <i>p-value</i>	
Sebelum di <i>Differencing</i>	Setelah di <i>Differencing</i>
0,7984	0,02494

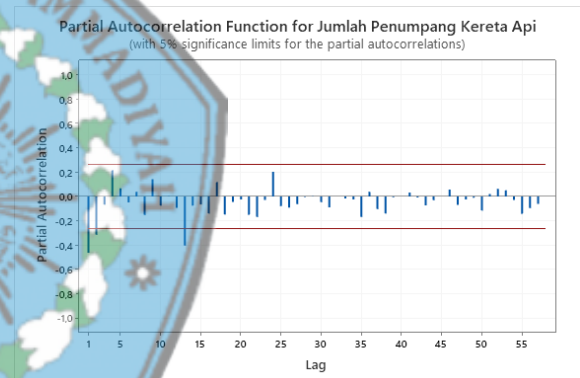
Berdasarkan tabel 2 Hasil Uji *Augmented Dikey Fuller* diperoleh nilai *p-value* sebesar $0,7984 > \alpha = 0,05$ yang artinya data belum stasioner terhadap rataaan. Oleh karena itu, untuk menstasionerkan data diperlukan *differencing* terhadap data. Setelah dilakukan *differencing* diperoleh nilai *p-value* sebesar $0,02494 < \alpha = 0,05$. Berdasarkan hasil tersebut dapat disimpulkan data sudah stasioner terhadap rataaan.

Identifikasi Model ARIMA

Langkah selanjutnya setelah data sudah stasioner terhadap varians dan rataaan adalah membuat plot ACF (*Autocorrelation Function*) dan plot PACF (*Partial Autocorrelation Function*) untuk mengidentifikasi model ARIMA sementara dengan melihat lag yang signifikan dengan sedikit mungkin parameter agar model lebih stabil.



Gambar 5. Plot ACF setelah Stasioner



Gambar 6. Plot PACF setelah stasioner

Berdasarkan gambar 5 menunjukkan plot ACF berbentuk pola *cut off* pada lag ke 2 sehingga diidentifikasi MA musiman, sedangkan pada gambar 6 terlihat plot PACF berbentuk pola *cut off* sehingga diidentifikasi AR musiman. Pada diferensi musiman dan non-musiman yang dilakukan adalah sebanyak satu kali dan banyak periode yang dilakukan adalah 12 bulan. Maka model awal yang diperoleh adalah $ARIMA(1,1,1)(1,1,1)^{12}$. Namun tidak menutup kemungkinan terdapat model alternatif lainnya yang mungkin terbentuk yaitu $ARIMA(0,1,1)(1,1,1)^{12}$, $ARIMA(1,1,0)(1,1,1)^{12}$, $ARIMA(2,1,0)(1,1,0)^{12}$, $ARIMA(2,1,0)(1,1,1)^{12}$, $ARIMA(2,1,0)(0,1,1)^{12}$, $ARIMA(3,1,1)(0,1,1)^{12}$, $ARIMA(3,1,3)(1,1,0)^{12}$, $ARIMA(3,1,1)(1,1,1)^{12}$.

ARIMA (3,1,0)(1,1,0)¹², ARIMA (3,1,0)(0,1,1)¹², ARIMA (3,1,0)(1,1,1)¹².

Setelah didapatkan model sementara, langkah selanjutnya yaitu melakukan estimasi parameter untuk mendapatkan model yang tepat. Estimasi parameter dilakukan untuk setiap model dan dilihat apakah parameter tersebut signifikan atau tidak. Model yang baik adalah model yang memiliki nilai *p-value* kurang dari nilai $\alpha = 0,05$.

Tabel 3. Estimasi Parameter Model Sementara

Model (1,1,0)(1,1,1) ¹²			
Parameter	AIC	P-value	Kesimpulan
AR 1		0,0001	Signifikan
SAR 1	991,01	0,0000	Signifikan
SMA 1		0,0201	Signifikan

Dari duabelas model yang diujikan, hanya terdapat 1 model yang memiliki parameter yang signifikan yaitu model ARIMA (1,1,0)(1,1,1)¹² dengan nilai AIC sebesar 991,01.

Diagnosing Checking

Pada tahap selanjutnya, dilakukan pemeriksaan tentang pengujian asumsi residual dari model ARIMA agar suatu model tersebut mampu mewakili pola data. Pengujian asumsi yang harus terpenuhi yaitu uji *white noise* residual dan uji normalitas residual

Tabel 4. Statistik Ljung-Box nilai residual model ARIMA

Model (1,1,0)(1,1,1) ¹²	
P-value	Kesimpulan
0,181	White Noise

Berdasarkan tabel 4 diatas, dapat diketahui bahwa model ARIMA (1,1,0)(1,1,1)¹² yang diujikan memenuhi asumsi *white noise*. Hal ini dikarenakan nilai *p-value* $> \alpha$ (0,05).

Tabel 5. Hasil Uji Normalitas Residual

Model (1,1,0)(1,1,1) ¹²	
P-value	Kesimpulan
0,203	Berdistribusi normal

Berdasarkan tabel 4.5 dapat dijelaskan bahwa model ARIMA (1,1,0)(1,1,1)¹² yang diujikan, model tersebut berdistribusi normal. Hal ini dikarenakan nilai *p-value* $> \alpha$ (0,05). Oleh

karena itu, model ARIMA (1,1,0)(1,1,1)¹² akan digunakan pada metode selanjutnya.

Uji Linieritas Data

Uji linieritas dilakukan untuk mengetahui apakah data berpola linier. Uji linieritas pada data jumlah penumpang kereta api di Jabodetabek menggunakan Uji Teraesvirta. Hasil uji linieritas data dapat dilihat pada tabel 6 berikut:

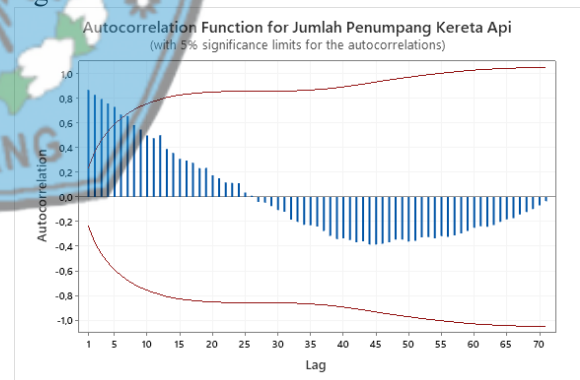
Tabel 6. Hasil Uji Linieritas Data

Uji Teraesvirta	
<i>p-value</i>	0,01917252

Berdasarkan tabel 4.6 hasil uji linieritas data, menunjukkan nilai *p-value* sebesar $0,01917252 < \alpha = 0,05$. Hal ini menunjukkan bahwa data jumlah penumpang kereta api di Jabodetabek terdapat pola nonlinier.

Penentuan Input RBFNN

Pada tahap awal pembentukan model RBFNN yaitu menentukan matriks input. Matriks input merupakan matrik yang meliputi variable input. Variabel input dalam model ini adalah lag yang memiliki nilai autokorelasi yang signifikan.



Gambar 7. Plot ACF Data

Berdasarkan pada Gambar 7 Plot ACF Data jumlah penumpang kereta api di Jabodetabek terlihat bahwa nilai autokorelasi yang signifikan terdapat pada lag 1 sampai lag 6. Dengan demikian diperoleh 6 variabel *input* yang akan digunakan dalam model RBFNN untuk data jumlah penumpang kereta api di Jabodetabek yaitu Z_{t-1} , Z_{t-2} , Z_{t-3} , Z_{t-4} , Z_{t-5} , dan Z_{t-6} . Banyak variabel *input* juga merupakan banyaknya *neuron* yang terbentuk

pada lapisan *input* jaringan RBFNN, sehingga pada penelitian ini terdapat 6 *neuron* pada lapisan *input*.

Menentukan Jaringan Optimum

Setelah menentukan data *input* tahap selanjutnya yaitu membentuk kelompok dengan menggunakan metode *k-menas clustering*. Penentuan banyak kelompok data dilakukan dengan menggunakan metode *trial and error* yaitu dengan melakukan pengelompokan data *training* ke dalam 2 kelompok, 3 kelompok, dan seterusnya hingga terbentuk kelompok-kelompok yang terdiri dari dua anggota. Setiap kelompok yang terbentuk diberikan nilai pembobot yaitu nilai *center* dan *varians*.

Banyaknya kelompok pada data *training* yaitu banyaknya neuron pada lapisan tersembunyi. Setiap *neuron* diberikan satu nilai pembobot (*w*). Nilai pembobot yang didapatkan dari data *training* diujikan pula pada data *testing*. Nilai pembobot antara lapisan tersembunyi dengan lapisan *output* dihitung dengan menggunakan metode kuadrat terkecil. Dari semua kelompok yang diujikan, diambil satu kelompok yang terbaik dengan melihat nilai MAPE terkecil pada data *training* dan data *testing*.

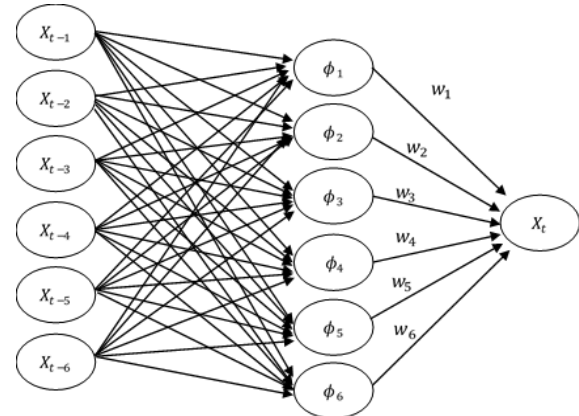
Setelah mendapatkan parameter dan model untuk masing-masing variasi kalender dan lokasi, maka langkah selanjutnya adalah pengujian asumsi apakah galat atau residual memenuhi asumsi *white noise*. Untuk menguji asumsi residual memenuhi *white noise* digunakan uji *Ljung-Box* dengan hasil yang tersaji pada Tabel 10. sebagai berikut :

Tabel 7. Hasil Nilai MAPE untuk Penentuan banyak *Neuron* Tersembunyi.

Kluster	MAPE <i>training</i>	MAPE <i>testing</i>
2 Kluster	206,29	100,74
3 Kluster	967,49	105,14
4 Kluster	978,33	103,92
5 Kluster	793,98	101,96
6 Kluster	201,81	100,58

Berdasarkan tabel 7 terlihat bahwa nilai MAPE terkecil untuk data *training* dan data *testing* terletak pada 6 kluster. Dengan demikian, model RBFNN yang digunakan untuk

meramalkan jumlah penumpang kereta api di Jabodetabek mempunyai arsitektur jaringan 6 *neuron input*, 6 *neuron* pada lapisan tersembunyi dan 1 *neuron* pada lapisan *output*. Arsitektur jaringan RBFNN yang terbentuk dapat ditunjukkan pada gambar 8 berikut ini:



Gambar 8. Arsitektur Jaringan RBFNN Jumlah Penumpang Kereta Api (6-6-1)

Pada hasil pembelajaran selain terbentuknya model jaringan RBFNN juga diperoleh bobot-bobot yang akan digunakan untuk meramalkan jumlah penumpang kereta api di Jabodetabek pada periode selanjutnya. Pembobot yang digunakan pada jaringan RBFNN adalah sebagai berikut:

$$c = \begin{pmatrix} -174,1750 & -224,9400 & -1116,6925 & 694,1375 & 431,4878 & -429,2910 \\ -45,4027 & 4,3380 & -628,0900 & -1126,4225 & 933,7667 & 500,5130 \\ -129,2786 & -196,3240 & 746,9312 & -311,0800 & -1097,7611 & 549,44390 \\ -191,2895 & 489,2420 & -873,1713 & 648,6333 & -512,5889 & -72,9740 \\ 249,4527 & 1319,4780 & -629,3663 & -412,8858 & 162,2856 & -823,8610 \\ 440,7591 & -1783,8880 & 350,1725 & 3678,442 & -69,6844 & -86,9340 \end{pmatrix}$$

$$w_b = -6,635; w_1 = -5,7508; w_2 = -2,527; w_3 = -2,6777; w_4 = -2,5976; w_5 = -6,5969; w_6 = -4,0913$$

Berdasarkan pada gambar 8 arsitektur jaringan RBFNN, model yang terbentuk dinyatakan sebagai berikut:

$$\hat{X}_t = \sum_{j=1}^m w_j \phi_j(x) = \sum_{j=1}^6 w_j \phi_j(x)$$

$$\hat{X}_t = -5,7508 \phi_1(x) - 2,527 \phi_2(x) - 2,6777 \phi_3(x) - 2,5976 \phi_4(x) - 6,5969 \phi_5(x) - 4,0913 \phi_6(x) - 6,635$$

Nilai $\phi_j(x)$ dinyatakan sebagai berikut:

$$\phi_j(x) = \exp\left(-\sum_{i=1}^k \left(\frac{|x_i - c_{ji}|}{2\sigma_j^2}\right)^2\right)$$

$$\phi_j(x) = \exp\left[-\frac{1}{2}\left(\left(\frac{|x_1 - c_{j1}|}{\sigma_1}\right)^2 + \left(\frac{|x_i - c_{ji}|}{\sigma_j}\right)^2 + \dots + \left(\frac{|x_m - c_{jm}|}{\sigma_m}\right)^2\right)\right]$$

Dari model yang terbentuk, maka pembentukan model RBFNN pada bulan Januari 2021 yaitu sebagai berikut:

$$\phi_1(x) = \exp\left[-\frac{1}{2}\left\{\left(\frac{|-17932,84-(-174,1750)|}{3659,845}\right)^2 + \left(\frac{|-17200,1-(-45,4027)|}{3659,845}\right)^2 + \left(\frac{|-18838,34-(-191,2895)|}{3659,845}\right)^2 + \left(\frac{|-16903,93-249,4527|}{3659,845}\right)^2 + \left(\frac{|-18499,28-440,7591|}{3659,845}\right)^2\right\}\right]$$

$$\phi_1(x) = \exp(-73.97685248)$$

$$\phi_1(x) = -4.28538E - 32$$

$$\phi_2(x) = \exp\left[-\frac{1}{2}\left\{\left(\frac{|-17932,84-(-224,9400)|}{2997,849}\right)^2 + \left(\frac{|-17200,1-(-4,3380)|}{2997,849}\right)^2 + \left(\frac{|-18838,34-(-489,2420)|}{2997,849}\right)^2 + \left(\frac{|-16903,93-1319,4780|}{2997,849}\right)^2 + \left(\frac{|-18499,28-(-1783,8880)|}{2997,849}\right)^2\right\}\right]$$

$$\phi_2(x) = \exp(-109.23803)$$

$$\phi_2(x) = -9.14393E - 48$$

$$\phi_3(x) = \exp\left[-\frac{1}{2}\left\{\left(\frac{|-17932,84-(-1116,6925)|}{3559,529}\right)^2 + \left(\frac{|-17200,1-(-628,0900)|}{3559,529}\right)^2 + \left(\frac{|-18838,34-(-873,1713)|}{3559,529}\right)^2 + \left(\frac{|-16903,93-(-692,3663)|}{3559,529}\right)^2 + \left(\frac{|-18499,28-(-350,1725)|}{3559,529}\right)^2\right\}\right]$$

$$\phi_3(x) = \exp(-75.14654195)$$

$$\phi_3(x) = 2.31351E - 33$$

$$\phi_4(x) = \exp\left[-\frac{1}{2}\left\{\left(\frac{|-17932,84-(-694,1375)|}{3678,442}\right)^2 + \left(\frac{|-17200,1-(-1126,4225)|}{3678,442}\right)^2 + \left(\frac{|-18838,34-(-648,6333)|}{3678,442}\right)^2 + \left(\frac{|-16903,93-(-412,8858)|}{3678,442}\right)^2 + \left(\frac{|-18499,28-(-671,2117)|}{3678,442}\right)^2\right\}\right]$$

$$\phi_4(x) = \exp(-71.66246682)$$

$$\phi_4(x) = 7.54026E - 32$$

$$\phi_5(x) = \exp\left[-\frac{1}{2}\left\{\left(\frac{|-17932,84-(-431,4878)|}{3148,49}\right)^2 + \left(\frac{|-17200,1-(-933,7667)|}{3148,49}\right)^2 + \left(\frac{|-18838,34-(-512,5889)|}{3148,49}\right)^2 + \left(\frac{|-16903,93-(-162,2856)|}{3148,49}\right)^2 + \left(\frac{|-18499,28-(-69,6884)|}{3148,49}\right)^2\right\}\right]$$

$$\phi_5(x) = \exp(-99.2567951)$$

$$\phi_5(x) = 7.82207E - 44$$

$$\phi_6(x) = \exp\left[-\frac{1}{2}\left\{\left(\frac{|-17932,84-(-429,2910)|}{3559,529}\right)^2 + \left(\frac{|-17200,1-(-500,5130)|}{3559,529}\right)^2 + \left(\frac{|-18838,34-(-72,9740)|}{3559,529}\right)^2 + \left(\frac{|-16903,93-(-823,8610)|}{3559,529}\right)^2 + \left(\frac{|-18499,28-(-86,9340)|}{3559,529}\right)^2\right\}\right]$$

$$\phi_6(x) = \exp(-218.094315)$$

$$\phi_6(x) = 1.91797E - 95$$

$$\widehat{X}_t = \sum_{j=1}^6 w_j \phi_j(x)$$

$$\widehat{X}_t = -5,7508(-4.28538E - 32) - 2,527(-9.14393E - 48) - 2,6777(2.31351E - 33) - 2,5976(7.54026E - 32) - 6,5969(7.82207E - 44) - 4,0913(1.91797E - 95) - 6,635$$

$$\widehat{X}_t = -6.635$$

Dengan demikian, hasil pembentukan model RBFNN pada bulan Januari 2021 sesuai perhitungan diatas yaitu sebesar -6,635. Sehingga dengan cara yang sama dapat dihitung pembentukan model RBFNN pada bulan Februari 2021 hingga Februari 2020.

Perbandingan ARIMA dan Hybrid ARIMA-RBFNN

Setelah melakukan pemodelan pada data training dengan metode ARIMA dan metode Hybrid ARIMA-RBFNN. Selanjutnya dilakukan peramalan untuk data *testing* jumlah penumpang kereta api periode Januari 2020-Desember 2020. Untuk membandingkan hasil peramalan pada data testing digunakan nilai *Mean Absolute Percentage Error* (MAPE) sebagai kriteria kebaikan model. Hasil perbandingan nilai MAPE dari kedua model dapat dilihat pada tabel 8 berikut ini:

Tabel 8. Nilai MAPE model ARIMA dan Hybrid ARIMA-RBFNN

	ARIMA (1,1,0)(1,1,1) ¹²	Hybrid ARIMA- RBFNN
MAPE	182,3232	182,2542684

Berdasarkan tabel 8 diatas, menunjukkan bahwa peramalan dengan model ARIMA didapatkan nilai MAPE sebesar 182,3232 sedangkan peramalan dengan model Hybrid ARIMA-RBFNN didapatkan nilai MAPE sebesar 182,2542684. Dengan demikian, dapat disimpulkan bahwa model Hybrid ARIMA-RBFNN lebih baik dibandingkan dengan model ARIMA untuk meramalkan jumlah penumpang kereta api di Jabodetabek.

Peramalan Jumlah Penumpang Kereta Api di Jabodetabek

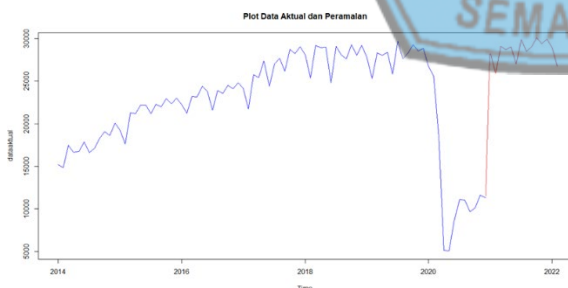
Dari proses sebelumnya didapatkan model terbaik yang digunakan untuk

meramalkan jumlah penumpang kereta api di Jabodetabek periode 14 bulan kedepannya yaitu model Hybrid ARIMA-RBFNN. Hasil peramalan menggunakan model terbaik disajikan pada tabel 9 dibawah ini:

Tabel 9. Hasil Peramalan dengan menggunakan metode *Hybrid* ARIMA-RBFNN
Jumlah Penumpang Kereta Api di Jabodetabek

Tahun	Bulan	Hasil Peramalan
2021	Januari	28252
2021	Februari	25947
2021	Maret	29075
2021	April	28672
2021	Mei	29060
2021	Juni	27047
2021	Juli	29914
2021	Agustus	28509
2021	September	29071
2021	Oktober	30108
2021	November	29405
2021	Desember	29960
2022	Januari	28927
2022	Februari	26764

Berdasarkan tabel 9 hasil peramalan jumlah penumpang kereta api di Jabodetabek pada periode Januari 2021 sampai Februari 2022 mengalami fluktuasi setiap bulannya. Jumlah penumpang tertinggi terjadi pada bulan Oktober 2021 sedangkan jumlah penumpang terendah terjadi pada bulan Februari 2021. Grafik peramalan periode 14 bulan ke depan dapat dilihat pada gambar 9 berikut ini:



SIMPULAN dan SARAN

Simpulan

Pada model ARIMA (1,1,0)(1,1,1)¹² didapatkan nilai *Mean Square Percentage Error* (MAPE) sebesar 182,3232 sedangkan model *Hybrid* ARIMA-RBFNN didapatkan nilai MAPE sebesar 182.2542684. Dengan demikian model terbaik yang digunakan adalah model

Hybrid ARIMA-RBFNN. Model terbaik tersebut merupakan gabungan dari model ARIMA (1,1,0)(1,1,1)¹² dan residualnya digunakan kembali untuk dimodelkan dengan metode *Radial Basis Function Neural Network*. Pemodelan dengan menggunakan RBFNN didapatkan arsitektur jaringan terbaik yaitu 6-6-1 (6 *neuron input*, 6 *neuron hidden* dan 1 *neuron output*).

Hasil peramalan jumlah penumpang kereta api di Jabodetabek pada bulan Januari 2021 hingga Februari 2022 mengalami fluktuasi setiap bulannya. Jumlah penumpang tertinggi terjadi pada bulan Oktober 2021 dan terendah terjadi pada bulan Februari 2021.

Saran

Saran yang dapat diberikan pada penelitian selanjutnya dapat menggunakan metode *hybrid* yang lain seperti SARIMAX-ANN, ARIMAX-RBFNN. Pada metode RBFNN, untuk menentukan banyaknya neuron pada lapisan tersembunyi dapat menggunakan algoritma *clustering* lain seperti *Fuzzy C Means*, algoritma SOM dan lainnya.

Daftar Pustaka

- Afif, Irman. 2014. Peramalan Indeks Saham Syariah Indonesia (ISSI) Menggunakan Model Backpropagation Neural Network dan Radial Basis Function Neural Network. Skripsi. Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Yogyakarta : Yogyakarta.
- Agustin, Rima Ika. 2018. Peramalan Data Intermiten menggunakan metode Autoregressive integrated moving average – artificial neural network (ARIMA-NN), Skripsi, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember, Surabaya.
- Aulia, Firda. 2020. Peramalan Jumlah Pengiriman Barang Melalui Kereta Api di Indonesia dengan Metode Hybrid ARIMA-ANN. Skripsi. Universitas Muhammadiyah Semarang.
- Badan pusat Statistik. Transportasi. Data Jumlah Penumpang Kereta Api di Indonesia.

- BPS Statistik Indonesia. Jakarta Indonesia.
- Banurea, Monika. 2020. Jaringan Syaraf Tiruan Untuk Memprediksi Jumlah Permintaan Pemasangan Indihome Dengan Menggunakan Metode Radial Basic Function Neural Netrwok (RBFNN). Jurnal Majalah Ilmiah Informasi dan Teknologi Ilmiah (2) :126-129.
- Budiastawan, I Gede., et.al. 2019. Prediksi dan Akurasi Nilai Tukar Mata Uang Rupiah Terhadap US Dollar Menggunakan Radial Basis Function Neural Network. Jurnal Elektronik Ilmu Komputer Udayana (4).
- Hikmah, AL. 2017. Peramalan Deret Waktu Dengan Menggunakan Autoregressive (AR), Jaringan Syaraf Tiruan Radial Basis Function (RBF) Dan Hybrid AR-RBF Pada Inflasi Indonesia, Skripsi, Program S1 Matematika Universitas Negeri, Semarang.
- Juliaristi, Fajarani. 2014. Peramalan Banyak Kasus Demam Berdarah di D.I.Yogyakarta dengan Model Radial Basis Function Neural Network, Skripsi, Program Studi Matematika, Universitas Negeri Yogyakarta, Yogyakarta.
- Lucianna, Pramita. 2017. Peramalan jumlah penumpang penerbangan di terminal bandara internasional juanda menggunakan metode ARIMA Box – Jenkins dan hybrid Autoregressive integrated moving average – artificial neural network (ARIMA-NN), Skripsi, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember, Surabaya.
- Purba, Novelina. 2016. Pemodelan Data Curah Hujan di Kabupaten Banyuwangi dengan Menggunakan ARIMA dan Radial Basis Function Neural Network, Skripsi, Program Studi S1 Matematika Institut Teknologi Sepuluh Nopember, Surabaya.
- Sani, Zulfiar. 2010. Transportasi : Suatu Pengantar. Jakarta: Universitas Indonesia-Press.
- Sari, Virgania et. Al. 2020. Prediksi Kecepatan Angin Dalam Mendeteksi Gelombang Air Laut Terhadap Skala Beaufort dengan Metode Hybrid ARIMA-NN (studi kasus: kabupaten Lombok Barat 2019). Jurnal Statistika Vol 8, No 1.
- Susilokarti, Dyah et al. 2015. Studi Komparasi Prediksi Curah Hujan Metode Fast Fourier Transformation (FFT), Autoregressive Integrated Moving Average (ARIMA) dan Artificial Neural Network (ANN). AGRITECH (2).
- Wahida Yanti. 2015. Peramalan jumlah penumpang dari pelayanan dalam negeri di pelabuhan kota Makasa menggunakan metode Seasonal Autoregressive Integrated Moving Average (SARIMA), Skripsi, Jurusan Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Alauddin, Makassar.
- Wiyanti, DT. 2012. Peramalan Deret Waktu Menggunakan Model Fungsi Radial Basis (RBF) dan Autoregressive Integrated Moving Average (ARIMA). Jurnal MIPA. Universitas Gajah Mada: Yogyakarta.
- Wulandari, Ayu. 2017. Peramalan Harga Minyak Mentah Dunia (Crude Oil) Menggunakan Metode Radial Basis Function Neural Network (RBFNN). Skripsi. Universitas Mulawarman: Samarinda.
- Warsito, B. 2009. Kapita Selekta Statistika Neural Network. BP UNDIP: Semarang.
- Zhang, G.P. 2003. Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. Neurocomputing, 50, 159-175.

