

PERAMALAN JUMLAH PENUMPANG KERETA API EKONOMI KERTAJAYA MENGGUNAKAN METODE ARIMA DAN ANFIS

Oleh:

Ilafi Andalita (1311100038)

Dosen Pembimbing:
Dr. Irhamah, M.Si.

JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER

Surabaya, 2015

AGENDA



Pendahuluan



Tinjauan Pustaka



Metodologi
Penelitian



Analisis dan
Pembahasan



Kesimpulan
dan Saran



Daftar Pustaka

Lampiran

Latar Belakang



Kereta Api Kelas Ekonomi Kertajaya

Salah satu cara mengetahui kebutuhan pengguna alat transportasi kereta api kelas ekonomi adalah dengan peramalan jumlah penumpang

Sepanjang tahun 2014 hingga 2015 terdapat penyesuaian tarif kereta api kelas ekonomi mengakibatkan jumlah penumpang berfluktuatif tinggi

Berdasarkan pengujian linieritas, *time series* jumlah penumpang termasuk dalam nonlinier *time series*

Pemodelan menggunakan pendekatan linier seperti ARIMA Box-Jenkins tidak selalu memberikan peramalan yang baik karena terikat beberapa asumsi dalam membangun model.

Oleh karena itu penelitian ini mengusulkan penggunaan metode ANFIS sebagai pendekatan soft computing yang diharapkan memberikan kinerja lebih baik dalam pemodelan nonlinier dan dibandingkan dengan hasil dari ARIMA.

Latar Belakang

No.	Peneliti	Judul Penelitian	Hasil Penelitian
1	(Widiyanto 2013)	Peramalan Jumlah Penumpang Kereta Api Logawa Dan Sritanjung Di Daop (Daerah Operasional) IX Jember	model regresi time series dengan efek variasi kalender baik digunakan
2	(Widhianti dan Wutaqa, 2013)	Peramalan Banyak Penumpang Kereta DAOP VI Yogyakarta Menggunakan Model Time Series dengan Variasi Kalender Islam Regarima	model penggabungan model ARIMA dan regresi time series dengan efek variasi kalender baik digunakan
3	(Padang, Tarigan, dan Sinulingga, 2013)	Peramalan Jumlah Penumpang Kereta Api Medan-Rantau Prapat dengan Metode Pemulusan Eksponensial Holt-Winters	Metode Pemulusan Eksponensial Holt-Winters baik untuk peramalan

**Penelitian Sebelumnya
Menggunakan Metode Statistik**

Latar Belakang



SOFT COMPUTING

Pendekatan untuk melakukan komputasi dengan **meniru akal manusia** dan memiliki kemampuan untuk **menalar dan belajar** pada lingkungan yang penuh dengan **ketidakpastian dan ketidaktepatan**

(Zadeh, 1992)

ANFIS memiliki arsitektur dari kedua komponen *soft computing* yaitu jaringan syaraf tiruan dan sistem/ logika fuzzy, sehingga ANFIS memiliki kelebihan yang dimiliki oleh kedua komponen tersebut .

- jaringan syaraf tiruan: melakukan pembelajaran dan adaptasi
- Sistem Fuzzy : Aturan *if-then* dari sistem fuzzy membentuk komponen kunci dari *fuzzy inference system*

(Avci, 2008).

Latar Belakang

No.	Peneliti	Judul Penelitian	Hasil Penelitian
1	(Nurvitasari, 2012)	Pendekatan Fungsi Transfer sebagai Input Adaptive Neuro-Fuzzy Inference System (ANFIS) dalam peramalan Kecepatan ANgin Rata-rata Harian di Sumenep	model ANFIS dengan input dari fungsi transfer memberikan akurasi lebih baik dibandingkan dengan model fungsi transfer
2	(Varedi, 2011)	Forecasting Seat Sales in Passenger Airlines Using the Reverse Traffic	meramlakan efek penjualan
3	(Gyu-jin, Jin-pyung, Jee-hyong, & Moon-hyun, 2011)	ANFIS based Short-term Prediction of Passenger Volume in Urban Rail Transit	model ANFIS memberikan akurasi >93%
4	(Atsalakis, Nezis, dan Matalliotakis, 2008)	Forecasting Mortality Rate Using a Neural Network with Fuzzy Inference System	model ANFIS memberikan akurasi yang lebih baik dibandingkan model AR dan ARIMA
5	Galavi, Mirzaei, Shui, dan Valizadeh (2013)	Klang River-Level Forecasting Using ARIMA and ANFIS	kedua model memiliki nilai MAPE sebesar <0.5% namun model ARIMA memberikan nilai yang lebih kecil
6	Yayar, Hekim, Yilmaz, & Bakirci (2011)	A Comparison of ANFIS and ARIMA Techniques in the Forecasting of Electric Energy Consumption of Tokat Province in Turkey	model ANFIS memberikan akurasi yang lebih baik dibandingkan model ARIMA

**Penelitian Sebelumnya
Menggunakan ANFIS**

Permasalahan

Jumlah penumpang KA Kertajaya merupakan time series yang nonlinier,

Pemodelan menggunakan pendekatan linier seperti ARIMA Box-Jenkins tidak selalu memberikan peramalan yang baik karena terikat beberapa asumsi dalam membangun model.

Sehingga, bagaimana peramalan jumlah penumpang KA Kertajaya dan ANFIS yang merupakan pendekatan *soft computing*? Lalu bagaimana perbandingan hasil pemodelan ANFIS dan ARIMA untuk meramalkan jumlah penumpang?

Tujuan

1. Mengetahui karakteristik jumlah penumpang berdasarkan eksplorasi data.

2. Meramalkan jumlah penumpang Kereta Api Ekonomi Kertajaya dengan pendekatan linier yaitu ARIMA Box-Jenkins.

3. Menerapkan teknik *soft computing* yaitu ANFIS untuk meramalkan jumlah penumpang Kereta Api Ekonomi Kertajaya yang merupakan *time series* yang nonlinier.

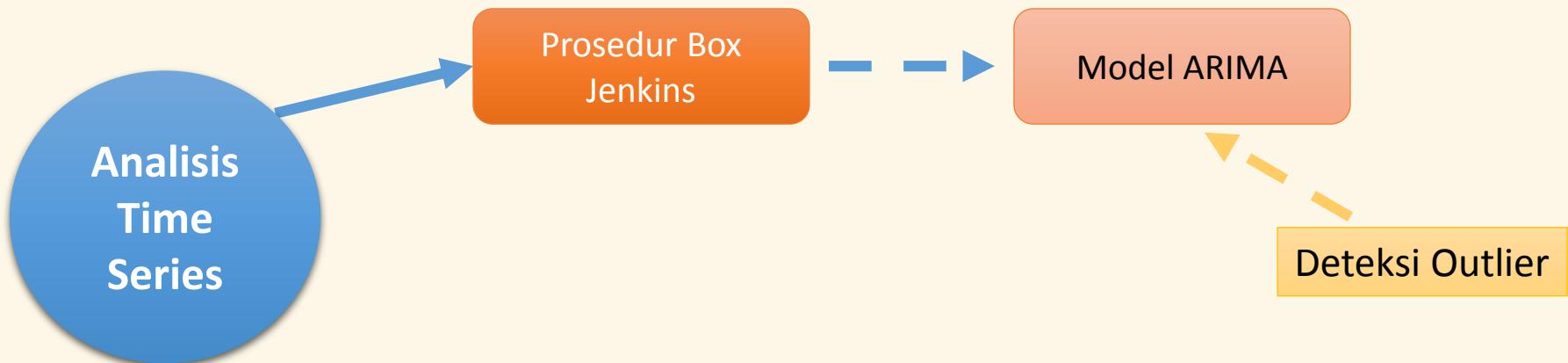
4. Melakukan perbandingan model dari pendekatan linier dan teknik *soft computing* yaitu model ARIMA Box-Jenkins dan ANFIS serta memperoleh model yang terbaik.

Manfaat

- Memberikan **informasi tertulis** kepada PT.KAI DAOP VIII Surabaya atas hasil penelitian.
- Menerapkan **teknik soft computing** sebagai alternatif dari pendekatan statistik.

Batasan Masalah

Penelitian ini dilakukan terhadap data jumlah penumpang harian Kereta Api kertajaya yang berangkat dari Stasiun Pasar Turi, Surabaya menuju Stasiun Pasar Senin, Jakarta



Metodologi statistik yang dapat menganalisis suatu rangkaian pengamatan yang didasarkan oleh **urutan waktu** pengamatan dengan **interval waktu yang sama** dimana setiap pengamatan tersebut **berkorelasi**

Prosedur Box-Jenkins

Identifikasi Model

1. Membuat plot *time series* dan memilih transformasi yang tepat.
- ✓ Transformasi untuk menstasionerkan data dalam variansi (Transformasi Box-cox)

$$T(Z_t) = \frac{Z_t^\lambda - 1}{\lambda}, \text{ untuk } \lambda \neq 0$$

$$\lim_{\lambda \rightarrow 0} T(Z_t) = \frac{Z_t^\lambda - 1}{\lambda} \text{ untuk } \lambda = 0$$

Nilai Lambda yang Umum digunakan dan Transformasi Terkait:

Nilai λ	Transfomasi
-1,0	$\frac{1}{Z_t}$
-0,5	$\frac{1}{\sqrt{Z_t}}$
0,0	$\ln Z_t$
0,5	$\sqrt{Z_t}$
-0,5	Z_t (tidak ada transfomasi)

- ✓ *Differencing* untuk menstasionerkan data dalam rata-rata

Prosedur Box-Jenkins

Identifikasi Model

2. Memeriksa sampel ACF dan PACF untuk mengonfirmasi orde yang tepat untuk *differencing*.

✓ Sampel ACF :

$$\hat{\gamma}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2}$$

✓ Sampel PACF :

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{\sum_{j=1}^k \hat{\phi}_{kj}}$$

Apabila ACF menurun atau merosot dengan lambat dan PACF putus setelah lag 1, maka hal hal tersebut mengindikasikan bahwa *differencing* diperlukan atau dapat melakukan uji *unit root*

3. Memeriksa sampel ACF dan PACF dari data yang telah stasioner.

Proses	ACF	PACF
AR(p)	Turun cepat secara eksponensial/ sinusoidal	Terputus setelah lag p
MA(q)	Terputus setelah lag q	Turun cepat secara eksponensial/ sinusoidal
ARMA (p,q)	Turun cepat setelah lag (q-p)	Turun cepat setelah lag (p-q)

Prosedur Box-Jenkins

Estimasi Parameter

Uji Hipotesis:

$$H_0: \phi_i = 0 \text{ atau } \theta_j = 0, i=1,2,\dots,p; j=1,2,\dots,q$$

$$H_1: \phi_i \neq 0 \text{ atau } \theta_j \neq 0, i=1,2,\dots,p; j=1,2,\dots,q$$

Statistik Uji:

$$t_{hitung} = \frac{\hat{\phi}}{s.e.(\hat{\phi})} \text{ atau } t_{hitung} = \frac{\hat{\theta}}{s.e.(\hat{\theta})}$$

H_0 ditolak apabila $|t_{hitung}| \geq t_{\frac{\alpha}{2}, n-p}$ dengan n menunjukkan banyaknya data dan p menunjukkan banyak parameter pada model .

Prosedur Box-Jenkins

Pemeriksaan Diagnostik

1

Pengujian Asumsi Residual Distribusi Normal (Uji Kolmogorov-Smirnov)

Uji Hipotesis:

$$H_0: F(x) = F_0(x)$$

$$H_1: F(x) \neq F_0(x).$$

Statistik Uji:

$$D = \sup_x |S(x) - F_0(x)|$$

H_0 ditolak apabila $D > D_{(1-\alpha;n)}$

Keterangan:

$F(x)$: fungsi distribusi belum diketahui

$F_0(x)$: fungsi distribusi yang dihipotesiskan (normal)

$S(x)$: fungsi peluang kumulatif yang dihitung dari data sampel

n : ukuran sampel

2

Pengujian Asumsi Residual White Noise (uji portmanteau lack of fit)

Uji Hipotesis:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_K = 0$$

H_1 : minimal terdapat $\rho_k \neq 0$

Statistik Uji:

$$Q = n(n+2) \sum_{k=1}^K (n-k)^{-1} \hat{\rho}_k^2$$

H_0 ditolak apabila $Q > X_{(1-\alpha), df=K-m}^2$

Keterangan:

$\hat{\rho}_k$: autokorelasi pada lag ke- k dengan $k=1,2,\dots,K$

df : derajat kebebasan $df=K-m$ dengan $m=p+q$

Prosedur Box-Jenkins

Pemilihan Model Terbaik

Tujuan pemodelan adalah peramalan data untuk periode selanjutnya, maka kriteria berdasarkan kesalahan peramalan pada *out sample*:

- ✓ *Mean Absolute Percentage Error*

$$\text{MAPE} = \left(\frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \right) \times 100\%$$

- ✓ *Mean Square Error*

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2}$$

Keterangan:

n : banyak *outsample*

Y_t : *outsample* ke- t dengan $t=1,2,\dots,n$

\hat{Y}_t : ramalan *outsample* ke- t dengan

$t=1,2,\dots,n$

Model Autoregressive Integrated Moving Average MUSIMAN

Model ARIMA Musiman:

$$\Phi_p(B^s) \phi_p(B)(1-B)^d (1-B^s)^D Z_t = \theta_q(B) \Theta_Q(B^s) a_t$$

Keterangan:

$\phi_p(B)$: Operator AR orde p

$\theta_q(B)$: Operator MA orde q

$(1-B)^d$: operator untuk differencing

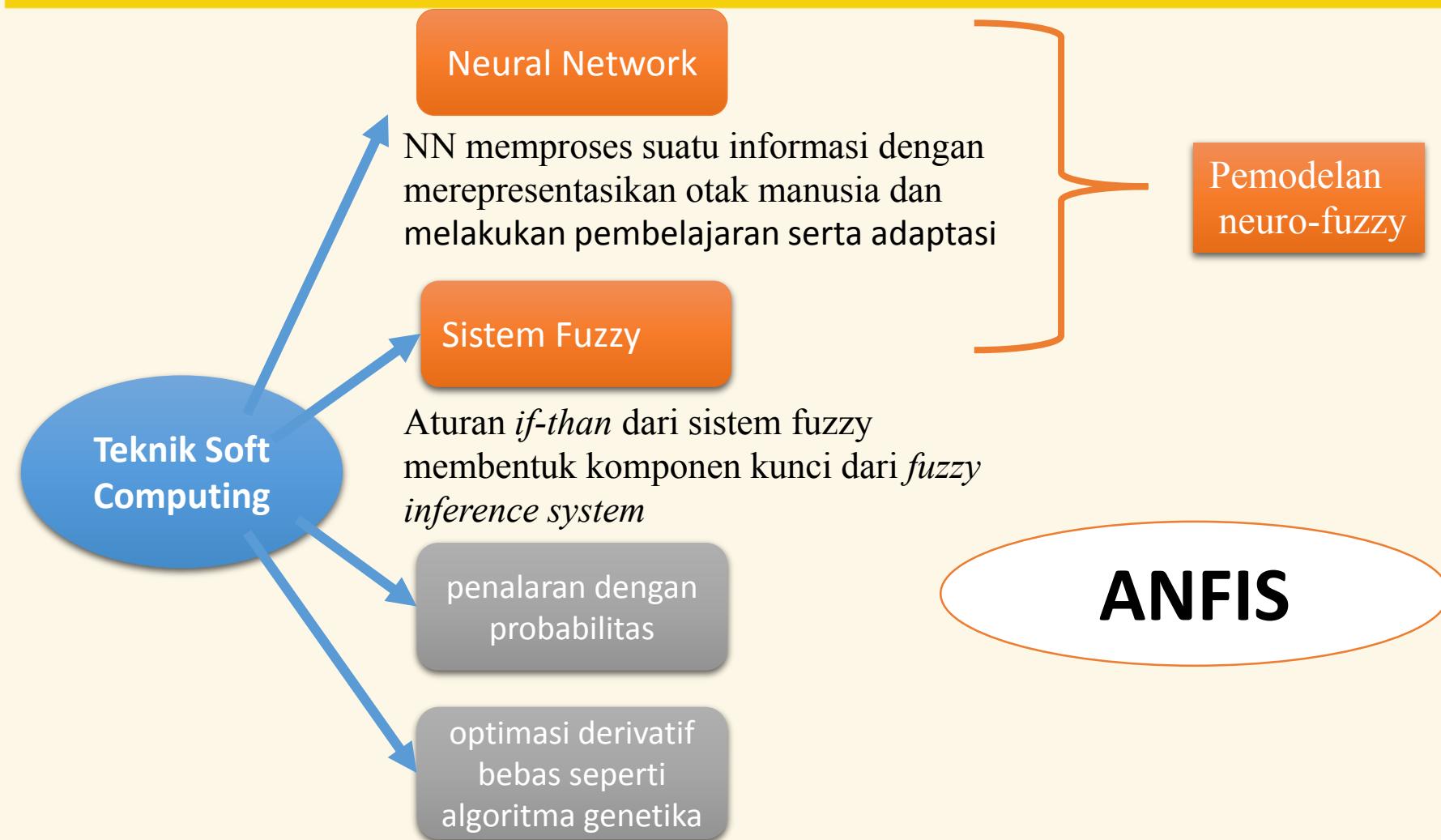
$\Phi_p(B^s)$: Operator AR musiman orde P

$\Theta_Q(B^s)$: Operator MA musiman orde Q

Deteksi Outlier

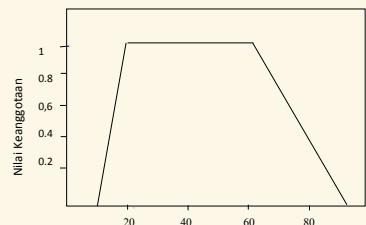
1. Model *Additive Outlier* (AO)
2. Model *Innovational Outlier* (IO)
3. Model *Level Shift* (LS)
4. Model *Temporary Change* (TC)

Teknik Soft Computing



Jang, Sun, & Mizutani, 1997.

Fungsi Keanggotaan Trapezoiedal

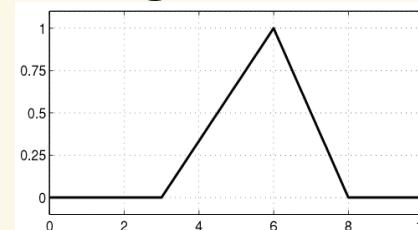


$$\text{trap}(Z; a, b, c, d) = \begin{cases} 0, & Z \leq a \\ (Z-a)/(b-a), & a \leq Z \leq b \\ 1, & b \leq Z \leq c \\ (d-Z)/(d-c), & c \leq Z \leq d \\ 0, & d \leq Z \end{cases}$$

Atau

$$\text{trap}(Z; a, b, c, d) = \max \left(\min \left(\frac{(Z-a)}{(b-a)}, 1, \frac{(d-Z)}{(d-c)} \right), 0 \right)$$

Fungsi Keanggotaan Triangular

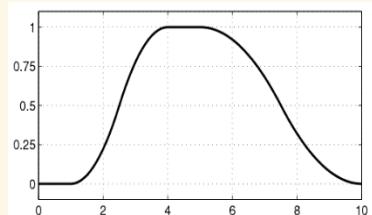


$$\text{tri}(Z; a, b, c) = \begin{cases} 0, & Z \leq a \\ (Z-a)/(b-a), & a \leq Z \leq b \\ (c-Z)/(c-b), & b \leq Z \leq c \\ 0, & c \leq Z \end{cases}$$

Atau

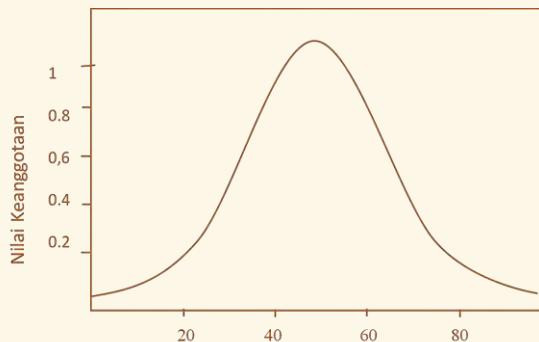
$$\text{tri}(Z; a, b, c) = \max \left(\min \left(\frac{(Z-a)}{(b-a)}, \frac{(c-Z)}{(c-b)} \right), 0 \right)$$

Fungsi Keanggotaan Pi



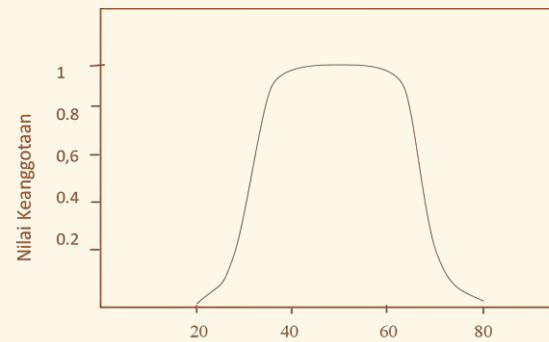
$$\text{pi}(Z; a, b, c, d) = \begin{cases} 0, & Z \leq a \\ 2((Z-a)/(b-a))^2, & a \leq Z \leq (a+b)/2 \\ 1 - 2((Z-b)/(b-a))^2, & (a+b)/2 \leq Z \leq b \\ 1, & b \leq Z \leq c \\ 1 - 2((Z-c)/(d-c))^2, & c \leq Z \leq (c+d)/2 \\ 2((Z-d)/(d-c))^2, & (c+d)/2 \leq Z \leq d \\ 0, & Z \geq d \end{cases}$$

Fungsi Keanggotaan Gaussian



$$\text{gaussian}(x; c, \sigma) = e^{-\frac{1}{2} \left(\frac{x-c}{\sigma} \right)^2}$$

Fungsi Keanggotaan Generalized Bell



$$\text{bell}(x; a, b, c) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}}$$

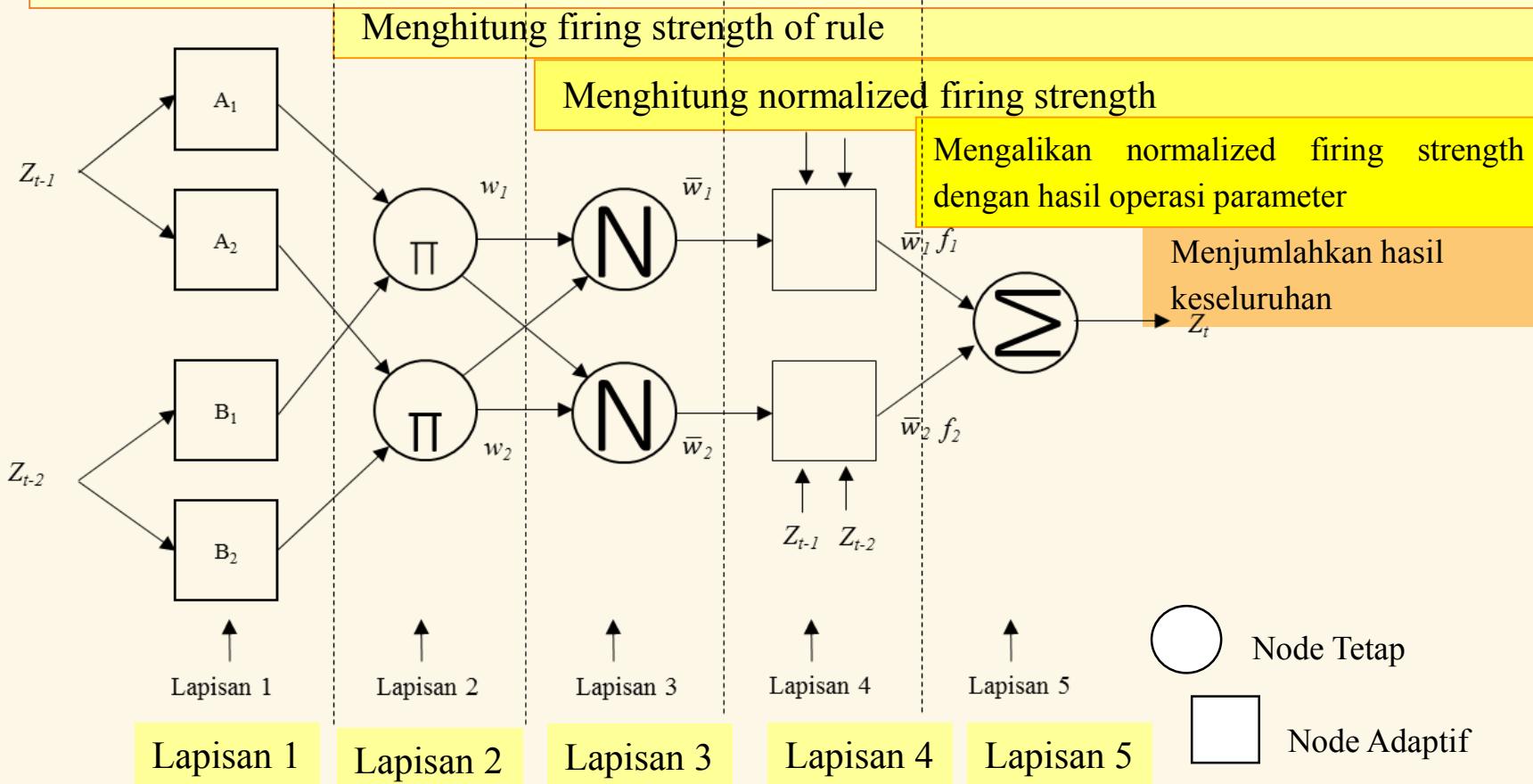
fungsi keanggotaan yang memberikan nilai keanggotaan pada setiap objek dengan rentang nilai antara nol dan satu

Model Adaptive Neuro Fuzzy Inference Systems (ANFIS)

Dua aturan fuzzy *if-then*
Model fuzzy Sugeno

if Z_{t-1} adalah A_1 and Z_{t-2} adalah B_1 , then $f_1 = p_1 Z_{t-1} + q_1 Z_{t-2} + r_1$
if Z_{t-1} adalah A_2 and Z_{t-2} adalah B_2 , then $f_2 = p_2 Z_{t-1} + q_2 Z_{t-2} + r_2$

Menghitung derajat keanggotaan berdasarkan fungsi keanggotaan



(Sugeno & Kang, 1988; Takagi & Sugeno, 1983; Takagi & Sugeno, 1985).

Sumber Data



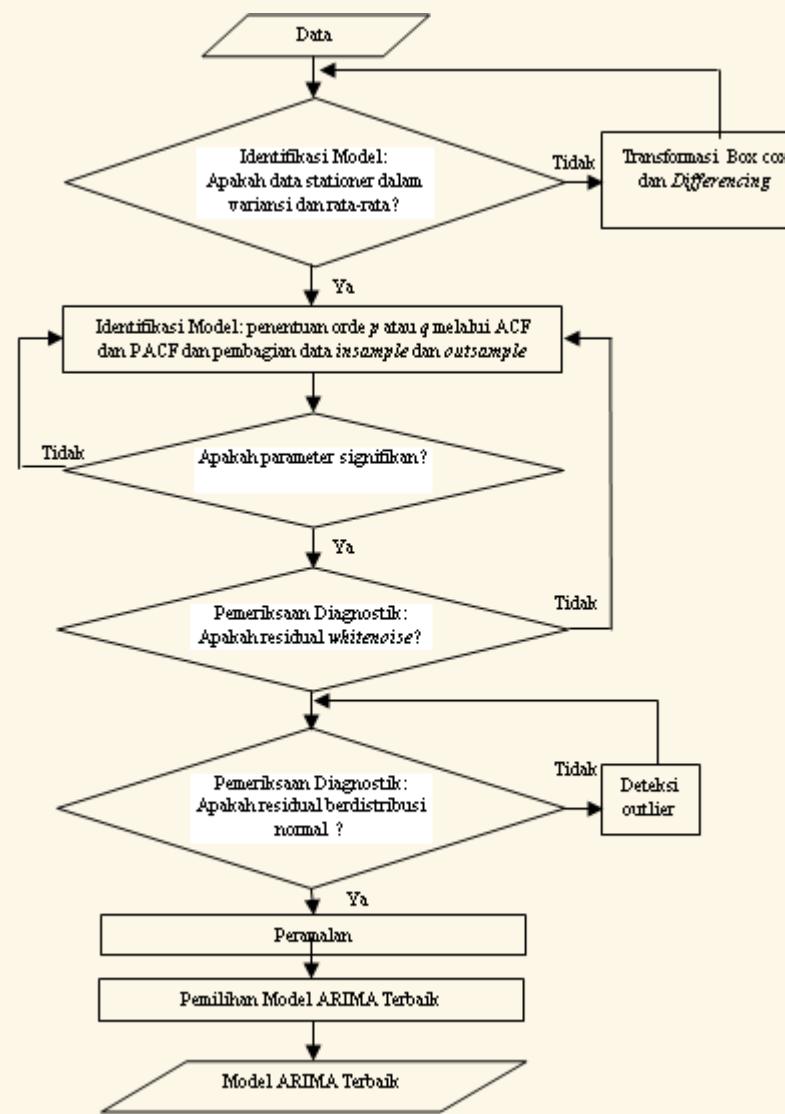
jumlah penumpang Kereta Api Kelas Ekonomi kertajaya mulai 1 Mei 2014 s.d 6 Mei 2015 yang berasal dari PT. KAI DAOP VIII Pasar Turi, Surabaya.

jumlah penumpang harian Kereta Api kertajaya pemberangkatan Surabaya (Stasiun Pasar Turi)

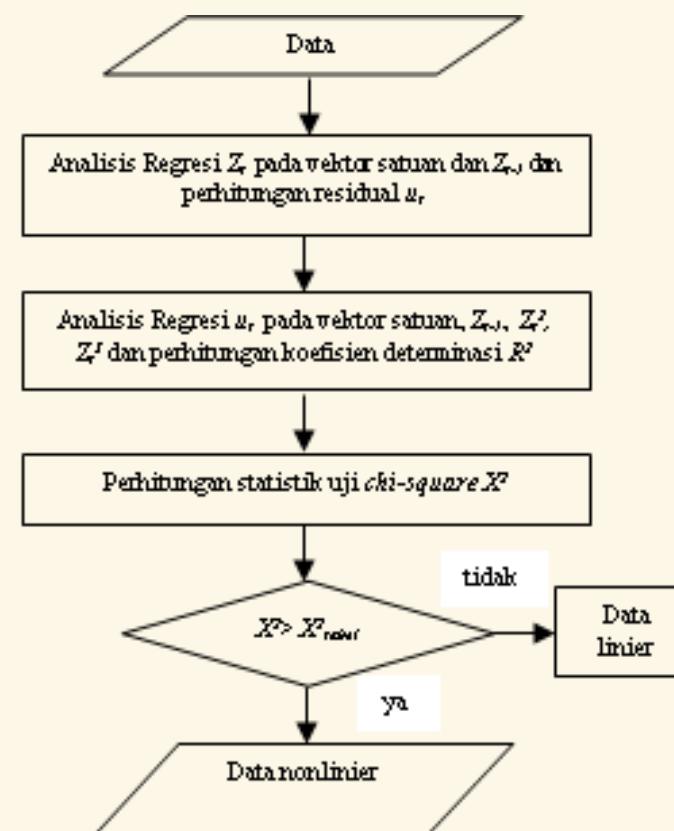
Variabel Penelitian



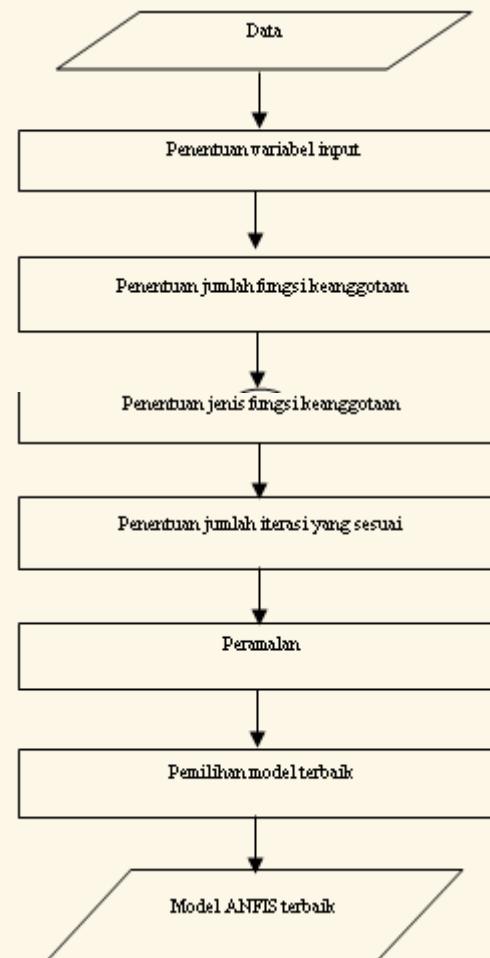
Langkah-langkah Analisis (ARIMA)



Langkah-langkah Uji Linieritas



Langkah-langkah Analisis (ANFIS)



Eksplorasi Data Jumlah Penumpang

Hari	Jumlah Hari (n)	Rata-rata	Variansi	Min	Median	Maks
Senin	53	947	898.44	893	946	1046
Selasa	53	947	1908.60	897	938	1176
Rabu	53	950	1558.87	884	941	1100
Kamis	53	967	1929.03	904	959	1127
Jumat	53	1016	4131.60	922	1002	1242
Sabtu	53	994	1927.05	918	982	1105
Minggu	53	1031	4731.70	892	1026	1172

Statistik Deskriptif

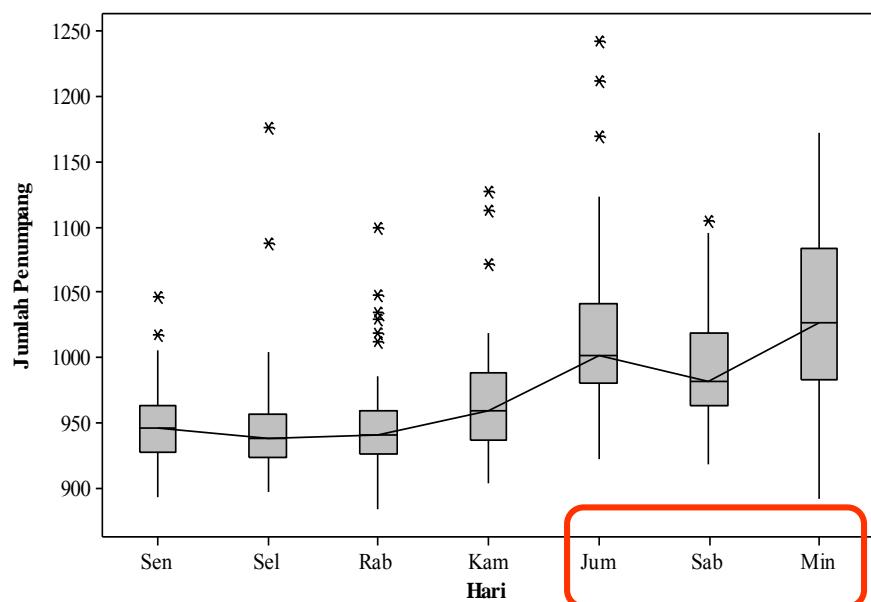
Bulan	Jumlah Hari	Rata-rata	Variansi	Min.	Median	Maks.
Mei 2014	31	1007	3221.30	934	982	1160
Juni 2014	30	966	1411.89	898	963	1063
Juli 2014	31	979	3074.49	884	984	1176
Agustus 2014	31	946	712.42	910	938	1005
September 2014	30	983	4575.50	904	970	1242
Oktober 2014	31	982	3765.20	897	959	1170
November 2014	30	984	2655.91	914	975	1095
Desember 2014	31	995	3540.20	934	968	1105
Januari 2015	31	955	1871.49	893	939	1071
Februari 2015	28	961	1958.44	901	947	1061
Maret 2015	31	1013	5367.00	936	976	1172
April 2015	30	972	5960.70	900	939	1212
Mei 2015	6	965	3315.80	920	939	1061

Variansi Tinggi

Jumlah penumpang beragam

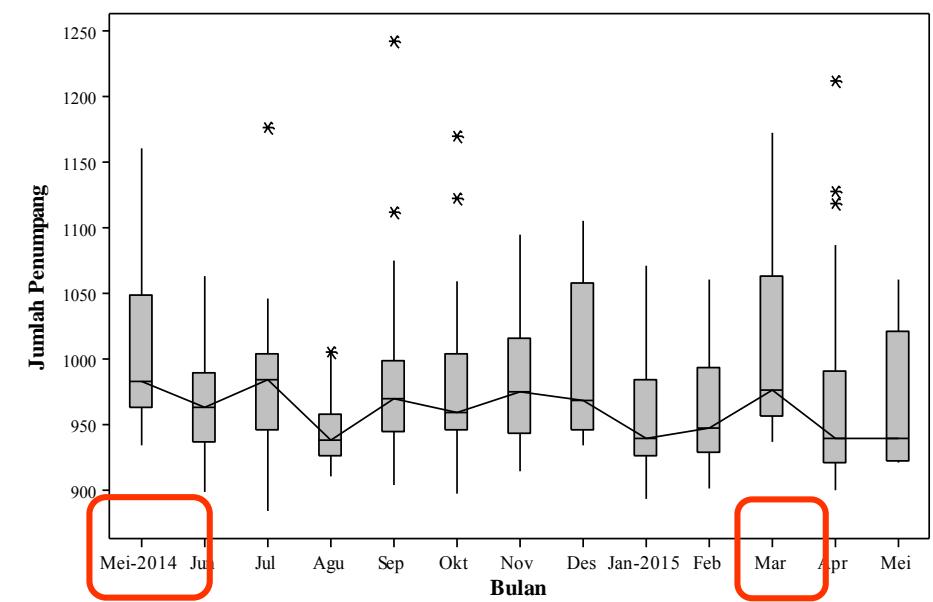
Eksplorasi Data Jumlah Penumpang

Boxplot



Jumat, Sabtu, dan Minggu

Jumlah penumpang cenderung tinggi



Mei 2014 & Maret 2015

Jumlah penumpang cenderung tinggi, periode tersebut merupakan Bulan di mana tariff rendah

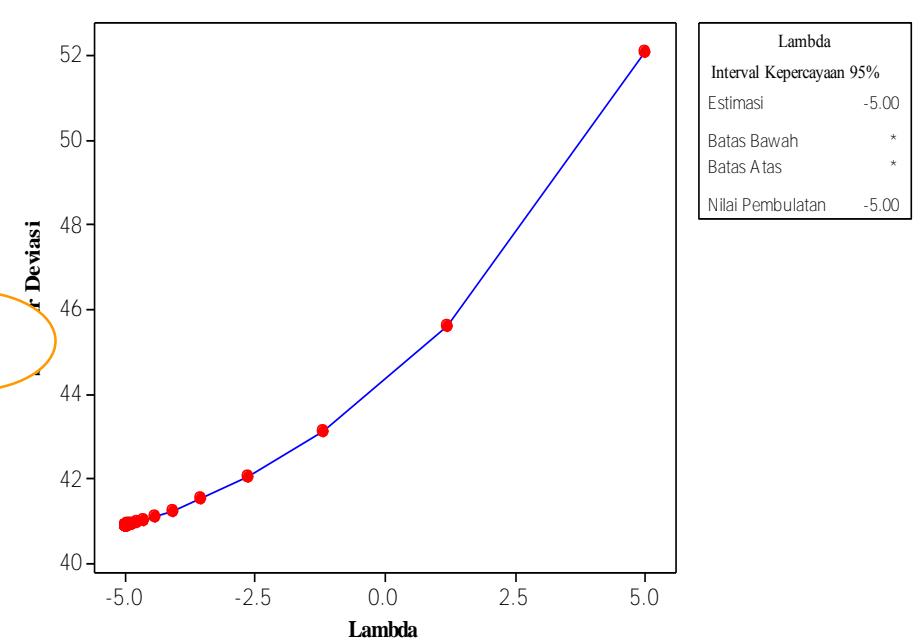
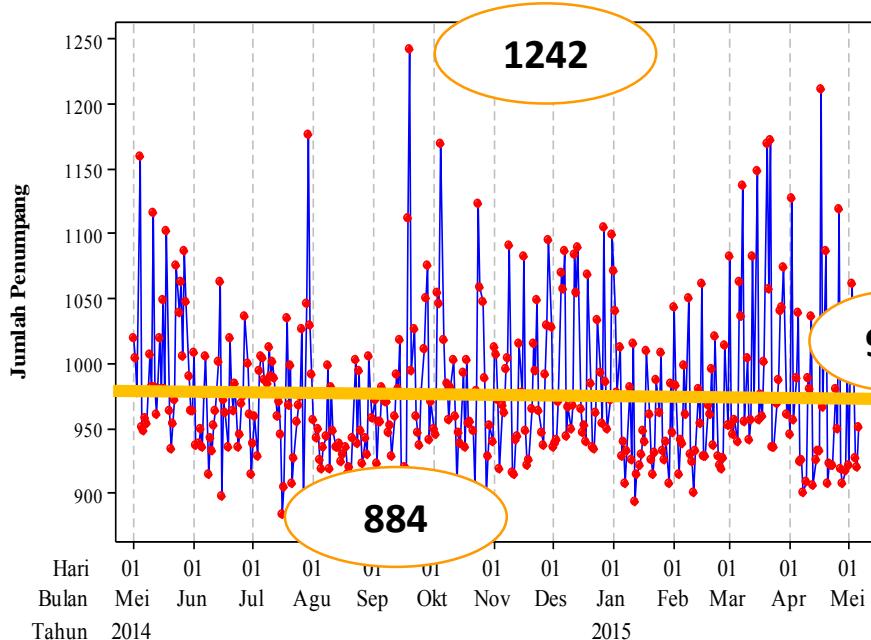


Peramalan Jumlah Penumpang Kereta Api Kertajaya Menggunakan ARIMA

Identifikasi Model Sementara

Visualisasi Plot Time Series

Pemeriksaan Stasioneritas dalam Variansi



In-sample

357 data (1 Mei 2014- 22 April 2015)

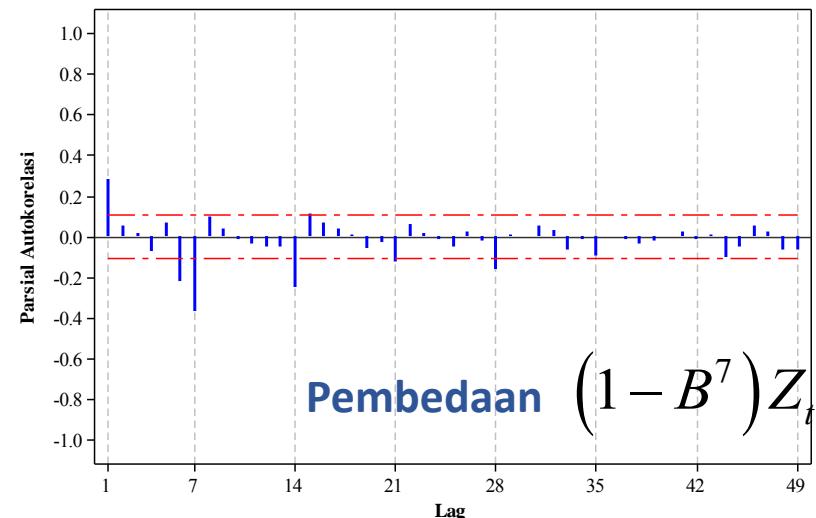
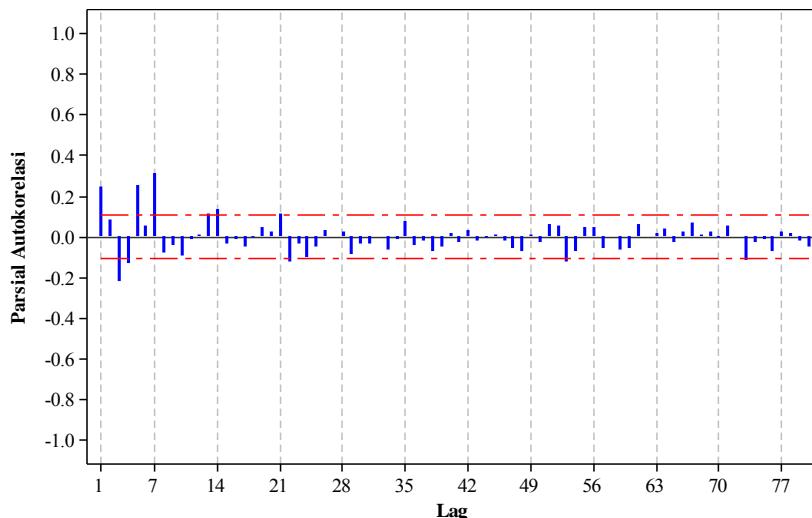
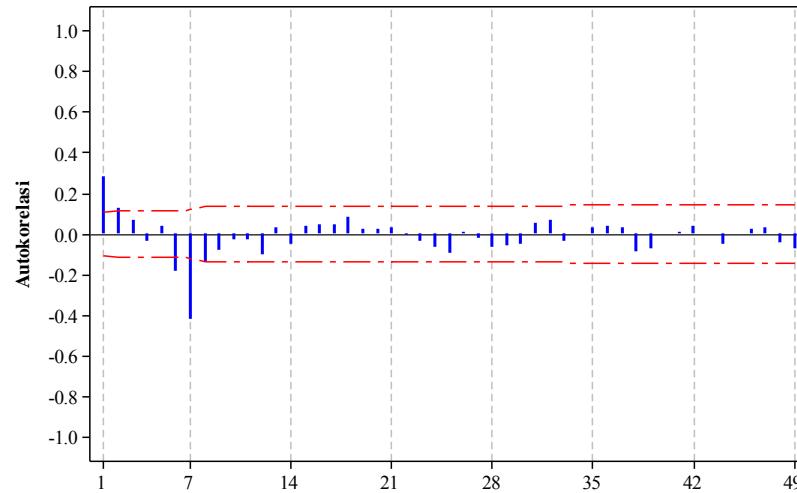
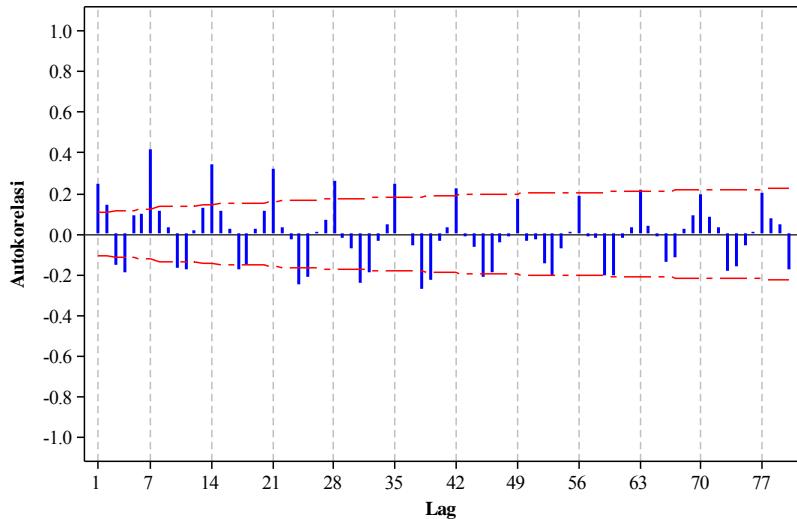
Out-sample

14 data (23 April 2014- 6 Maret 2015)

Data Tidak Stasioner dalam variansi,
tidak dilakukan transformasi

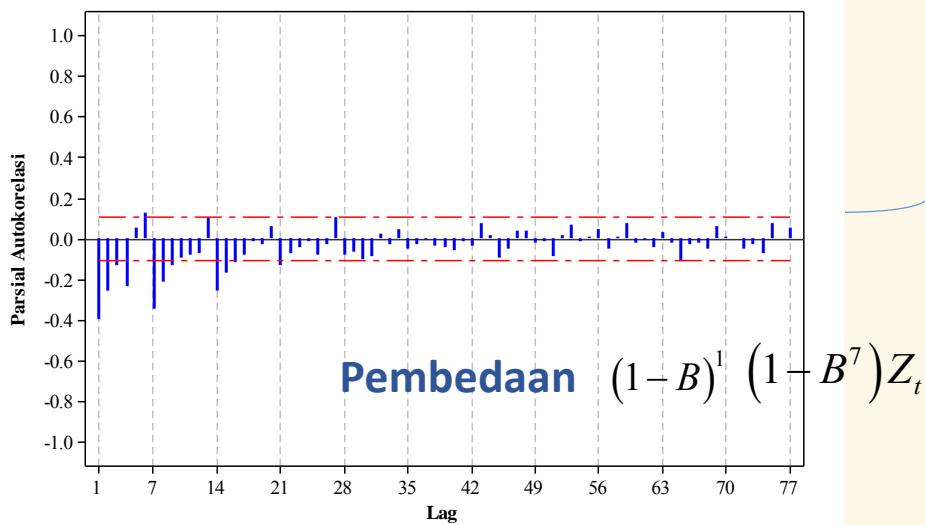
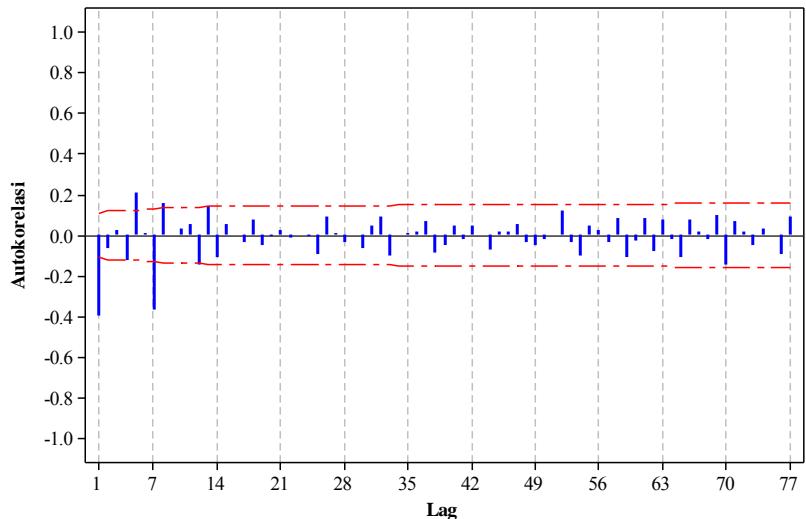
Identifikasi Model Sementara

Pemeriksaan Stasioneritas dalam Rata-rata





Identifikasi Model Sementara



ARIMA $([1,2,3,6],1,1) (0,1,1)^7$.
ARIMA $([4,6],1,1) (0,1,1)^7$
ARIMA $(0,1, [1,4]) (0,1,1)^7$.

Estimasi Parameter

ARIMA ([1,2,3,6],1,1) (0,1,1)⁷.

	Estimasi	Standar Error	t		P-Value
θ_4	0.4694	0.0667	7.03	1.9667	<0.0001
Θ_1	0.7457	0.0401	18.60	1.9667	<0.0001
ϕ_1	-0.6655	0.0510	-13.06	1.9667	<0.0001
ϕ_2	-0.5139	0.0639	-8.05	1.9667	<0.0001
ϕ_3	-0.4514	0.0674	-6.70	1.9667	<0.0001
ϕ_6	-0.0986	0.0479	-2.06	1.9667	0.0404

ARIMA ([4,6],1,1) (0,1,1)⁷

Parameter	Estimasi	Standar Error	t		P-Value
θ_1	0.7281	0.0384	18.96	1.9667	<0.0001
Θ_1	0.7414	0.0371	19.97	1.9667	<0.0001
ϕ_4	-0.1190	0.0546	-2.18	1.9667	0.0298
ϕ_6	-0.1116	0.0554	-2.01	1.9667	0.0448

ARIMA (0,1, [1,4]) (0,1,1)⁷.

Parameter	Estimasi	Standar Error	t		P-Value
θ_1	0.7285	0.0387	18.81	1.9667	<0.0001
θ_4	0.0811	0.0390	2.08	1.9667	0.0382
Θ_1	0.7307	0.0386	18.94	1.9667	<0.0001

Parameter Signifikan dalam mode

Pemeriksaan Diagnostik

ARIMA ([1,2,3,6],1,1) (0,1,1)⁷.

Sampai Lag ke-	X^2_{hitung}	db	$X^2_{(1-0.05),db}$	P-Value
6	.	0		.
12	3.00	6	12.59	0.8087
18	7.68	12	21.03	0.8094
24	9.53	18	28.87	0.9462
30	12.71	24	36.42	0.9707
36	17.34	30	43.77	0.9683
42	19.16	36	51.00	0.9904
48	21.46	42	58.12	0.9964

ARIMA (0,1, [1,4]) (0,1,1)⁷.

Sampai Lag ke-	X^2_{hitung}	db	$X^2_{(1-0.05),db}$	P-Value
6 _{D_(1-0.05;n>40)}	7.54	3	7.81	0.0566
12	9.41	9	16.92	0.4000
18	14.5	15	25.00	0.4877
24	16.96	21	32.67	0.7133
30	21.47	27	40.11	0.7639
36	25.6	33	47.40	0.8175
42	27.96	39	54.57	0.9058
48	31.34	45	61.66	0.9389

ARIMA ([4,6],1,1) (0,1,1)⁷

Sampai Lag ke-	X^2_{hitung}	db	$X^2_{(1-0.05),db}$	P-Value
6	3.96	2	5.99	0.1379
12	8.02	8	15.51	0.4311
18	11.83	14	23.68	0.6201
24	14.69	20	31.41	0.7941
30	19.1	26	38.89	0.8323
36	23.27	32	46.19	0.8697
42	25.24	38	53.38	0.9442
48	29.17	44	60.48	0.9583

Diagnostik Distribusi Normal

Model ARIMA	D	$D_{(1-0.05;n>40)}$	P-Value
([1,2,3,6],1,1) (0,1,1) ⁷	0.0892	0.0728	<0.01
([4,6],1,1) (0,1,1) ⁷	0.0942	0.0728	<0.01
(0,1, [1,4]) (0,1,1) ⁷	0.0939	0.0728	<0.01

Residua white noise namun tidak berdistribusi normal



Deteksi Outlier

ARIMA ([1,2,3,6],1,1) (0,1,1)⁷.

Par- meter	Estimasi	Standar Error	T	P-Value	Tipe Outlier
θ_4	0.4263	0.0663	6.43	<0.0001	-
θ_6	0.1396	0.0527	2.65	0.0085	-
Θ_1	0.6813	0.0433	15.72	<0.0001	-
ϕ_1	-0.6151	0.0545	-11.29	<0.0001	-
ϕ_2	-0.4800	0.0641	-7.49	<0.0001	-
ϕ_3	-0.4224	0.0673	-6.27	<0.0001	-
ω_{142}	197.8149	40.3184	4.91	<0.0001	Additive
ω_{172}	-95.0264	39.5052	-2.41	0.0167	Additive
ω_{338}	-146.2274	40.5297	-3.61	0.0004	Additive
ω_{351}	-89.5601	43.5885	-2.05	0.0407	Additive

ARIMA ([4,6],1,1) (0,1,1)⁷

Par- meter	Estimasi	Standar Error	t	P-value	Outlier
θ_1	0.8101	0.0482	16.80	<.0001	-
Θ_1	0.6802	0.0428	15.91	<.0001	-
ϕ_1	0.2205	0.0746	2.96	0.0033	-
ϕ_6	-0.1525	0.0571	-2.67	0.0079	-
ω_{142}	199.7619	39.3498	5.08	<.0001	Additive
ω_{338}	-163.2736	39.9802	-4.08	<.0001	Additive
ω_{172}	-86.2940	38.6318	-2.23	0.0262	Additive
ω_{351}	-93.5305	42.6860	-2.19	0.0291	Additive
ω_{38}	-161.2107	38.9028	-4.14	<.0001	Additive
ω_{91}	-77.6835	32.9762	-2.36	0.0191	Level Shift

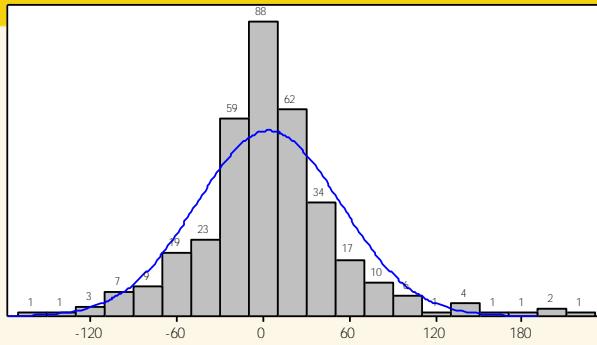
ARIMA (0,1, [1,4]) (0,1,1)⁷.

Par- meter	Estimasi	Standar Error	t	P-value	Outlier
θ_1	0.7190	0.0398	18.06	<0.0001	-
θ_4	0.0886	0.0395	2.24	0.0258	-
Θ_1	0.7065	0.0403	17.54	<0.0001	-
ω_{142}	220.1660	41.9212	5.25	<0.0001	Additive
ω_{338}	-141.2870	42.7856	-3.30	0.0011	Additive

Residual tetap tidak berdistribusi normal

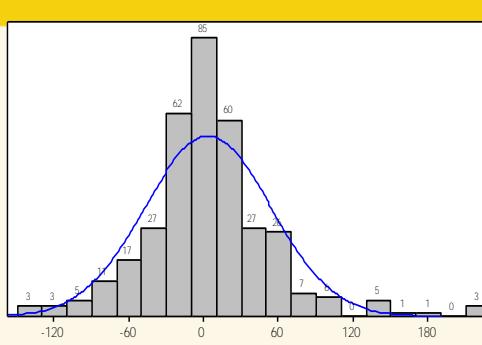
Ringkasan Grafis pada Residual Model ARIMA:

(a); (b); (c) $(0,1,[1,4])$ $(0,1,1)^7$



Anderson-Darling Normality Test	
A-Squared	5.27
P-Value <	0.005
Mean	3.634
StDev	50.072
Variance	2507.223
Skewness	0.61734
Kurtosis	2.95875
N	349
Minimum	-153.320
1st Quartile	-19.545
Median	1.150
3rd Quartile	24.865
Maximum	219.160

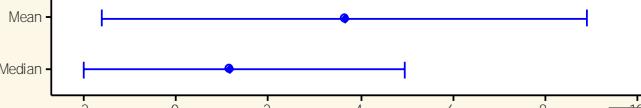
95% Confidence Interval for Mean: -1.637 to 8.906
95% Confidence Interval for Median: -2.031 to 4.921
95% Confidence Interval for StDev: 46.613 to 54.091



Anderson-Darling Normality Test	
A-Squared	5.07
P-Value <	0.005
Mean	3.161
StDev	50.499
Variance	2550.199
Skewness	0.71460
Kurtosis	3.32094
N	349
Minimum	-147.750
1st Quartile	-21.980
Median	2.450
3rd Quartile	25.020
Maximum	226.670

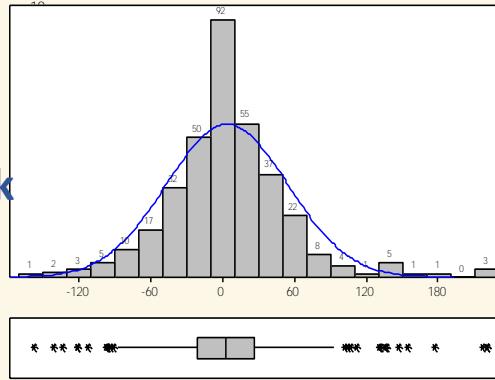
95% Confidence Interval for Mean: -2.156 to 8.478
95% Confidence Interval for Median: -3.211 to 6.629
95% Confidence Interval for StDev: 47.010 to 54.552

95% Confidence Intervals



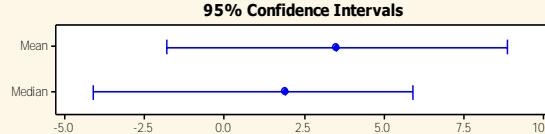
$([1,2,3,6],1,1)$ $(0,1,1)^7$

Kurva Leptokurtik: puncak lebih runcing daripada puncak kurva normal



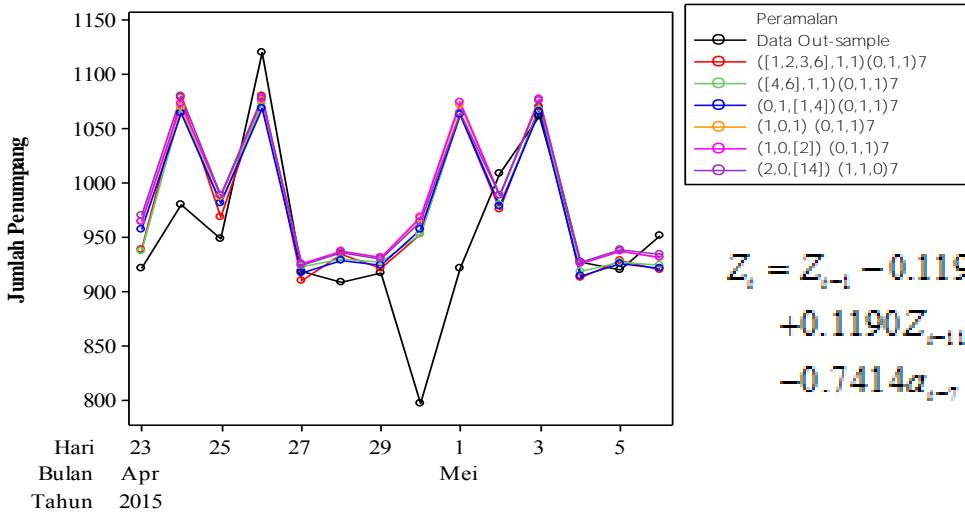
$([4,6],1,1)$ $(0,1,1)^7$

Residual tersebar banyak disekitar nol,



$[1,4])$ $(0,1,1)^7$

Pemilihan Model ARIMA Terbaik



ARIMA $([4,6],1,1)(0,1,1)^7$

$$Z_t = Z_{t-1} - 0.1190Z_{t-4} + 0.1190Z_{t-5} - 0.1116Z_{t-6} + 1.1116Z_{t-7} - Z_{t-8} \\ + 0.1190Z_{t-11} - 0.1190Z_{t-12} + 0.1116Z_{t-13} - 0.1116Z_{t-14} - 0.7281a_{t-4} \\ - 0.7414a_{t-7} + 0.5398a_{t-8} + a_t$$

Model ARIMA	Kriteria In-sample		Kriteria Out-sample
	RMSE	RMSE	MAPE(%)
([1,2,3,6],1,1)(0,1,1)7	50.1323	65.0723	4.7277
([4,6],1,1)(0,1,1) ⁷	50.5261	64.2621	4.6681
(0,1,[1,4])(0,1,1) ⁷	50.8103	65.3818	4.8032
(1,0,1) (0,1,1) ⁷	48.7163	69.1343	5.1474
(1,0,[2]) (0,1,1) ⁷	48.7381	69.7480	5.2009
(2,0,[14]) (1,1,0) ⁷	48.5756	68.5717	5.1297

Peramalan Jumlah Penumpang Kereta Api Kertajaya Menggunakan ANFIS

Uji Linieritas

Uji	X^2_{hitung}	db	$X^2_{(1-0.05),db}$	p-value
Terasvirta	10.0616	2	5.9915	0.006534
White	11.4871	2	5.9915	0.003203

nonlinier

Kombinasi Input untuk ANFIS

Kode Input	Kombinasi Variabel Input
1	$x_{i-8}, x_{i-7}, x_{i-1}$
2	x_{i-7}, x_{i-1}
3	x_{i-8}, x_{i-7}
4	$x_{i-7}, x_{i-4}, x_{i-1}$
5	$x_{i-7}, x_{i-6}, x_{i-1}$
6	x_{i-7}, x_{i-4}
7	x_{i-7}, x_{i-6}

Jenis Fungsi Keanggotaan: Gaussian, generalized bell, trapezoidal, triangular, dan pi

Jumlah Fungsi Keanggotaan : 2

Epoch: 500



Pemodelan ANFIS dengan Fungsi Keanggotaan Bentuk Kurva Pi

Input Data In-sample

Data ke-	Z_{t-3}	Z_{t-2}	Z_{t-1}	Target Output (Z_t)
1	1019	1004	954	1007
2	1004	977	1007	982
3	977	1160	982	1116
:	:	:	:	:
348	906	925	907	923
349	925	933	923	922

Parameter nonlinier pada Lapisan ke-1

Input	a	b	c	d
Input1 mf1 (A1)	633.4	776.6	1044.0	1045.0
Input1 mf2 (A2)	997.6	1150.0	1349.0	1493.0
Input2 mf1 (B1)	633.4	776.6	1076.0	1088.0
Input2 mf2 (B2)	985.0	1179.0	1349.0	1493.0
Input3 mf1 (C1)	633.4	776.6	938.6	1130.0
Input3 mf2 (C2)	1038.0	1039.0	1349.0	1493.0

Pemodelan ANFIS dengan Fungsi Keanggotaan Bentuk Kurva Pi

Output pada Lapisan ke-1

Data ke-	Derajat Keanggotaan					
	μ_{A1}	μ_{A2}	μ_{B1}	μ_{B2}	μ_{C1}	μ_{C2}
1	1.0000	0.0394	1.0000	0.0192	0.9871	0.0000
2	1.0000	0.0035	1.0000	0.0000	0.7446	0.0000
3	1.0000	0.0000	0.0000	0.9808	0.8972	0.0000
:	:	:	:	:	:	:
348	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
349	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000

Output pada Lapisan ke-2

Data ke-	Firing Strength			
	$w_{1,t}$	$w_{2,t}$	$w_{3,t}$	$w_{4,t}$
1	0.9871	0.0000	0.0189	0.0000
2	0.7446	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.8800	0.0000
:	:	:	:	:
348	1.0000	0.0000	0.0000	0.0000
349	1.0000	0.0000	0.0000	0.0000

Data ke-	Firing Strength			
	$w_{5,t}$	$w_{6,t}$	$w_{7,t}$	$w_{8,t}$
1	0.0389	0.0000	0.0007	0.0000
2	0.0026	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.0000	0.0000
:	:	:	:	:
348	0.0000	0.0000	0.0000	0.0000
349	0.0000	0.0000	0.0000	0.0000

Delapan Aturan if-then

if Z_{t-8} is A1 and Z_{t-7} is B1 and Z_{t-1} is C1, then $w_{1,t} = \mu_{A1} \mu_{B1} \mu_{C1}$

if Z_{t-8} is A1 and Z_{t-7} is B1 and Z_{t-1} is C2, then $w_{2,t} = \mu_{A1} \mu_{B1} \mu_{C2}$

if Z_{t-8} is A1 and Z_{t-7} is B2 and Z_{t-1} is C1, then $w_{3,t} = \mu_{A1} \mu_{B2} \mu_{C1}$

if Z_{t-8} is A1 and Z_{t-7} is B2 and Z_{t-1} is C2, then $w_{4,t} = \mu_{A1} \mu_{B2} \mu_{C2}$

if Z_{t-8} is A2 and Z_{t-7} is B1 and Z_{t-1} is C1, then $w_{5,t} = \mu_{A2} \mu_{B1} \mu_{C1}$

if Z_{t-8} is A2 and Z_{t-7} is B1 and Z_{t-1} is C2, then $w_{6,t} = \mu_{A2} \mu_{B1} \mu_{C2}$

if Z_{t-8} is A2 and Z_{t-7} is B2 and Z_{t-1} is C1, then $w_{7,t} = \mu_{A2} \mu_{B2} \mu_{C1}$

if Z_{t-8} is A2 and Z_{t-7} is B2 and Z_{t-1} is C2, then $w_{8,t} = \mu_{A2} \mu_{B2} \mu_{C2}$

Pemodelan ANFIS dengan Fungsi Keanggotaan Bentuk Kurva Pi

Output pada Lapisan ke-3

Output pada Lapisan ke-4

Data ke-	Normalized Firing Strength				Data ke-	$\bar{w}_{1,t}y_1$	$\bar{w}_{2,t}y_2$	$\bar{w}_{3,t}y_3$	$\bar{w}_{4,t}y_4$
	$\bar{w}_{1,t}$	$\bar{w}_{2,t}$	$\bar{w}_{3,t}$	$\bar{w}_{4,t}$		$\bar{w}_{1,t}y_1$	$\bar{w}_{2,t}y_2$	$\bar{w}_{3,t}y_3$	$\bar{w}_{4,t}y_4$
1	0.9440	0.0000	0.0181	0.0000	1	937.7224	0.0000	18.5861	0.0000
2	0.9965	0.0000	0.0000	0.0000	2	979.1336	0.0000	0.0000	0.0000
3	0.0000	0.0000	1.0000	0.0000	3	0.0000	0.0000	1080.7812	0.0000
:	:	:	:	:	:	:	:	:	:
348	1.0000	0.0000	0.0000	0.0000	348	940.4869	0.0000	0.0000	0.0000
349	0.1232	0.0464	0.4145	0.1561	349	947.1562	0.0000	0.0000	0.0000
Data ke-									
$\bar{w}_{5,t}y_5$									
1	0.0372	0.0000	0.0007	0.0000	1	36.0664	0.0000	0.6154	0.0000
2	0.0035	0.0000	0.0000	0.0000	2	3.3885	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.0000	0.0000	3	0.0000	0.0000	0.0000	0.0000
:	:	:	:	:	:	:	:	:	:
348	0.0000	0.0000	0.0000	0.0000	348	0.0000	0.0000	0.0000	0.0000
349	0.0000	0.0000	0.0000	0.0000	349	0.0000	0.0000	0.0000	0.0000

Pemodelan ANFIS dengan Fungsi Keanggotaan Bentuk Kurva Pi

Output pada Lapisan ke-5

$$\begin{aligned}
 \hat{Z}_t &= \bar{w}_{1,t} f_1 + \bar{w}_{2,t} f_2 + \bar{w}_{3,t} f_3 + \bar{w}_{4,t} f_4 + \bar{w}_{5,t} f_5 + \bar{w}_{6,t} f_6 + \bar{w}_{7,t} f_7 + \bar{w}_{8,t} f_8 \\
 &= \bar{w}_{1,t} (p_1 Z_{t-2} + q_1 Z_{t-1} + r_1 Z_{t-1} + s_1) + \bar{w}_{2,t} (p_2 Z_{t-2} + q_2 Z_{t-1} + r_2 Z_{t-1} + s_2) \\
 &\quad + \bar{w}_{3,t} (p_3 Z_{t-2} + q_3 Z_{t-1} + r_3 Z_{t-1} + s_3) + \bar{w}_{4,t} (p_4 Z_{t-2} + q_4 Z_{t-1} + r_4 Z_{t-1} + s_4) \\
 &\quad + \bar{w}_{5,t} (p_5 Z_{t-2} + q_5 Z_{t-1} + r_5 Z_{t-1} + s_5) + \bar{w}_{6,t} (p_6 Z_{t-2} + q_6 Z_{t-1} + r_6 Z_{t-1} + s_6) \\
 &\quad + \bar{w}_{7,t} (p_7 Z_{t-2} + q_7 Z_{t-1} + r_7 Z_{t-1} + s_7) + \bar{w}_{8,t} (p_8 Z_{t-2} + q_8 Z_{t-1} + r_8 Z_{t-1} + s_8)
 \end{aligned}$$

t	Gaussian	Generalizedbell	Trapezoidal	Triangular	Pi
1	977	978	984	989	993
2	977	977	975	986	983
3	1062	1041	1043	1076	1081
:	:	:	:	:	:
348	942	937	943	924	940
349	945	945	948	942	947

Pemilihan Model ANFIS terbaik

Perbandingan RMSE Kebaikan Model untuk *In-sample*

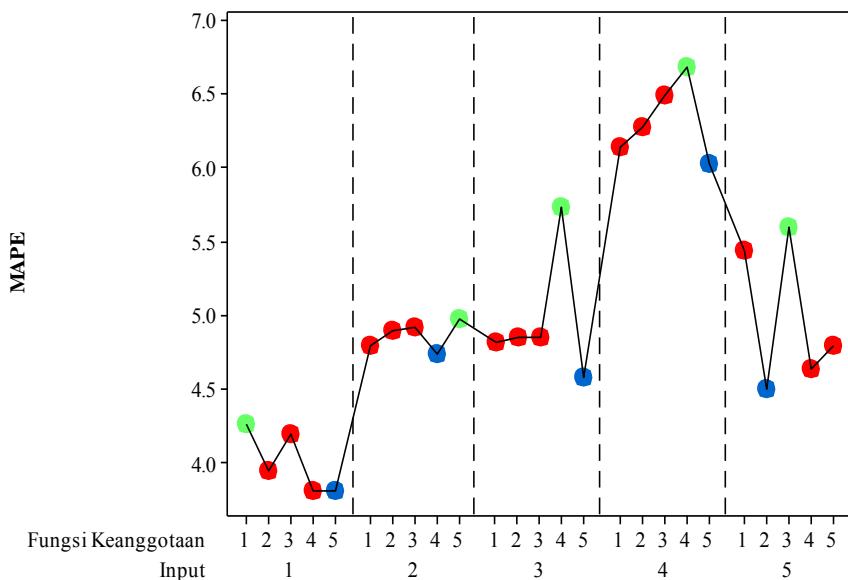
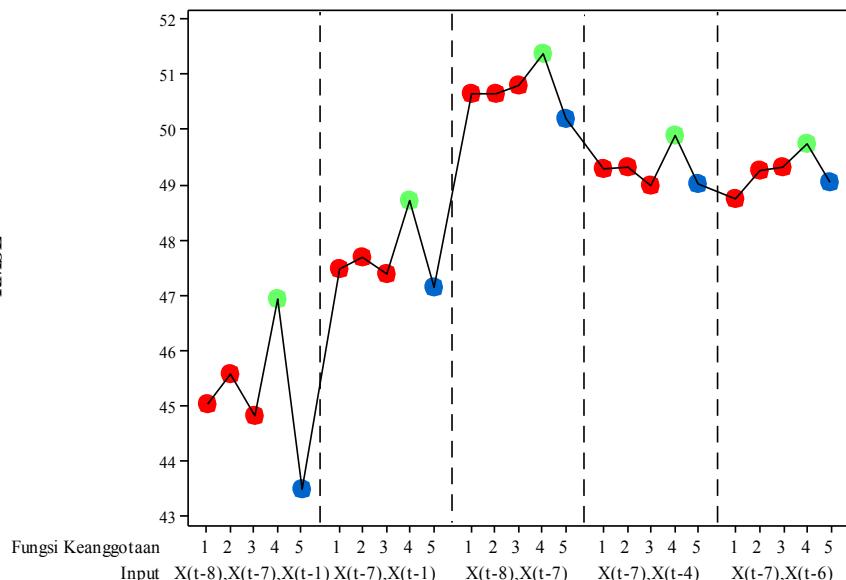
Jenis Input	Gauss	Gbell	Trap	Tri	Pi
$X_{t-8}, X_{t-7}, X_{t-1}$	45.057	45.592	44.838	46.933	43.492
X_{t-7}, X_{t-1}	47.496	47.696	47.409	48.706	47.150
X_{t-8}, X_{t-7}	50.648	50.643	50.798	51.368	50.192
X_{t-7}, X_{t-4}	49.280	49.312	49.001	49.909	49.008
X_{t-7}, X_{t-6}	48.741	49.255	49.321	49.756	49.059

Perbandingan RMSE Kebaikan Model untuk *Out-sample*

Jenis Input	kriteria	Gauss	Gbell	Trap	Tri	Pi
$x_{t-8}, x_{t-7}, x_{t-1}$	MAPE	4.27	3.94	4.19	3.81	3.81
	RMSE	57.04	52.18	53.64	51.90	51.30
x_{t-7}, x_{t-1}	MAPE	4.79	4.89	4.91	4.74	4.98
	RMSE	59.06	58.77	59.53	58.46	60.21
x_{t-8}, x_{t-7}	MAPE	4.81	4.85	4.85	5.74	4.58
	RMSE	58.35	58.04	58.59	66.21	57.80
x_{t-7}, x_{t-4}	MAPE	6.14	6.28	6.49	6.68	6.03
	RMSE	75.60	78.67	82.22	88.59	76.74
x_{t-7}, x_{t-6}	MAPE	5.44	4.50	5.60	4.64	4.80
	RMSE	65.98	56.89	67.78	59.62	61.49

Plot Individual RMSE In-sample dan MAPE Out-sample

Fungsi Keanggotaan: (1) Gauss, (2) Gbell, (3) Trap,(4) Tri, (5) Pi



- **Tertinggi**
- **Terendah**

Secara umum, pemodelan ANFIS dengan *input* $x_{t-8}, x_{t-7}, x_{t-1}$ memberikan MAPE Outsample dan RMSE In-sample yang cenderung kecil

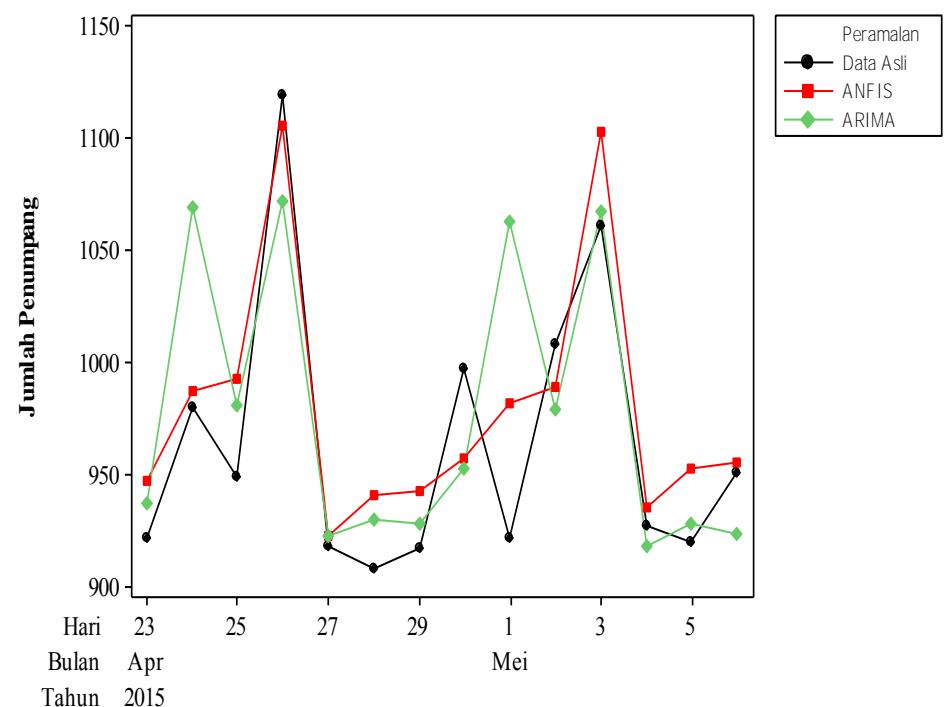
Fungsi Keanggotaan Pi cenderung memberikan RMSE Insample terkecil , namun pada MAPE Out-sample fungsi keanggotaan Pi, Triangular, dan Generalized Bell cenderung memberi MAPE terkecil

Perbandingan Pemodelan ARIMA dan ANFIS

Model ARIMA terbaik yaitu ARIMA $([4,6],1,1)(0,1,1)^7$ d

Model ANFIS terbaik yaitu dengan jenis *input* $x_{t-8}, x_{t-7}, x_{t-1}$, dua fungsi keanggotaan, dan jenis fungsi keanggotaan P_i .

Model	RMSE	MAPE
ANFIS	51.30	3.81 %
ARIMA	64.63	4.67 %



Hasil Peramalan

Periode	Hari	Data Asli	ANFIS (Pi)	ARIMA ([4,6],1,1) (0,1,1) ⁷
23-Apr-15	Kamis	922	947	937
5 Mei 2015	Jumat	980	987	1069
5 Mei 2015	Sabtu	949	993	981
5 Mei 2015	Minggu	1119	1106	1072
5 Mei 2015	Senin	918	922	923
5 Mei 2015	Selasa	908	941	930
5 Mei 2015	Rabu	917	943	928
5 Mei 2015	Kamis	997	957	953
1 Mei 2015	Jumat	922	982	1063
2 Mei 2015	Sabtu	1008	989	979
3 Mei 2015	Minggu	1061	1102	1067
4 Mei 2015	Senin	927	935	918
5 Mei 2015	Selasa	920	953	928
6 Mei 2015	Rabu	951	956	924

KESIMPULAN DAN SARAN

1. Jumlah penumpang kereta api Kertajaya mencapai 1242 orang perhari dan paling sedikit terdapat 884 orang perhari. Adapun jumlah penumpang memiliki pola yang sama dan berulang pada satu minggu atau dapat disebut sebagai musiman tujuh. Model ARIMA ([4,6],1,1) (0,1,1)⁷ memiliki MAPE dan RMSE yang lebih kecil yaitu 4.67 % dan 64.26 dibandingkan lima model ARIMA yang lain.
2. Model ANFIS dengan $input Z_{t-8}, Z_{t-7}, Z_{t-1}$ dan fungsi ke-anggotaan bentuk kurva P_i memiliki akurasi yang lebih baik dibandingkan model ANFIS yang lainnya ya-itu dengan MAPE 3.81% dan RMSE 51.30.
3. Perbandingan MAPE dan RMSE menunjukkan bah-wa model ANFIS memiliki nilai yang lebih kecil da-ripada model ARIMA hal tersebut dikarenakan data jumlah penumpang merupakan data yang nonlinier yang sulit apabila dianalisis menggunakan pendekatan ARIMA Box-Jenkins yang merupakan pende-katan linier. Berdasarkan model terbaik yaitu ANFIS, jumlah penumpang pada waktu ke- t dipengaruhi oleh jumlah penumpang satu, tujuh, dan delapan hari sebelumnya.

Daftar Pustaka

- Atsalakis, G., Nezis, D., & Matalliotakis, G. (2008). Forecasting Mortality Rate Using a Neural Network with Fuzzy Inference System. *International Workshop on Applied Probability* (p. 806). Rethymnon: Department of Economics, University of Crete.
- Avci, E. (2008). Comparison of Wavelet Families for Texture Classification by Using Wavelet Packet Entropy Adaptive Network Based Fuzzy Inference System. . *Applied Soft Computing*, 8(1), 225-231.
- BPS. (2014, Oktober 1). Berita Resmi Statistik. *Perkembangan Pariwisata dan Transportasi Nasional Agustus 2014*, pp. 1-10.
- Cryer, J. D., & Chan, K. S. (2008). *Time Series Analysis with Applications in R*. New York: Springer.
- Daniel, W. W. (2000). *Applied Nonparametric Statistics, second edition*. Dallas: Cengage Learning.
- Dickey, D., & Fuller, W. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 427-431.
- Galavi, H., Mirzaei, M., Shui, L. T., & Valizadeh, N. (2013). Klang River-Level Forecasting Using ARIMA and ANFIS. *American Water Works Association Journal*, E496-E506.
- Gyu-jin, J., Jin-pyung, K., Jee-hyong, L., & Moon-hyun, K. (2011). ANFIS based Short-term Prediction of Passenger Volume in Urban Rail Transit. *International Symposium on Advanced Intelligent systems (ISIS)*. USA: Springer.

Daftar Pustaka

- Jang, J., Sun, C., & Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing*. London: Prentice Hall.
- KAI. (2014, Maret 06). *PT. Kereta Api Indonesia*. Retrieved from Pemerintah Subsidi Penumpang KA Kelas Ekonomi : <https://www.kereta-api.co.id>
- Nurvitasari, Y. (2012). Pendekatan Fungsi Transfer sebagai Input Adaptive Neuro-Fuzzy Inference System (ANFIS) dalam peramalan Kecepatan ANgin Rata-rata Harian di Sumenep. *Jurnal Sains dan Seni ITS*, 62-68.
- Padang, E., Tarigan, G., & Sinulingga, U. (2013). Peramalan Jumlah Penumpang Kereta Api Medan-Rantau Prapat dengan Metode Pemulusan Eksponensial Holt-Winters. *Jurnal Saintika Matematika*, 161-174.
- Sugeno, M., & Kang, G. T. (1988). Structure Identification of Fuzzy Model . *Fuzzy Sets and Systems*, 28:15-33.
- Takagi, T., & Sugeno, M. (1983). Derivation of Fuzzy Control Rules from Human Operator's Control Actions. *The IFAC Symposium on Fuzzy Information, Knowledge Representation and Decision Analysis*, (pp. 55-60).
- Takagi, T., & Sugeno, M. (1985). Fuzzy Identification of Systems and Its Applications to Modelling and Control. *IEEE Transactions on Systems, Man, and Cybernetics*, (pp. 15:116-132).

Daftar Pustaka

- Varedi, M. (2011). *Forecasting Seat Sales in Passenger Airlines Using the Reverse Traffic.* German: VDM Verlag.
- Wei, W. (2006). *Time Series Univariate and Multivariate Methods.* Canada: Addison Wesley.
- Widhianti, N., & Wutaqa, D. U. (2013). Peramalan Banyak Penumpang Kereta DAOP VI Yogyakarta Menggunakan Model Time Series dengan Variasi Kalender Islam Regarima. *Seminar Nasional Matematika dan Pendidikan Matematika* (pp. 181-190). Yogyakarta: FMIPA UNY.
- Widiyanto, T. A. (2013). Peramalan Jumlah Penumpang Kereta Api Logawa dan Sritanjung di DAOP IX Jember. *Jurnal Sains dan Seni ITS*, 2337-3520.
- Yayar, R., Hekim, M., Yilmaz, V., & Bakirci, F. (2011). A Comparison of ANFIS and ARIMA Techniquesin the Forecasting of Electric Energy Consumption of Tokat Province in Turkey. *Journal of Economic and Social Studies*, 87-112.
- Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, 8:338-353.
- Zadeh, L. A. (1992). Fuzzy Logic. Neural Networks and Soft Computing. *One-page Course Announcement of CS 294-4, Spring 1993.* Berkeley: The University of California.

Selasa, 30 Juni 2015

PERAMALAN JUMLAH PENUMPANG KERETA API EKONOMI KERTAJAYA MENGGUNAKAN METODE ARIMA DAN ANFIS

Oleh:

Ilafi Andalita (1311100038)

Dosen Pembimbing:
Dr. Irhamah, M.Si.

JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER