

Pemanfaatan Metode *Smoothing Whittaker-Henderson* untuk Meningkatkan Akurasi *Neural Network Forecasting*

Hans Pratyaksa¹, Adhistya Erna Permanasari^{2*}, Silmi Fauziati³

Intisari—Dalam pelayanan kesehatan, institusi kesehatan perlu memastikan ketersediaan obat untuk pasien. Terdapat beberapa tantangan yang berkaitan dengan ketidakpastian jumlah penggunaan obat pada periode yang akan datang. Ketidakpastian dapat dikurangi dengan menganalisis histori data obat untuk memprediksi kebutuhan mendatang. Runtun waktu dapat mengandung *spikes* atau pola fluktuasi yang *spikes*-nya dapat menyamarkan informasi utama. Hal tersebut dapat memengaruhi akurasi model prediksi. Salah satu metode *forecasting* pada data runtun waktu yang banyak digunakan adalah metode *artificial neural network* (ANN). Metode ANN membutuhkan tahap praproses data sebelum dilakukannya proses pelatihan. Tahap praproses penting dilakukan untuk mendapatkan informasi atau pengetahuan. Makalah ini berfokus pada pengaplikasian metode *smoothing* dalam tahap praproses metode ANN. Pengaplikasian metode *smoothing* diharapkan dapat meningkatkan kualitas data pembelajaran ANN yang akan memperbaiki akurasi prediksi. Makalah ini berfokus pada implementasi metode *smoothing* dalam tahap praproses data untuk metode ANN. Metode *smoothing* yang digunakan adalah *exponential smoothing* (ES) dan *smoothing Whittaker-Henderson* (WH) yang diaplikasikan pada dua dataset runtun waktu. Metode *refining* yang digunakan adalah metode WH, yang diuji pada dua data runtun waktu dataset obat-obatan. Hasil menunjukkan bahwa *mean squared error* (MSE) yang diperoleh dengan mengaplikasikan metode WH lebih rendah daripada ANN tanpa *smoothing*, untuk kedua dataset. Hasil evaluasi menunjukkan bahwa implementasi metode *smoothing* WH dalam tahap praproses data untuk ANN (WH+ANN) menghasilkan MSE yang lebih rendah secara signifikan daripada hasil ANN, dengan *confidence level* sebesar 94% untuk dataset 1 dan 85% untuk dataset 2.

Kata Kunci—*Forecasting*, Praproses, Metode *Smoothing*, *Whittaker-Henderson*, *Artificial Neural Network*.

I. PENDAHULUAN

Pemberian obat, sebagai salah satu cara utama pakar kesehatan menangani penyakit, adalah kebutuhan esensial bagi pasien. Oleh karena itu, rumah sakit atau apotek harus memastikan ketersediaan obat yang akan diberikan kepada pasien yang membutuhkan. Ketersediaan obat-obatan merupakan salah satu tantangan bagi institusi kesehatan dalam memberikan pelayanan yang berkualitas. Maka, institusi kesehatan perlu mengendalikan ketersediaan obat dengan baik untuk menghindari efek yang tidak diinginkan [1], [2].

^{1,2,3} Department of Electrical and Information Engineering, Faculty of Engineering, Universitas Gadjah Mada, Jln. Grafika No. 2, Kampus UGM, Yogyakarta, 55281, INDONESIA (phone: 0274-552305; email: ¹*hans_ti10@mail.ugm.ac.id*, ²*adhistya@ugm.ac.id*, ³*silmi@ugm.ac.id*)

*Corresponding Author

[Diterima: 30 November 2021, Revisi: 7 Januari 2022]

Kemampuan memprediksi kebutuhan di masa mendatang penting dalam mengantisipasi ketersediaan sumber daya. *Forecasting* yang akurat memungkinkan pembuat keputusan mengantisipasi kebutuhan pelayanan dan membuat keputusan terkait manajemen pemenuhan sumber daya dan persediaan dari waktu ke waktu [1], [3]-[6]. *Forecasting* dapat dilakukan dengan intuisi atau dengan menganalisis data runtun waktu. Pada *forecasting* berdasarkan analisis data runtun waktu, *estimator* menggunakan pola data runtun waktu yang merepresentasikan jumlah permintaan terhadap runtun waktu.

Dalam penambangan data (*data mining*), banyak penelitian tentang *forecasting* yang telah dilakukan. Berbagai pengembangan metode atau kombinasi beberapa metode telah diusulkan untuk menghasilkan prediksi yang akurat. Salah satu metode untuk memprediksi data runtun waktu adalah metode *neural network* (NN). Penelitian-penelitian tentang peningkatan *predictive accuracy* pada *artificial neural network* (ANN) menawarkan beberapa metode, termasuk kombinasi metode optimisasi, seperti *particle swarm optimization* (PSO) atau algoritme genetika (*genetic algorithm*, GA), dengan metode ANN [7], [8]; atau kombinasi dua metode *forecasting*, seperti kombinasi metode *autoregressive integrated moving average* (ARIMA) dan ANN [9].

Selain pendekatan-pendekatan di atas, penelitian-penelitian terbaru menunjukkan bahwa peningkatan *predictive accuracy* metode ANN dapat dilakukan dengan memfokuskan pada tahap praproses. Pada data runtun waktu, pola-pola yang berfluktuasi sering terjadi, yang hal ini dapat menyamarkan nilai informasi atau fenomena utama, sehingga menurunkan kualitas data. Oleh karena itu, tahap praproses sangat penting untuk mengekstraksi informasi penting yang memengaruhi kinerja akurasi *forecasting* [10]. Salah satu tahap praproses untuk data runtun waktu adalah *data smoothing*. Penggunaan teknik *smoothing* dalam data runtun waktu diharapkan dapat mengekstraksi tren atau pola utama dalam data [11]. *Smoothing* juga dapat digunakan untuk menghilangkan derau (*noise*) [12] atau fluktuasi data [11], tetapi tidak menghilangkan informasi penting pada data [13].

Sejalan dengan pengembangan teknik peningkatan *predictive accuracy* ANN, metode-metode *smoothing* juga berkembang. Dalam [14], disebutkan bahwa salah satu metode *smoothing* yang dapat diandalkan untuk *smoothing* data runtun waktu adalah metode *smoothing Whittaker-Henderson* (WH) karena metode tersebut mempertimbangkan *fidelity* dan *smoothness* dari *smoothing data* [14], [15]. Selain itu, [14] menjelaskan dan mengembangkan sebuah algoritme dari metode WH berdasarkan *generalized cross-validation* (GCV). Pada penelitian tersebut, metode *smoothing* WH yang diusulkan [14] diaplikasikan pada tahap praproses metode ANN untuk memprediksi jumlah obat yang digunakan pasien.

Penelitian tentang penggunaan metode *smoothing* WH penting dilakukan untuk mengetahui metode tersebut dapat

meningkatkan kualitas data untuk meningkatkan akurasi *forecasting* metode ANN atau tidak. Pada makalah ini, dua metode *data smoothing* digunakan untuk memprediksi jumlah penggunaan obat: metode transformasi data metode *smoothing* WH [14] dan metode *exponential smoothing* (ES) [16]. Kemudian, kedua hasil akurasi *forecasting* dibandingkan dan dievaluasi.

Bagian-bagian makalah ini selanjutnya disusun sebagai berikut. Bagian 2 menyajikan penelitian yang terkait. Bagian 3 mendeskripsikan data yang digunakan. Bagian 4 menjelaskan metodologi. Bagian 5 menyajikan skenario pengujian, Bagian 6 memamparkan hasil pengujian, dan Bagian 7 menyimpulkan hasil keseluruhan penelitian.

II. NEURAL NETWORK FORECASTING

Forecasting adalah aktivitas meramalkan kejadian yang akan terjadi di masa mendatang berdasarkan pengetahuan yang diperoleh dari kejadian-kejadian yang telah terjadi. *Forecasting* dapat dilakukan pada data runtun waktu. Data runtun waktu adalah sebuah kumpulan data yang diukur berturut-turut pada interval waktu yang berbeda [17]. Data runtun waktu dapat digunakan untuk berbagai tujuan, salah satunya adalah untuk memprediksi kejadian-kejadian yang akan datang dan untuk memahami fenomena yang akan terjadi [18]. Prediksi data runtun waktu telah digunakan dan diteliti dalam berbagai bidang. Sudah ada berbagai metode untuk memprediksi kejadian di masa mendatang dalam data runtun waktu dari model linear dan nonlinear.

Penelitian-penelitian tentang *forecasting* runtun waktu menyatakan bahwa model-model *forecasting* nonlinear dapat dilakukan. Salah satu model nonlinear untuk *forecasting* runtun waktu menggunakan metode NN [19], [20]. *Machine learning* dianggap dapat memberikan hasil prediksi yang lebih baik. Terdapat banyak pendekatan dalam peningkatan akurasi *forecasting* metode NN.

Pemodelan NN adalah pemodelan dengan sistem *black box* yang sangat bergantung pada data masukan dan keluaran. Oleh karena itu, kualitas dan distribusi set sampel sangat penting dalam menggeneralisasi jaringan [16]. Dengan demikian, tahap praproses memainkan peranan penting dalam pemodelan NN. Tahap praproses adalah langkah esensial untuk mengekstraksi informasi penting sehingga akan memengaruhi kinerja akurasi *forecasting* [10]. Salah satu tahap praproses untuk data runtun waktu adalah *data smoothing*. Penggunaan teknik *smoothing* dalam data runtun waktu diharapkan dapat mengekstraksi tren atau pola utama dalam data [11]. *Smoothing* juga dapat digunakan untuk menghilangkan derau yang tidak diinginkan [12] atau menghilangkan data yang tidak tepat [11] tetapi tidak menghilangkan informasi penting pada data [13]. Oleh karena itu, tahap praproses menggunakan metode *smoothing* dapat meningkatkan kualitas data runtun waktu.

Sebuah penelitian tentang efek praproses pada *predictive accuracy* telah dilakukan [21]. Penelitian tersebut mempelajari efek *data smoothing* runtun waktu dalam tahap praproses pada *multi-layer perception neural network* (MLP NN). Metode-metode *smoothing* yang dibandingkan adalah metode *seasonal adjustment series* (SAS), *seasonal trend cycle* (STC), ES (sederhana, linear, Holt, dan Winters), dan Box-Jenkins (*moving average*, *auto regression*, dan ARIMA). Hasil

menunjukkan bahwa *smoothing* dalam tahap praproses dapat meningkatkan akurasi *forecasting* NN dan penggunaan metode *smoothing* dapat menghilangkan derau yang terjadi.

Referensi [21] membuktikan bahwa salah satu tahap praproses data runtun waktu, yaitu tahap *data smoothing* yang diaplikasikan sebelum masukan ANN, dapat meningkatkan *predictive accuracy*. Selain itu, sebuah penelitian tentang peningkatan akurasi *forecasting* data runtun waktu juga telah dilakukan [22]. Penelitian ini berfokus pada praproses data untuk mengekstraksi informasi yang bersifat *redundant* dari isyarat dan mentransformasikan data dalam tahap praproses menggunakan *discrete cosine transform* (DCT) untuk meningkatkan akurasi *forecasting feedforward neural network* (FFNN).

Dalam [16], metode ES digunakan untuk mentransformasi data runtun waktu untuk digunakan sebagai masukan NN. Hasilnya menunjukkan peningkatan akurasi *forecasting* harga emas. Penelitian tersebut menunjukkan bahwa metode ES meningkatkan kualitas data yang akan digunakan oleh ANN.

Dalam [10], beberapa metode *smoothing* dibandingkan pada tahap praproses. Metode-metode yang dibandingkan adalah *simple moving average* (SMA), *centered moving average* (*centered MA*), *double moving average* (*double MA*), *weighted moving average* (*weight MA*), dan *modified weighted moving average algorithm* (*modified-WMA*). Hasil prediksi menunjukkan bahwa *smoothing centered MA* dapat meningkatkan kinerja akurasi prediksi. Penelitian sebelumnya menggunakan metode ES untuk mentransformasi data sebelum digunakan sebagai masukan NN [16]. Metode ES digunakan untuk meningkatkan kualitas runtun data yang akan digunakan sebagai data latih pada ANN. Model ANN digunakan untuk memprediksi harga emas.

III. DATA

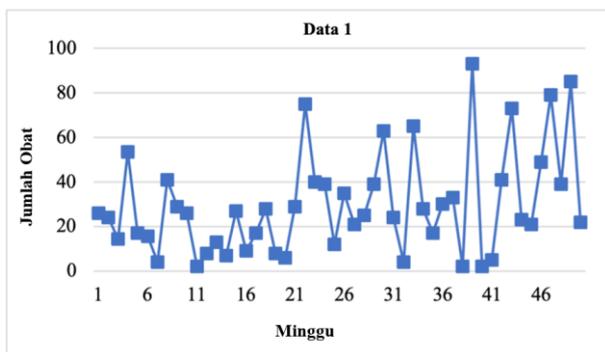
Data yang digunakan dalam penelitian ini adalah data jumlah obat-obatan yang digunakan di rumah sakit "X". Dua data runtun waktu menunjukkan penggunaan dua tipe obat, yaitu Cephalexin (dalam mililiter) dan Enxyplex (dalam gram). Data runtun waktu yang digunakan berupa data mingguan selama tahun 2015. Data runtun waktu untuk data 1 dan data 2 diperlihatkan pada Gbr. 1 dan Gbr. 2. Sumbu x menunjukkan minggu, sedangkan sumbu y menunjukkan jumlah obat yang digunakan.

IV. METODOLOGI

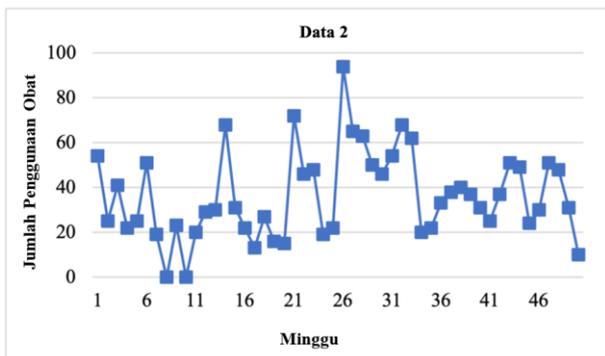
Tahap-tahap yang dilakukan dalam penelitian ini dijelaskan sebagai berikut.

A. Deteksi Lag

Dalam tahap ini, *time lag* dideteksi pada data runtun waktu menggunakan *autocorrelation function* (ACF) dan *partial autocorrelation function* (PACF). Ketergantungan antara data aktual dan nilai di masa lalu adalah representasi prinsip dasar *forecasting* runtun waktu. Kesamaan data untuk tiap nilai dalam runtun data yang saling berdekatan diamati untuk menemukan ketergantungan tersebut. Sebuah uji statistik untuk menentukan ketergantungan antara dua entitas yang dinyatakan dengan analisis korelasi, yang menghasilkan sebuah koefisien [12]. ACF dapat digunakan untuk menganalisis korelasi dalam



Gbr. 1 Runtun waktu 1.



Gbr. 2 Runtun waktu 2.

data runtun waktu. ACF mendeteksi ketergantungan yang penting dalam data runtun waktu [12].

Dalam runtun waktu, nilai data pertama sangat bergantung pada nilai data kedua, nilai data kedua sangat bergantung pada nilai data ketiga, dan seterusnya. Hal ini dapat menyebabkan tidak adanya ketergantungan yang signifikan pada runtun tersebut. Untuk menyelesaikan masalah ini, PACF dapat digunakan [12]. Analisis ACF dan PACF digunakan untuk menentukan masukan yang relevan dan jumlah tunda yang digunakan sebagai masukan bagi ANN.

B. Transformasi Data Runtun Waktu

Pada tahap ini, data runtun waktu dinormalisasi sebelum digunakan sebagai masukan ANN. Normalisasi adalah tahapan untuk menyesuaikan data ke dalam jangkauan yang kecil. Dalam makalah ini, nilai aktual diubah ke dalam jangkauan antara 0,1-0,9 menggunakan (1).

$$X' = \frac{0,8(x-a)}{b-a} + 0,1 \quad (1)$$

dengan x adalah nilai yang akan ditransformasi, a adalah nilai terkecil dari data runtun waktu, dan b adalah nilai terbesar dari data runtun waktu.

C. Penentuan Artificial Neural Network (ANN) Terstruktur

Jaringan *multilayer perceptron* (MLP) adalah salah satu arsitektur NN yang populer. Jaringan ini dapat memiliki satu atau lebih lapisan tersembunyi yang terletak di antara lapisan masukan dan lapisan keluaran. *Node* masukan adalah sebuah *time lag*, sedangkan *node* keluaran adalah nilai yang diprediksi. Model ANN yang digunakan dalam makalah ini adalah FFNN.

D. Metode Smoothing Whittaker-Henderson (WH)

Metode WH adalah salah satu metode *smoothing* yang banyak digunakan [23]. Untuk n adalah urutan pengukuran $\{y_1, y_2, \dots, y_n\}$, sebuah nilai real positif untuk λ , dan sebuah bilangan bulat positif $p < n$, dicari nilai $\{x_1, x_2, \dots, x_n\}$ yang meminimalkan (2).

$$\lambda \sum_{j=1}^n (y_j - x_j)^2 + \sum_{j=1}^{n-p} (\Delta^p x_j)^2 \quad (2)$$

dengan variabel Δ adalah operator beda maju.

Dengan n sebagai urutan pengukuran $\{y_1, y_2, \dots, y_n\}$, λ adalah bilangan real positif, dan $p < n$ adalah bilangan bulat positif, nilai $\{x_1, x_2, \dots, x_n\}$ yang meminimalkan persamaan berikut dihitung.

$$\Delta x_j = x_{j+1} - x_j \quad (3)$$

$$\Delta^2 x_j = \Delta(\Delta x_j) = x_{j+2} - 2x_{j+1} + x_j \quad (4)$$

dengan Δ adalah operator beda maju.

Dalam (3), jumlahan pertama digunakan untuk menghitung *fidelity* data, sedangkan jumlahan kedua pada (4) digunakan untuk menghitung *smoothness*. Dalam banyak kasus, nilai p yang digunakan adalah $p = 2$. Parameter λ mengendalikan *tradeoff* antara *fidelity* dan *smoothness* [14]. Ketika $\lambda \rightarrow 0$, nilai *smoothness* yang diestimasi konvergen ke $p-1$ derajat polinomial. Sementara itu, nilai yang diestimasi konvergen ke data real ketika $\lambda \rightarrow \infty$. Pendekatan yang digunakan untuk memilih parameter *smoothness* λ adalah *trial and error* atau GCV [15]. Nilai λ yang dipilih digunakan untuk meminimalkan skor GCV [14], [15], [23], [24].

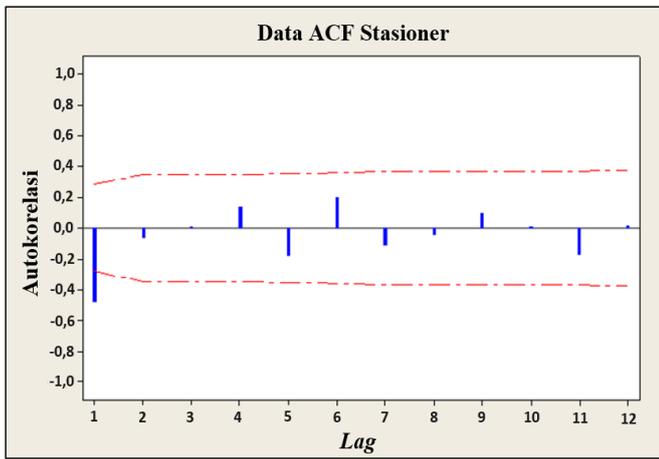
$$GCV = \frac{1}{n} \sum_{j=1}^n \left(\frac{y_j - \bar{x}_j}{1 - n^{-1} \text{trace}(\lambda A^{-1})} \right)^2 \quad (5)$$

dengan y adalah data real, \bar{x} adalah data yang dihasilkan dari proses *smoothing*, dan *trace* dari (λA^{-1}) disebut juga matriks "hat".

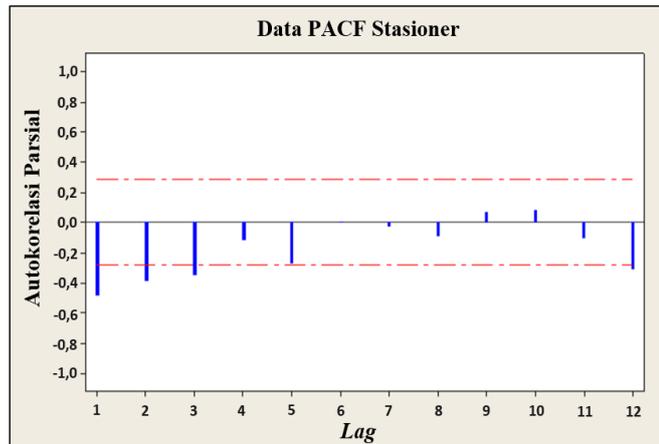
V. HASIL EKSPERIMEN

A. Deteksi Lag

1) *Dataset 1*: Data runtun waktu harus stasioner agar dapat mencari *lag-lag* yang signifikan. Hal ini berarti runtun tersebut stasioner terhadap rerata dan varians. Jumlah *lag* dari ACF dan PACF yang signifikan akan digunakan sebagai jumlah *node* masukan ANN. Gbr. 3 menunjukkan hasil deteksi *lag* dari *dataset 1*. *Plot* ACF dalam Gbr. 3(a) menunjukkan bahwa *lag* 1 adalah *lag* yang signifikan, sedangkan dalam *plot* PACF, Gbr. 3(b) menunjukkan bahwa *lag* 1, *lag* 2, *lag* 3, dan *lag* 12 adalah *lag-lag* yang signifikan. Referensi [25] menyatakan bahwa lebih baik dipilih *lag* signifikan pada permulaan ACF atau PACF, sehingga *lag* 1, *lag* 2, dan *lag* 3 dipilih sebagai masukan ANN. Arsitektur ANN diatur oleh Aturan 1 dan Aturan 2. Dalam Aturan 1, jumlah neuron dalam lapisan tersembunyi sama dengan masukan, sedangkan dalam Aturan 2, jumlah neuron dalam lapisan tersembunyi adalah setengah jumlah neuron masukan. Persamaan untuk Aturan 1 dan Aturan 2 ditunjukkan dalam (6) dan (7). Gbr. 4 menunjukkan dua arsitektur yang digunakan untuk *dataset 1* dengan sejumlah neuron masukan. Gbr. 4(a) menunjukkan jumlah neuron dalam



(a)



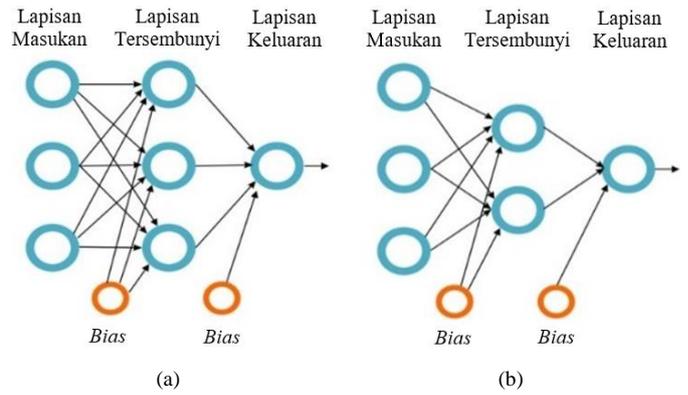
(b)

Gbr. 3 Grafik data stasioner, (a) *plot* data stasioner ACF, (b) *plot* data stasioner PACF

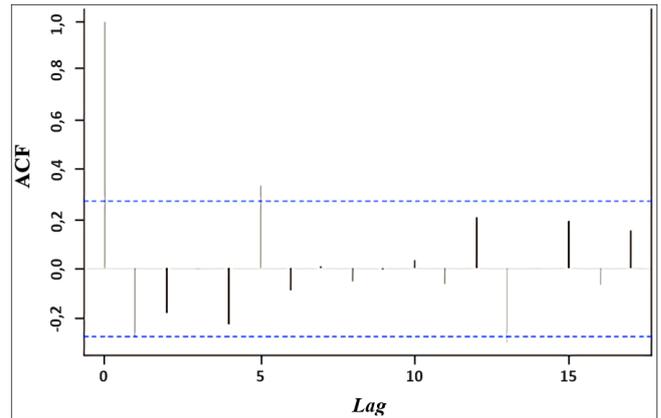
lapisan tersembunyi = 3, mengacu pada Aturan 1, sedangkan Gbr. 4(b) menunjukkan jumlah neuron dalam lapisan tersembunyi = 2, mengacu pada Aturan 2.

2) *Dataset 2*: Gbr. 5 menunjukkan hasil deteksi *lag* pada *dataset 2*. Gbr. 5(a) memperlihatkan *plot* ACF data stasioner, dengan *lag* 1 and *lag* 5 merupakan *lag* signifikan. Sementara itu, Gbr. 5(b), yang merupakan *plot* PACF data stasioner, menunjukkan bahwa *lag* yang signifikan adalah *lag* 1, *lag* 2, *lag* 4, dan *lag* 5, serta jumlah masukan FFNN adalah empat *node* masukan.

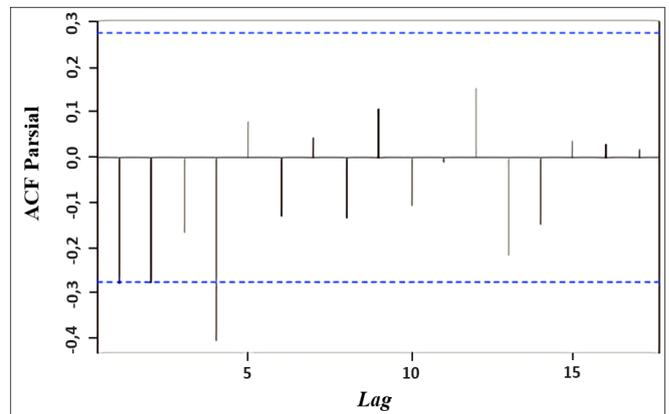
Arsitektur ANN architecture diatur oleh Aturan 1 dan Aturan 2. Dalam Aturan 1, jumlah neuron pada lapisan tersembunyi sama dengan masukan, sedangkan pada Aturan 2, jumlah neuron pada lapisan tersembunyi adalah setengah jumlah neuron masukan. Formula untuk Aturan 1 dan Aturan 2 ditunjukkan dalam (2) dan (3). Terdapat dua model arsitektur ANN untuk *dataset 2* dengan jumlah neuron masukan = 4, seperti ditunjukkan pada Gbr. 6. Gbr. 6(a) menunjukkan jumlah neuron pada lapisan tersembunyi = 4, mengacu pada Aturan 1, sedangkan Gbr. 6(b) menunjukkan jumlah neuron pada lapisan tersembunyi = 2, mengacu pada Aturan 2.



Gbr. 4 Arsitektur ANN untuk *dataset 1*, (a) jumlah neuron pada lapisan tersembunyi = 3, (b) jumlah neuron pada lapisan tersembunyi = 2.



(a)



(b)

Gbr. 5 Hasil deteksi *lag* *dataset 2*, (a) *plot* ACF data stasioner, (b) *plot* PACF data stasioner.

B. Transformasi Data Runtun Waktu

1) *Dataset 1*: Sebelum transformasi atau tahap *smoothing* diimplementasikan, data runtun waktu perlu dinormalisasi terlebih dahulu. Dalam tahap ini, ES sederhana (Gbr. 7) dan *smoothing* WH (Gbr. 8) diimplementasikan pada data runtun waktu, lalu akan dihasilkan *dataset* runtun waktu baru.

2) *Dataset 2*: Gbr. 9 dan Gbr. 10 mengilustrasikan hasil ES sederhana dan *smoothing* WH pada *dataset 2* runtun waktu.

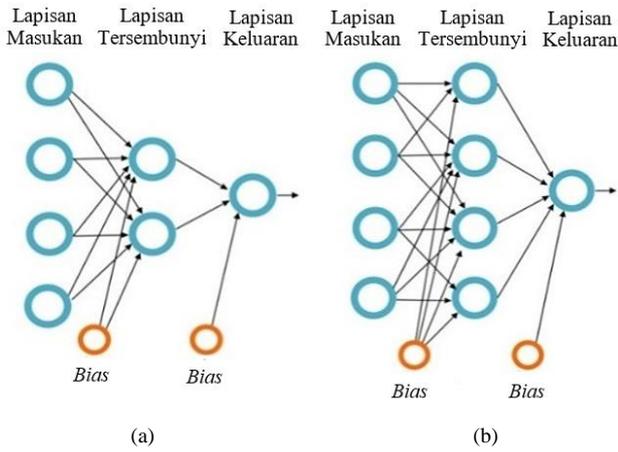


Fig. 6 Arsitektur ANN untuk dataset 2, (a) jumlah neuron pada lapisan tersembunyi = 4, (b) jumlah neuron pada lapisan tersembunyi = 2.

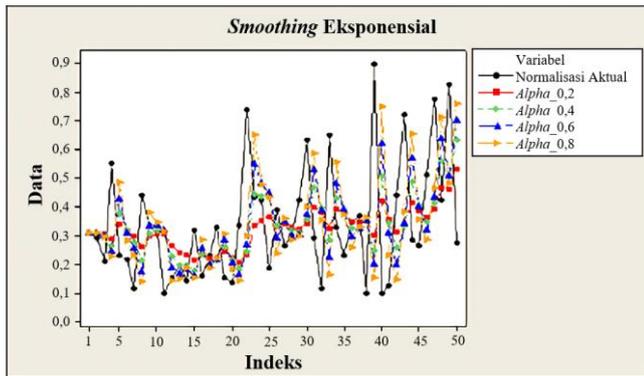
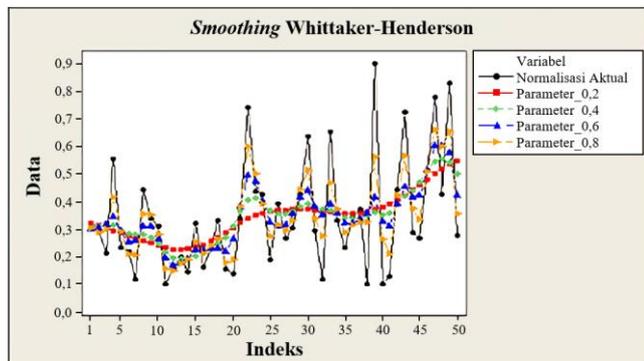


Fig. 7 Dataset 1 smoothing menggunakan ES sederhana.

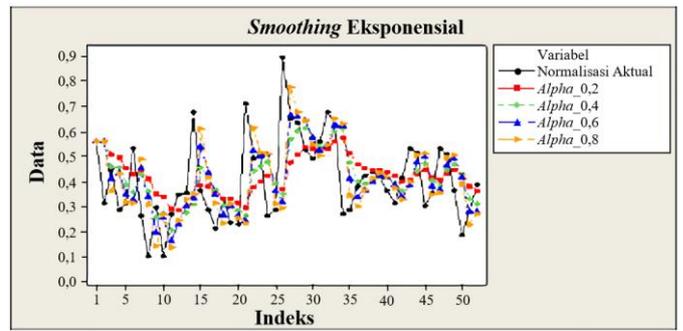


Gbr. 8 Smoothing WH untuk dataset 1, dengan $\sigma = 0,2; 0,4; 0,6; 0,8$.

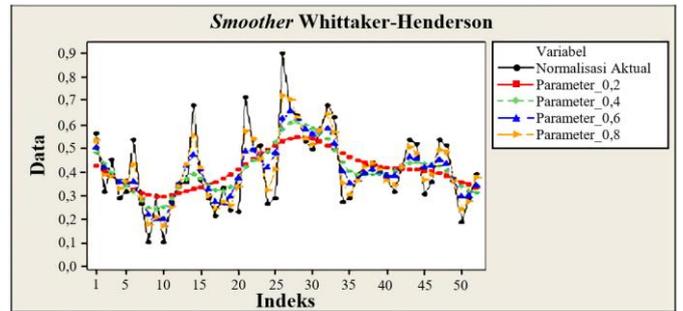
C. Perbandingan Model

Penentuan jumlah masukan ANN dilakukan menggunakan analisis ACF-PACF. Pengujian dilakukan pada data dengan dan tanpa *smoothing* WH. Kemudian, hasil MSE results dibandingkan untuk menemukan perbedaan kinerja *forecasting*.

Pada makalah ini, terdapat dua fase implementasi ANN. Fase 1 dilakukan untuk menemukan struktur ANN terbaik dengan menguji laju pembelajaran (*learning rate*, LR) yang berbeda. Setelah struktur ANN terbaik diperoleh, struktur tersebut digunakan untuk fase 2. Pada fase 2, struktur ANN tersebut diuji menggunakan data yang telah ditransformasi



Gbr. 9 Smoothing dataset 2 menggunakan ES sederhana, dengan $\sigma = 0,2; 0,4; 0,6; 0,8$.



Gbr. 10 Smoothing dataset 2 menggunakan *smoothing* WH, dengan $\sigma = 0,2; 0,4; 0,6; 0,8$.

TABEL I
PARAMETER ANN YANG DIGUNAKAN

| Parameter | Nilai |
|----------------------------------|---------------|
| Maksium epoch | 1.000 |
| MSE target | 0 |
| Kegagalan validasi maksimum | 6 |
| Konstanta momentum | 0,7 |
| Konstanta laju pembelajaran (LR) | 0,1; 0,3; 0,5 |

TABEL II
NILAI PARAMETER *SMOOTHING* TERBAIK UNTUK TIAP ATURAN METODE WH+ANN (WH+ANN) UNTUK DATASET 1

| Aturan | LR | Parameter <i>Smoothing</i> (σ) |
|----------|-----|---|
| Aturan 1 | 0,3 | 0,1 |
| Aturan 1 | 0,5 | 0,1 |
| Aturan 2 | 0,5 | 0,1 |

TABEL III
NILAI PARAMETER *SMOOTHING* TERBAIK UNTUK TIAP ATURAN METODE WH+ANN (WH+ANN) UNTUK DATASET 2

| Aturan | LR | Parameter <i>Smoothing</i> (σ) |
|----------|-----|---|
| Aturan 1 | 0,3 | 0,8 |
| Aturan 1 | 0,5 | 0,7 |
| Aturan 2 | 0,5 | 0,9 |

dengan metode ES dan WH. Algoritme pembelajaran yang digunakan dalam penelitian ini adalah *gradient descent with momentum* (GDM). Parameter pembelajaran bagi algoritme ANN-GDM disajikan dalam Tabel I. Nilai parameter *smoothing* terbaik untuk dataset 1 dan dataset 2 ditunjukkan dalam Tabel II dan Tabel III.

TABEL IV
STATISTIK DESKRIPTIF DATASET 1 MODEL TERPILIH

| Model | ANN | | | WH+ANN | | |
|-----------------|--------|--------|----------------|------------|------------|----------------|
| | Rerata | Median | Simpangan Baku | Rerata MSE | Median MSE | Simpangan Baku |
| Aturan 1-LR 0,3 | 983,73 | 868,71 | 355,12 | 792,86 | 721,04 | 277,16 |
| Aturan 1-LR 0,5 | 949,80 | 830,05 | 630,78 | 768,26 | 733,31 | 142,79 |
| Aturan 2-LR 0,5 | 819,93 | 786,54 | 148,88 | 759,52 | 727,66 | 127,02 |

TABEL V
STATISTIK DESKRIPTIF DATASET 2 MODEL TERPILIH

| Model | ANN | | | WH+ANN | | |
|-----------------|--------|--------|----------------|------------|------------|----------------|
| | Rerata | Median | Simpangan Baku | Rerata MSE | Median MSE | Simpangan Baku |
| Aturan 1-LR 0,3 | 256,79 | 182,80 | 375,65 | 219,41 | 178,50 | 245,62 |
| Aturan 1-LR 0,5 | 300,60 | 191,50 | 410,90 | 276,65 | 192,20 | 192,20 |
| Aturan 2-LR 0,5 | 217,37 | 188,00 | 153,48 | 222,08 | 185,20 | 275,95 |

Pengujian dilakukan menggunakan satu lapisan tersembunyi dan dua kondisi neuron tersembunyi.

Aturan 1:

$$N_h = N_{in} \tag{6}$$

Aturan 2:

$$N_h = N_{in}/2 \tag{7}$$

dengan N_h = jumlah neuron tersembunyi dan N_{in} = jumlah masukan ANN.

Aturan 1 telah diusulkan pada [26], yang meneliti efek jumlah neuron tersembunyi terhadap kinerja *forecasting*. Dalam [27], beberapa penelitian menunjukkan bahwa dengan menggunakan jumlah neuron yang sama dengan jumlah neuron masukan, dihasilkan *forecasting* yang baik. Sementara itu, Aturan 2 diusulkan pada [28], dengan jumlah neuron lapisan tersembunyi adalah setengah jumlah neuron masukan jaringan. Fungsi aktivasi yang digunakan pada lapisan tersembunyi adalah *hyperbolic tangent*, sedangkan fungsi aktivasi yang digunakan pada neuron keluaran adalah fungsi aktivasi linear.

Dalam penelitian ini, data dibagi menjadi 70% data untuk pelatihan, 15% untuk validasi, dan 15% untuk pengujian. Akurasi diukur dengan pengujian data, sedangkan *error forecasting* diukur menggunakan MSE dengan formula seperti pada (8).

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - y'_t)^2 \tag{8}$$

dengan y_t adalah nilai aktual pada waktu t dan y'_t adalah nilai yang diprediksi pada waktu t . Pengujian dilakukan sebanyak seratus kali dengan bobot awal acak dalam rentang kecil antara -1 sampai 1. Sebuah mekanisme penghentian awal digunakan untuk mencegah *overfitting*. Implementasi metode ANN ini menggunakan perangkat lunak MATLAB.

Setelah LR terbaik tiap aturan diperoleh, arsitektur dengan LR tersebut dipilih sebagai arsitektur terbaik untuk data aktual. Kemudian dilakukan pengujian ulang data yang telah dikenai *smoothing* menggunakan metode WH (WH+ANN). Hasil MSE dari model ANN yang belum dikenai *smoothing* dibandingkan dengan MSE dari WH+ANN untuk menentukan terdapat perbedaan pada nilai MSE yang diprediksi atau tidak.

Hasil pengujian berdasarkan parameter LR terbaik pada kedua aturan untuk *dataset 1* ditunjukkan pada Tabel IV,

sedangkan hasil untuk *dataset 2* disajikan dalam Tabel V. Uji beda statistik antar grup data digunakan untuk menentukan nilai parameter terbaik dari arsitektur ANN. Uji statistik MSE untuk kedua *dataset* menunjukkan bahwa model ANN terbaik diperoleh dengan LR = 0,3 dan LR = 0,5 untuk Aturan 1 dan LR = 0,5 untuk Aturan 2. LR sebesar 0,3 dan 0,5 pada Aturan 1 dipilih karena hasil pengujian kedua LR menunjukkan bahwa akuisisi MSE untuk kedua LR tersebut tidak berbeda secara statistik dengan *confidence level* kurang dari 95%. Oleh karena itu, kedua model ANN tersebut digunakan untuk merepresentasikan model terbaik Aturan 1 dan digunakan untuk pengujian selanjutnya. Hasil akuisisi MSE dari model-model tersebut kemudian dibandingkan dengan akuisisi MSE dari ANN dengan metode WH (WH+ANN).

Perbedaan *confidence level* pada nilai MSE metode WH+ANN pada *dataset 2* dapat terjadi. Hal tersebut dapat terjadi karena perbedaan pola *spikes* antara kedua *dataset*. Pengamatan terhadap *plot* data runtun waktu menunjukkan bahwa *dataset 2* mengandung *spikes* atau pola yang berfluktuasi lebih sedikit daripada *dataset 1*, yang mengindikasikan bahwa proses *smoothing* pada *dataset 2* tidak memberikan efek *smoothing* yang signifikan. Maka, hasil perbandingan MSE antara WH+ANN dan ANN untuk *dataset 2* tidak menghasilkan perbedaan yang signifikan secara statistik dengan *significance level* 0,05.

Hasil pengujian untuk *dataset 1* diperlihatkan pada Tabel IV, yang menunjukkan bahwa akuisisi MSE dari WH+ANN lebih kecil daripada akuisisi MSE dari ANN untuk kedua aturan, Aturan 1 dan Aturan 2. Sementara itu, Tabel V menunjukkan hasil pengujian pada *dataset 2*. Terlihat bahwa akuisisi MSE dari WH+ANN cenderung memiliki nilai yang lebih kecil daripada akuisisi MSE dari ANN untuk Aturan 1 dan Aturan 2. Hasil menunjukkan bahwa metode ANN yang menggunakan metode *smoothing* WH (WH+ANN) memiliki MSE yang lebih rendah secara signifikan daripada metode ANN *non-smoothing* dengan *confidence level* sebesar 94% untuk *dataset 1* dan 85% untuk *dataset 2*.

Hasil eksperimen menunjukkan bahwa pemilihan yang tepat terhadap nilai parameter *smoothing* dapat menghasilkan nilai MSE yang rendah, dan sebaliknya pemilihan nilai parameter *smoothing* yang tidak benar dapat menghasilkan prediksi MSE yang tinggi. Penentuan nilai parameter *smoothing* yang sesuai

merupakan sebuah tantangan dalam mengaplikasikan metode *smoothing* pada fase praproses data runtun waktu untuk meningkatkan akurasi *forecasting* ANN. Hasil menunjukkan bahwa parameter *smoothing* WH dapat menghasilkan MSE terendah jika memiliki nilai GCV terkecil. Akan tetapi, perlu diselidiki GCV yang rendah selalu menghasilkan *forecasting* MSE untuk ANN yang rendah atau tidak.

VI. KESIMPULAN

Makalah ini berfokus pada pengaplikasian metode *smoothing* pada tahap praproses metode ANN. Data runtun waktu harus bersifat stasioner untuk mencari *lag* yang signifikan. Dalam hal ini, runtun bersifat stasioner terhadap rerata dan varians. Jumlah *lag* ACF dan PACF yang signifikan digunakan sebagai jumlah *node* masukan ANN.

Perbandingan prediksi dilakukan dengan membandingkan metode ANN dengan dan tanpa *smoothing*. Metode *smoothing* yang digunakan adalah WH. Hasil menunjukkan bahwa WH+ANN memiliki MSE yang lebih rendah secara statistik dibandingkan ANN (tanpa *smoothing*) dengan *confidence level* 94% untuk *dataset* 1 dan 85% untuk *dataset* 2. Hal ini menunjukkan bahwa proses *data smoothing* memengaruhi akurasi *forecasting* model ANN.

KONFLIK KEPENTINGAN

Penulis menyatakan tidak ada konflik kepentingan dalam penulisan makalah ini.

KONTRIBUSI PENULIS

Konseptualisasi dan metodologi, Hans Pratyaksa, Adhitya Erna Permanasari, Silmi Fauziati; komputasi, Hans Pratyaksa; analisis formal, Hans Pratyaksa; validasi dan evaluasi, Hans Pratyaksa, Adhitya Erna Permanasari, Silmi Fauziati; persiapan draf asli, Hans Pratyaksa, Adhitya Erna Permanasari; penulisan—peninjauan dan penyuntingan, Adhitya Erna Permanasari; pengawasan, Adhitya Erna Permanasari, Silmi Fauziati; administrasi proyek, Adhitya Erna Permanasari; akuisisi pendanaan, Adhitya Erna Permanasari.

REFERENSI

- [1] M.I. Ramos, J.J. Cubillas, dan F.R. Feito, "Improvement of the Prediction of Drugs Demand Using Spatial Data Mining Tools," *J. Med. Syst.*, Vol. 40, No. 1, hal. 1-9, Jan. 2016.
- [2] H. Pratyaksa, A.E. Permanasari, S. Fauziati, dan I. Fitriana, "ARIMA Implementation to Predict the Amount of Antiseptic Medicine Usage in Veterinary Hospital," *2016 1st Int. Conf. Biomed. Eng. (BIOMED)*, 2016, hal. 1-4.
- [3] C. Qingkui dan R. Junhu, "Study on the Demand Forecasting of Hospital Stocks Based on Data Mining and BP Neural Networks," *2009 Int. Conf. Electron. Commerce, Bus. Intell.*, 2009, hal. 284-289.
- [4] M. Jalalpour, Y. Gel, dan S. Levin, "Forecasting Demand for Health Services: Development of A Publicly Available Toolbox," *Oper. Res. Health Care*, Vol. 5, hal. 1-9, Jun. 2015.
- [5] N.K. Zadeh, M.M. Sepehri, dan H. Farvaresh, "Intelligent Sales Prediction for Pharmaceutical Distribution Companies: A Data Mining Based Approach," *Math. Probl. Eng.*, Vol. 2014, hal. 1-15, Mei 2014.
- [6] Á. Lublóy, "Factors Affecting the Uptake of New Medicines: A Systematic Literature Review," *BMC Health Serv. Res.*, Vol. 14, hal. 1-25, Okt. 2014.
- [7] D. Alba-Cuellar, *et al.*, "Time Series Forecasting with PSO-Optimized Neural Networks," *2014 13th Mexican Int. Conf. Artif. Intell.*, 2014, hal. 102-111.
- [8] F. Yu dan X. Xu, "A Short-Term Load Forecasting Model of Natural Gas Based on Optimized Genetic Algorithm and Improved BP Neural Network," *Appl. Energy*, Vol. 134, hal. 102-113, Des. 2014.
- [9] I. Khandelwal, R. Adhikari, dan G. Verma, "Time Series Forecasting Using Hybrid ARIMA and ANN Models Based on DWT Decomposition," *Procedia Comput. Sci.*, Vol. 48, hal. 173-179, 2015.
- [10] F. Nhita, D. Saepudin, Adiwijaya, dan U.N. Wisesty, "Comparative Study of Moving Average on Rainfall Time Series Data for Rainfall Forecasting Based on Evolving Neural Network Classifier," *2015 3rd Int. Symp. Comput., Bus. Intell. (ISCBI)*, 2015, hal. 112-116.
- [11] D.I. Wilson, "The Black Art of Smoothing," *Elect., Automat. Technol.*, No. June/July, hal. 35-36, 2006.
- [12] O. Ostashchuk, "Time Series Data Prediction and Analysis," Master thesis, Czech Technical University in Prague, Prague, Czech Republic, 2017.
- [13] D. Garcia, "Robust Smoothing of Gridded Data in One and Higher Dimensions with Missing Values," *Comput. Statist., Data Anal.*, Vol. 54, No. 4, hal. 1167-1178, Apr. 2010.
- [14] H.L. Weinert, "Efficient Computation for Whittaker-Henderson Smoothing," *Computat. Statist., Data Anal.*, Vol. 52, No. 2, hal. 959-974, Okt. 2007.
- [15] J.J. Stickel, "Data Smoothing and Numerical Differentiation by A Regularization Method," *Comput., Chem. Eng.*, Vol. 34, No. 4, hal. 467-475, Apr. 2010.
- [16] I. Suryani dan R.S. Wahono, "Penerapan Exponential Smoothing untuk Transformasi Data dalam Meningkatkan Akurasi Neural Network pada Prediksi Harga Emas," *J. Intell. Syst.*, Vol. 1, No. 2, hal. 67-75, Des. 2015.
- [17] I.N. Soyiri dan D.D. Reidpath, "An Overview of Health Forecasting," *Environ. Health, Prev. Med.*, Vol. 18, hal. 1-9, Jan. 2013.
- [18] I.K. Utami, "Seleksi Input untuk Artificial Neural Network Menggunakan Binary Particle Swarm Optimization dalam Pemodelan Runtun Waktu Kasus Avian Influenza," Master thesis, Universitas Gadjah Mada, Yogyakarta, Indonesia, 2016.
- [19] G. Lachtermacher dan J.D. Fuller, "Backpropagation in Time-Series Forecasting," *J. Forecast.*, Vol. 14, No. 4, hal. 381-393, Jul. 1995.
- [20] F.S. Wong, "Time Series Forecasting Using Back Neural Networks," *Neurocomputing*, Vol. 2, No. 4, hal. 147-159, Jul. 1991.
- [21] A. Azadeh, M. Sheikhalishahi, M. Tabesh, dan A. Negahban, "The Effects of Pre-Processing Methods on Forecasting Improvement of Artificial Neural Networks," *Aust. J. Basic, Appl. Sci.*, Vol. 5, No. 6, hal. 570-580, Jun. 2011.
- [22] S. Anbazhagan dan N. Kumarappan, "Day-Ahead Deregulated Electricity Market Price Forecasting Using Neural Network Input Featured by DCT," *Energy Convers., Manage.*, Vol. 78, No. 2, hal. 711-719, Feb. 2014.
- [23] A.S. Nocon dan W.F. Scott, "An Extension of the Whittaker-Henderson Method of Graduation," *Scand. Actuar. J.*, Vol. 2012, No. 1, hal. 70-79, Mar. 2012.
- [24] P.H.C. Eilers, "A Perfect Smoother," *Anal. Chem.*, Vol. 75, No. 14, hal. 3631-3636, May 2003.
- [25] H. Abolfazli, S.M. Asadzadeh, dan S.M. Asadzadeh, "Forecasting Rail Transport Petroleum Consumption Using an Integrated Model of Autocorrelation Functions-Artificial Neural Network," *Acta Polytech. Hungarica*, Vol. 11, No. 2, hal. 203-214, Jan. 2014.
- [26] Z. Tang dan P.A. Fishwick, "Feedforward Neural Nets as Models for Time Series Forecasting," *ORSA J. Comput.*, Vol. 5, No. 4, hal. 374-385, Nov. 1993.
- [27] P.G. Zhang, E. Patuwo, dan M. Y. Hu, "Forecasting with Artificial Neural Networks: The State of the Art," *Int. J. Forecasting*, Vol. 14, No. 1, hal. 35-62, Mar. 1998.
- [28] S.Y. Kang, "An Investigation of the Use of Feedforward Neural Networks for Forecasting," Ph.D. dissertation, Kent State University, Ohio, USA, 1991.