

PENENTUAN UNIT *HIDDEN OPTIMAL* PADA MODEL *NEURAL NETWORK* DENGAN ANALISIS KONTRIBUSI *INCREMENTAL SEL*

Budi Warsito
Jurusan Matematika FMIPA UNDIP
Jl. Prof. H. Soedarto, S.H, Semarang 50275

Abstract. This paper discusses about choosing the optimal number of hidden units at neural network models which is applied to the Composite Stock Price Index data in Surabaya Stock Exchange. One of the problem in fitting NN models is an NN model which fits well may give poor out-of-sample forecasts. Thus it is required traditional methods such as ACF and PACF to select a good NN model, e.g. to select appropriate lagged variables as the 'inputs'. The Incremental Contribution of Cells methods which belong to the general-to-specific procedure is used to choose the optimal number of hidden units. The size and topology of networks is selected by reducing the size of the network through the use of quadratic correlation coefficients and graphical analysis of network output for every hidden layer cell. The resulted of NN model is compared with models with different architecture those obtain from the Box-Jenkins methods. The Akaike's Information Criterion and Schwartz Bayesian Criterion are used for comparing different models.

Keywords: Nonlinear Autoregressive, Neural Network, Incremental Contribution of Cells, Composite Stock Price Index

1. PENDAHULUAN

Artificial neural network (ANN atau NN saja) merupakan jaringan yang saling berhubungan antar node atau simpul dimana tiap-tiap hubungan mempunyai bobot koneksi (*weight*) yang dilatih untuk mencapai respon yang diinginkan. Masing-masing bobot dipropagasi ke seluruh simpul dengan suatu pelatihan untuk memperoleh output yang diinginkan. Beberapa tulisan seperti [4] dan [2] menyebutkan bahwa algoritma pembelajaran NN dapat menyelesaikan permasalahan model deret berkala nonlinear.

Beberapa buku teks menggambarkan arsitektur dan algoritma NN ([5],[3]). Beberapa publikasi ilmiah juga menunjukkan kemampuan NN untuk menyelesaikan berbagai permasalahan seperti klasifikasi observasi [6] dan regresi logistik [1]. Dalam pembentukan model NN untuk time series, [9] telah melakukan perbandingan performa model ARIMA, UCM dan NN berdasarkan MAE dan RMSE. Sedangkan [2] dan [4] membandingkan model ARIMA dan NN dengan input lag terpilih dari model ARIMA. [8] menggunakan

hukum Baum-Haussler untuk menentukan jumlah unit hidden. Sedangkan [7] memperkenalkan metode *pruning* dengan analisis kontribusi *incremental sel*, analisis grafik dan komponen utama.

Tulisan ini akan membahas cara menentukan jumlah unit hidden optimal pada model NN data time series dengan menggunakan prosedur *general-to-specific* melalui analisis kontribusi *incremental sel* dan analisis grafik dengan mengeluarkan salah satu sel dari jaringan secara bergantian. Sel dengan kontribusi kecil akan dikeluarkan dari model sehingga diperoleh jumlah sel (unit) hidden optimal.

2. MODEL NEURAL NETWORK

NN terdiri dari elemen-elemen pemrosesan sederhana yang disebut *neuron*. Setiap neuron dihubungkan ke neuron lain dengan link komunikasi yang disebut arsitektur jaringan. Bobot koneksi mewakili besarnya informasi yang digunakan. Operasi dasar sebuah neuron buatan meliputi jumlahan dari hasil perkalian sinyal input dengan bobot antar neuron, kemudian menggunakan fungsi aktivasi untuk

menghasilkan output. Fungsi aktivasi yang sering digunakan adalah *Logistic Sigmoid* $f(x)=(1+exp(-x))^{-1}$ Pada setiap pelatihan, jaringan menghitung respon dari unit output dan kesalahan dengan membandingkan output terhitung dengan nilai target.

Pelatihan jaringan dengan back-propagation meliputi tiga tahap yaitu umpan maju (*feedforward*) dari pola input, penghitungan dan propagasi balik dari error dan penyesuaian bobot. Pada tahap umpan maju setiap unit input menerima sinyal input (x_i) dan menyebarkannya ke unit tersembunyi z_1, \dots, z_p . Setiap unit tersembunyi menghitung aktivasinya dan jumlah terboboti dari input-inputnya dalam bentuk

$$z_{in_j} = \sum_i w_{ji}x_i + w_{bj}, \quad (2.1)$$

dengan x_i adalah aktivasi dari unit input ke- i yang mengirimkan sinyal ke unit hidden ke j , w_j adalah bobot dari sinyal yang terkirim dan $j = 1, 2, \dots, q$ adalah jumlah unit hidden. Hasil penjumlahan di-transformasi dengan fungsi aktivasi non-linear $f(\cdot)$

$$z_j = f(z_{in_j}). \quad (2.2)$$

Setelah semua unit tersembunyi menghitung aktivasinya kemudian mengirimkan sinyal (z_j) ke unit output. Kemudian unit output menghitung aktivasinya dalam bentuk

$$g(w, z) = \sum_j w_{jo}z_j + w_{bo}. \quad (2.3)$$

Fungsi pada (3.3) merupakan nilai output dari jaringan yaitu

$$o = \sum_j w_{jo}f(a_j) + w_{bo}, \quad (2.4)$$

dengan w_{bo} adalah bobot dari bias ke unit output.

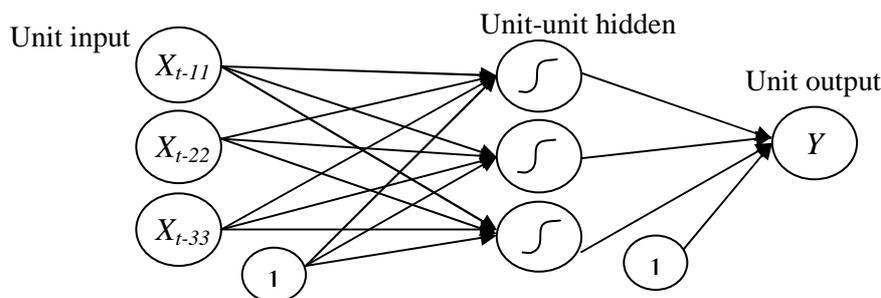
Arsitektur NN dengan unit input lag 1 sampai 3 dan unit konstan, satu *hidden layer* dengan 3 neuron dan 1 unit output diilustrasikan pada Gambar 1. Model NN dengan satu hidden dan input x_{t-1}, \dots, x_{t-p} ditulis dalam bentuk

$$\hat{x}_t = \psi_0 \left\{ w_{co} + \sum_n w_{no} \psi_n \left(w_{cn} + \sum_i w_{in} x_{t-j_i} \right) \right\} \quad (2.5)$$

dengan $\{w_{cn}\}$ adalah bobot antara unit konstan dan neuron dan w_{co} adalah bobot antara unit konstan dan output. $\{w_{in}\}$ dengan $\{w_{cn}\}$ adalah bobot antara unit konstan dan neuron dan w_{co} adalah bobot antara unit konstan dan output. $\{w_{in}\}$ dan $\{w_{no}\}$ masing-masing menyatakan bobot koneksi input dengan neuron dan antara neuron dengan output. Kedua fungsi ψ_n dan ψ_o masing-masing fungsi aktivasi yang digunakan pada neuron dan output. Notasi untuk model NN adalah $NN(j_1, \dots, j_k, n)$ yang menyatakan NN dengan input lag j_1, \dots, j_k dan n neuron. Bobot diestimasi dengan meminimumkan SSE, yaitu $S = \sum_t (\hat{x}_t - x_t)^2$.

3. JUMLAH UNIT HIDDEN

Penentuan jumlah unit hidden dengan analisis kontribusi incremental sel [7] dijelaskan sebagai berikut. Misalkan diberikan sejumlah H sel (unit) hidden layer maka akan diperoleh output jaringan dan kuantitas performa jaringan yang dinyatakan dengan kuadrat dari koefisien korelasi dari x dan \hat{x} yang dinyatakan sebagai



Gambar 1. Arsitektur NN untuk peramalan deret berkala dengan satu hidden layer yang terdiri 3 neuron dan variabel input nilai pada lag 1, 2, dan 3

$$R^2 = \frac{(\hat{x}'x)^2}{(x'x)(\hat{x}'\hat{x})}, \quad (3.1)$$

dengan \hat{x} adalah vektor nilai-nilai output jaringan.

Kita akan membandingkan data observasi dan output jaringan dimana satu sel dikeluarkan dan dengan semua sel dimasukkan dalam jaringan. Misalkan jika kontribusi dari sel hidden h sama dengan nol ($c_h=0$) maka jaringan akan menghasilkan output \hat{x}_{-h} dengan kesalahan

$$e_{-h} = x - \hat{x}_{-h}. \quad (3.2)$$

Koefisien korelasi kuadrat dari jaringan yang unit hiddennya telah dikurangi satu sel tersebut yaitu R_{-h}^2 antara x dan \hat{x}_{-h} dirumuskan dengan

$$R_{-h}^2 = \frac{(\hat{x}_{-h}'x)^2}{(x'x)(\hat{x}_{-h}'\hat{x}_{-h})}. \quad (3.3)$$

Kontribusi *incremental* dari sel h dinyatakan sebagai selisih dari R^2 dan R_{-h}^2 yaitu

$$R_{inc}^2 = R^2 - R_{-h}^2. \quad (3.4)$$

Jika R_{inc}^2 dari sel ke h bernilai rendah dibandingkan R_{inc}^2 dari semua sel yang lain maka sel ini menjadi kandidat untuk dikeluarkan dari jaringan. Dengan demikian jika dari sejumlah H sel tersebut terdapat H' sel yang mempunyai R_{inc}^2 yang bernilai rendah dibandingkan dengan sel-sel yang lain, maka jumlah sel yang tetap dipertahankan dalam jaringan sebanyak $H-H'$ sel. Ini merupakan jumlah unit hidden optimal dalam jaringan tersebut karena sel dengan R_{inc}^2 rendah pada dasarnya tidak mempunyai kontribusi yang nyata dalam arti penambahan jaringan dengan sel tersebut tidak terlalu mempengaruhi performa jaringan.

Reduksi sel juga dapat dilakukan dengan analisis grafik. Grafik dari $\{t, \hat{x}_{-h}(t)\}$ dibandingkan dengan grafik dari $\{t, x(t)\}$ dan perbandingan ini akan dapat memberi petunjuk kontribusi sel hidden h dalam menjelaskan variansi $x(t)$. Jika grafik $\{t, \hat{x}_{-h}(t)\}$ telah menunjukkan prediksi

yang baik untuk $\{t, x(t)\}$ maka sel h akan dikeluarkan dari model sebab jaringan tanpa sel h tersebut telah menghasilkan prediksi yang baik atau dengan kata lain penambahan sel h tidak mempunyai kontribusi yang nyata terhadap performa jaringan.

4. TERAPAN PADA DATA IHSG BES

Penelitian ini menggunakan data Indeks Harga Saham Gabungan (IHSG) pada Bursa Efek Surabaya (BES) mulai Juli 1995 – Maret 2005 sebanyak 105. Data diambil pada hari penutupan bursa setiap akhir bulan dengan perhitungan nilai IHSG pada hari ke- t

$$IHSG_t = \frac{\text{Nilai Pasar}_t}{\text{Nilai Dasar}} \times 100. \quad (3.5)$$

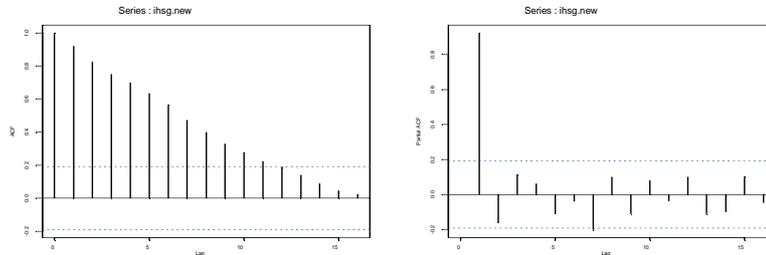
Data terlebih dahulu dibagi dengan 400 untuk menyesuaikan dengan fungsi aktivasi yang digunakan. Plot fungsi ACF dan PACF dari data hasil transformasi disajikan pada Gambar 2. Nampak bahwa plot ACF menurun menuju nol dan PACF signifikan pada lag pertama sehingga input yang digunakan adalah lag 1. Dalam hal ini asumsi stasioner tidak diperlukan.

Untuk menentukan model NN digunakan software Splus 2000 dibatasi untuk hidden layer tunggal dan fungsi aktivasi logistik sigmoid. Jumlah unit hidden ditentukan melalui prosedur “*general to specific*” menggunakan analisis kontribusi incremental sel dan analisis grafik dimulai dengan 6 unit dan diperoleh nilai prediksi model NN(1;6). Koefisien korelasi kuadrat yang diperoleh adalah $R^2=0,9885$. Koefisien korelasi kuadrat dari jaringan yang unit hiddennya telah dikurangi satu sel (sel ke h) yaitu R_{-h}^2 dan kontribusi dari masing-masing sel hidden h ($h=1,2, \dots, 6$) yaitu R_{inc}^2 disajikan pada Tabel 1. Plot data asli dan prediksi model NN masing-masing tanpa hidden layer ke 1, 2, 3, 4, 5 dan 6 ditunjukkan pada Gambar 3.

Dari Tabel 1 nampak bahwa nilai R_{inc}^2 dari sel 2 dan 4 secara signifikan lebih

kecil dari sel yang lain. Dari Gambar 3 juga nampak bahwa sel ke 2 dan 4 tidak memberi kontribusi yang cukup berarti dalam arti bahwa tanpa menggunakan kedua sel model telah menghasilkan prediksi yang cukup baik. Berdasarkan kedua hal

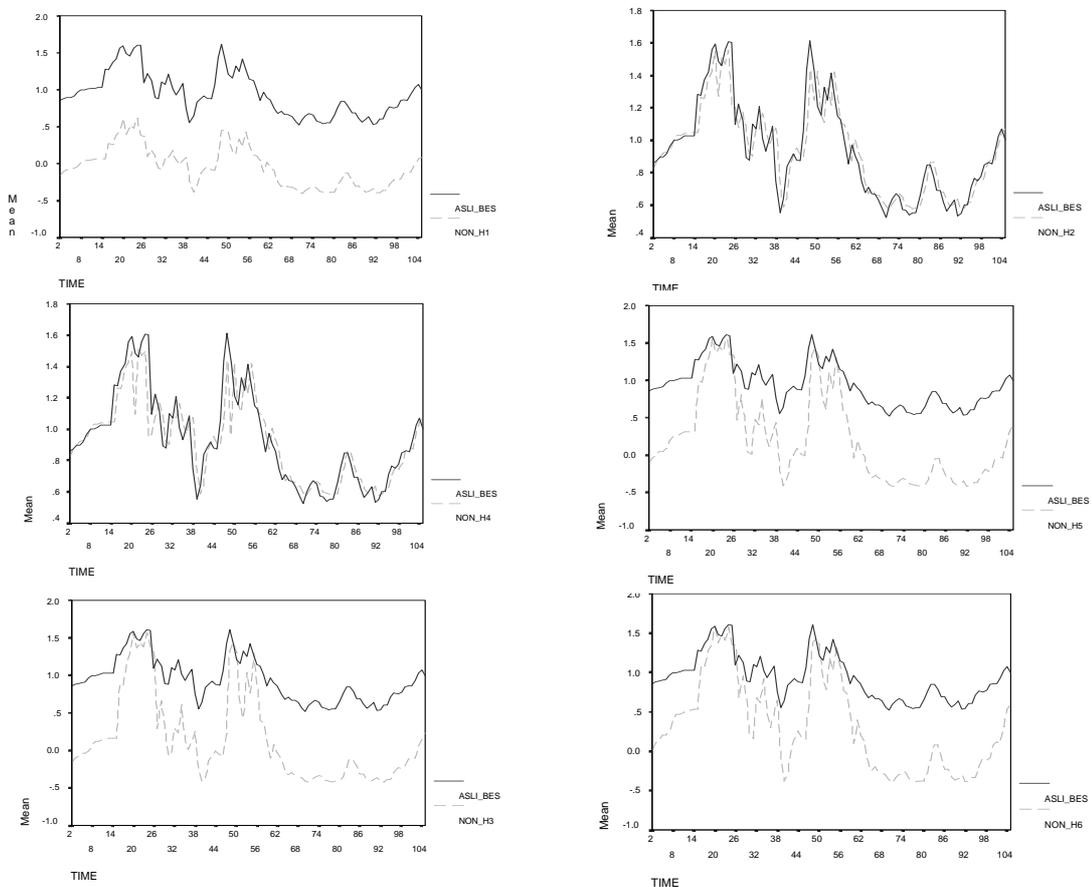
tersebut maka sel 2 dan 4 menjadi kandidat untuk dihapus dari model sehingga tinggal sel 1, 3, 5 dan 6. Model jaringan yang diperoleh terdiri dari empat unit hidden, yaitu NN(1;4) dengan jumlah parameter atau bobot 13.



Gambar 2. Plot ACF dan PACF data hasil transformasi IHSB BES

Tabel 1. Nilai R^2_{-h} dan R^2_{inc} dari masing-masing sel h

h	1	2	3	4	5	6
R^2_{-h}	0,0653	0,8487	0,1790	0,7024	0,2057	0,2463
R^2_{inc}	0,9237	0,1399	0,8095	0,2861	0,7828	0,7423



Gambar 3. Plot data asli dan prediksi model NN masing-masing tanpa sel hidden layer ke 1, 2, 3, 4, 5 dan 6

Dengan input lag 1 model ini mempunyai fungsi dengan bentuk

$$y_t = w_{bo} + \frac{w_{1o}}{1 + \exp(-w_{11}y_{t-1} - w_{b1})} + \frac{w_{2o}}{1 + \exp(-w_{12}y_{t-1} - w_{b2})} + \frac{w_{3o}}{1 + \exp(-w_{13}y_{t-1} - w_{b3})} + \frac{w_{4o}}{1 + \exp(-w_{14}y_{t-1} - w_{b4})} \quad (3.6)$$

5. PENUTUP

Berdasarkan hasil analisis dan pembahasan dapat disimpulkan sebagai berikut

1. Pada pembentukan model NN, jumlah unit hidden optimal dapat ditentukan berdasarkan kontribusi *incremental* sel dan analisis grafik.
2. Variabel input model NN untuk data IHSG pada Bursa Efek Surabaya berdasarkan plot ACF dan PACF adalah lag 1.
3. Jumlah unit hidden layer optimal berdasarkan kontribusi *incremental* tiap-tiap sel hidden adalah empat, yang dilihat dari nilai R_{inc}^2 dan analisis grafik sehingga model yang dihasilkan adalah NN(1; 4).

6. DAFTAR PUSTAKA

- [1] Afghohani, Afif (2003), *Perbandingan antara Regresi Logistik dengan Jaringan Syaraf Tiruan pada Kasus Berat Kelahiran Bayi Rendah*, Tesis, Universitas Gadjah Mada, Yogyakarta.
- [2] Allende, H., Moraga, C. and Salas, R. (1999), *Artificial Neural Networks in Time Series Forecasting: A Comparative Analysis*, Research Grant BMBF RCH99/023.

- [3] Bishop, Christopher, M. (1995), *Neural Networks for Pattern Recognition*, Oxford University Press, New York,
- [4] Faraway, J. and Chatfield, C. (1998), *Time Series Forecasting with Neural Networks: a Comparative Study Using the Airline Data*, Applied Statistics.
- [5] Fausett, Laurene, (1994), *Fundamentals of Neural Networks; Architectures, Algorithms and Applications*, Prentice-Hall Inc., Englewoods Cliffs, New Jersey.
- [6] Hakim, R.B.F. (2001), *Klasifikasi Penggunaan Analisis Diskriminan Linear dan Jaringan Syaraf Tiruan*, Tesis, Universitas Gadjah Mada, Yogyakarta.
- [7] Kaashoek, J.F., van Dick, H.K. (1998), *Neural Network Analysis of Varying Trends in Real Exchange Rates*, Report EI9915/A Econometric Institute Rotterdam
- [8] Lin, F., Yu, X.H., Gregor, S. and Irons, R. (1995), *Time Series Forecasting with Neural Networks*, Complexity International, **2**.
- [9] Portugal, M.S. (1995), *Neural Networks Versus Time Series Methods: a forecasting Exercise*, 14th International Symposium on Forecasting, Stockholm School of Econometrics, Stockholm, Sweden, 12-15 of June.
- [10] Wei, William, W.S. (1994), *Time Series Analysis: Univariate and Multivariate Methods*, Addison-Wesley Publishing Company Inc.