

Pembentukan model neural network untuk data time series dengan intervensi dan aplikasinya pada data IHK

(Formation of neural network models for time series data with intervention and its application in CPI data)

R. Kusumawati, D. U. Wutsqa, dan R. Subekti

Juridik Matematika, FMIPA, Universitas Negeri Yogyakarta (UNY),
Kampus Karangmalang, Sleman, DI Yogyakarta 55281
faks. (0274) 548203 dan e-mail: dhoriva@yahoo.com

diterima 25 Juli 2014, disetujui 30 September 2014

Abstrak

Tujuan penelitian ini adalah mendapatkan model peramalan yang optimal dari data IHK di Yogyakarta menggunakan model neural network yang dibangun dengan melibatkan pengaruh intervensi, dalam hal ini karena adanya kebijakan kenaikan harga BBM. Data IHK pendidikan, rekreasi, dan olah raga mempunyai pola seperti fungsi tangga, sehingga model RNN yang tepat adalah model dengan input sebagaimana pada model regresi spline polinomial *truncated*. Hasil prediksi model baik pada data training maupun testing menunjukkan tingkat keakuratan yang tinggi.

Kata kunci: neural network, intervensi, IHK

Abstract

The purpose of this study is to get an optimal forecasting model of CPI data in Yogyakarta using neural network models that is built with the involvement of the intervention effect. This case happens because of the policy of rising fuel prices. CPI education, recreation, and sport have a pattern like a step function, so that the exact RNN models are the models with the input as the truncated polynomial spline regression models. The results of model predictions both training and testing data showed high accuracy.

Key words: neural network, intervention, CPI

Pendahuluan

Indeks Harga Konsumen (IHK) merupakan salah satu indikator ekonomi di Indonesia. Badan yang mempunyai kewenangan untuk menghitung dan menerbitkannya adalah Badan Pusat Statistik (BPS). Pengertian dari IHK menurut BPS adalah suatu indeks yang menghitung rata-rata perubahan harga dalam suatu periode, dari suatu kumpulan harga barang dan jasa yang dikonsumsi oleh penduduk atau rumah tangga dalam kurun waktu tertentu. Jenis barang dan jasa tersebut dikelompokkan menjadi 7 kelompok, yaitu:

bahan makanan; makanan jadi, minuman, rokok, dan tembakau; perumahan; sandang; kesehatan; pendidikan, rekreasi dan olahraga; transpor dan komunikasi. Peramalan terhadap nilai IHK untuk waktu yang akan datang merupakan hal yang penting karena IHK berfungsi untuk mengetahui perubahan harga dari sekelompok tetap barang dan jasa yang pada umumnya dikonsumsi masyarakat, indeksasi upah (*wage-indexation*), tunjangan gaji pegawai, dan penyesuaian nilai kontrak (*contractual payment*). Hasil peramalan IHK juga dapat dijadikan sebagai pertimbangan bagi pemerintah untuk menentukan kebijakan-

kebijakan ekonomi yang sekiranya akan mempengaruhi harga barang dan jasa.

IHK merupakan data bulanan yang fluktuasi nilai tidak hanya dipengaruhi oleh nilai di masa lalu tetapi juga oleh faktor eksternal seperti kebijakan pemerintah menaikkan atau menurunkan harga bahan bakar minyak. Peramalan IHK dengan mempertimbangkan faktor eksternal berkaitan dengan peramalan time series dengan model intervensi. Model-model yang sering digunakan dalam analisis time series adalah model intervensi. Model khusus pada regresi *time series* yang dapat menjelaskan pola data yang berubah secara drastis adalah model intervensi [1]. Model intervensi ini merupakan gabungan dari fungsi intervensi dan *noise* yang mengikuti model *autoregressive moving average* (ARIMA) [2]. Aplikasi model intervensi ini untuk menjelaskan efek faktor eksternal telah dilakukan oleh para peneliti antara lain dalam [3] yang meneliti pengaruh embargo minyak Arab terhadap tingkat konsumsi listrik di United State, [4] menganalisa intervensi krisis ekonomi dan *travel warning* terhadap jumlah kedatangan wisman melalui bandara Juanda dan Ngurah Rai, dan [5] menyelidiki pengaruh bom Bali I pada tingkat hunian hotel berbintang lima di Bali.

Model intervensi adalah model yang dibangun melalui pendekatan parametrik, sehingga memerlukan asumsi-asumsi yang cukup ketat seperti stasioneritas dan normalitas. Di samping itu, proses identifikasi model cukup rumit, karena menggabungkan identifikasi model intervensi dan model ARIMA. Salah satu pendekatan yang fleksibel dan tidak memerlukan asumsi yang ketat adalah model *neural networks* (NN). Model NN tidak memerlukan asumsi-asumsi pada data yang seringkali sulit dipenuhi. Ada banyak model NN yang telah digunakan dalam pemodelan maupun peramalan data *time series*, diantaranya adalah *feed forward neural network* (FFNN) dan *recurrent neural network*. Yang termasuk dalam kelas FFNN adalah *back propagation neural network*, *radial basis function neural network*, *general regression neural network*. *Back propagation neural network* yang lebih dikenal sebagai FFNN merupakan model yang sangat populer dan banyak digunakan untuk menyelesaikan masalah *time series*, khususnya untuk data finansial.

Penggunaan FFNN untuk analisis data *time series* secara luas telah banyak dilakukan. Hasil-hasil studi empirik di bidang finansial menunjukkan bahwa prediksi data *time series* menggunakan model FFNN memberikan hasil yang lebih akurat dibandingkan model yang diperoleh dari metode statistik. [6] menggunakan model FFNN untuk memprediksi pertumbuhan ekonomi di Canada, Firdaus *et al.* memprediksi curah hujan di Johor, Malaysia [7], Chan *et al.* pada harga saham di Shanghai [8], dan Chen pada inflasi di Amerika Serikat [9]. Penerapan NN di bidang finansial dapat dilihat pada [10,11]. Tkacz [6] menggunakan NN untuk peramalan GDP Canada. Suhartono, Subanar, dan Sri Rejeki mengaplikasikan FFNN pada data penumpang pesawat dan memberikan hasil prediksi yang lebih baik dibandingkan dengan prediksi ARIMA dengan pola seasonal [12]. Wutsqa [13] menerapkan model NN pada data inflasi dan Wutsqa dan Abadi untuk peramalan kunjungan wisata candi Prambanan [14].

Penelitian tentang NN yang dipaparkan di atas tidak diterapkan pada data intervensi. Dengan demikian, masih menjadi pertanyaan terbuka bagaimana pemodelan NN untuk data intervensi. Oleh karena itu, dalam penelitian ini akan dikaji pembentukan model *neural networks* untuk data *time series* dengan intervensi. Dalam penerapannya akan digunakan data IHK yang secara teoritis dipengaruhi intervensi, yaitu kebijakan kenaikan harga BBM. Tujuan penelitian ini adalah mendapatkan model peramalan yang optimal dari data IHK di Yogyakarta berdasarkan metode yang dikembangkan dan meramalkan nilai IHK di Yogyakarta.

Hasil penelitian ini diharapkan dapat memberikan terobosan dalam pemodelan *time series*, khususnya pada model intervensi. Dalam terapannya akan menjadi alternatif untuk pemodelan data IHK. Dengan demikian, hasil penelitian ini diharapkan dapat menjadi referensi berkaitan dengan pemodelan data-data *time series*.

Metode Penelitian

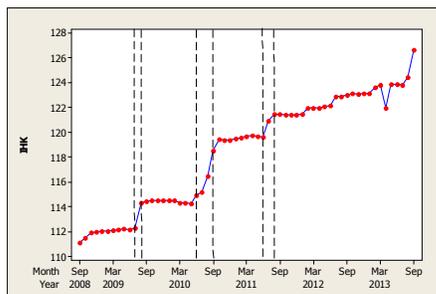
Data yang digunakan dalam penelitian ini adalah data sekunder, yaitu Data IHK kelompok pendidikan, rekreasi dan olahraga yang diambil dari Berita Resmi Statistik (BRS) dan pustaka

tercetak Badan Pusat Statistik (BPS) Daerah Istimewa Yogyakarta. Data IHK merupakan data bulanan, dan pada penelitian ini digunakan data bulan September 2008 sampai September 2013. Data IHK dimodelkan dengan NN.

Langkah-langkah dalam pembentukan model meliputi tahap identifikasi input, pembagian data menjadi data training dan testing, penentuan banyak neuron pada lapisan tersembunyi, penentuan input yang menghasilkan model optimal, penentuan bobot model, diagnostik model untuk melihat apakah error model sudah bersifat acak dengan plot ACF residual. Model yang baik adalah yang auto korelasi *errormya* tidak signifikan. Pemodelan data IHK dilakukan menggunakan beberapa paket program, yaitu MINITAB dan MATLAB.

Hasil dan Diskusi

Penelitian ini menggunakan data IHK kelompok pendidikan, rekreasi, dan olah raga, yang merupakan data *time series* sebanyak 61 periode. Untuk memodelkan data IHK tersebut, langkah pertama adalah identifikasi input model dengan mengguna plot *time series*. Berikut ini adalah plot *time series* data IHK tersebut.



Gambar 1. Plot *time series* data IHK kelompok pendidikan, rekreasi, dan olah raga.

Dari Gambar 1 terlihat bahwa data mengalami peningkatan yang cukup drastis pada periode 11, 12, 22, 25, 34, 36, sedang pola data antar periode tersebut cenderung konstan atau naik mengikuti trend linear. Oleh karena itu, input ditentukan berdasarkan titik-titik perubahan yang dalam model regresi spline polinomial *truncated* disebut dengan titik knot. Model yang digunakan pada data IHK ini adalah model RNN. Ada dua model yang dipertimbangkan dalam penelitian ini berdasarkan inputnya. Pertama

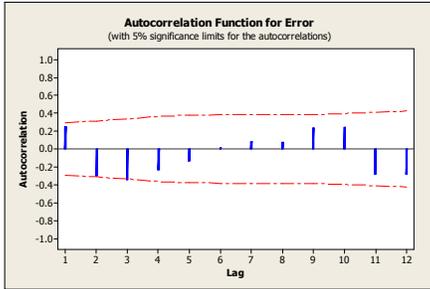
adalah variabel waktu $t = 1, \dots, 61$ sebagai input dan diberi notasi RNN1 Elman, dan yang kedua variabel waktu $t = 1, \dots, 61$ dan potongan polynomial (9) berdasarkan titik knot 11, 12, 22, 25, 34, 36 sebagai input, dan diberi notasi RNN2Elman. Tujuan menggunakan model RNN1 untuk menyelidiki seberapa baik RNN bekerja hanya dengan input t saja, dimana t tidak memperhatikan pola data yang menyerupai fungsi tangga, sedangkan pada RNN2 input t mengikuti perubahan yang pola data yang terpotong-potong atau tersegmentasi. Selanjutnya akan dijelaskan pembentukan kedua model mulai dari penentuan banyak neuron pada lapisan tersembunyi hingga peramalan.

Proses pembentukan model dilakukan dengan proses validasi silang dengan membagi data menjadi data training dan testing. Data training mulai periode 1 hingga periode 46, dan sisanya data testing (Lampiran 1 dan 2). Pembelajaran dilakukan untuk menentukan banyak neuron pada lapis tersembunyi dengan mempertimbangkan nilai MSE dan MAPE pada data training maupun testing. Hasil pembelajaran pada data training dan validasi pada data testing untuk model RNN1 Elman mulai dari neuron 1 hingga 8 disajikan pada Tabel 1.

Tabel 1. Nilai MSE dan MAPE hasil pembelajaran RNN Elman (untuk mencari neuron terbaik).

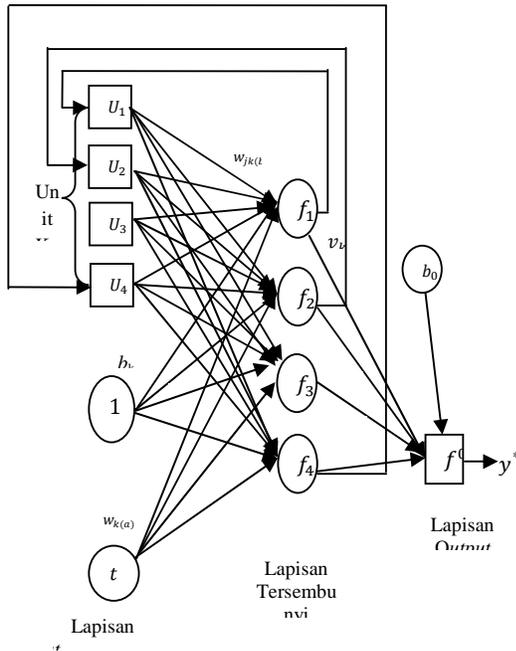
Neuron	Training		Testing	
	MSE	MAPE	MSE	MAPE
1	0.4767	0.4346	2.4113	1.3327
2	0.3936	0.4191	4.75	1.328
3	0.4755	0.4365	5.0005	1.3387
4	0.0993	0.1927	3.6098	0.6546
5	0.104	0.213	4.0239	0.6531

Berdasarkan hasil pada Tabel 1 ditentukan banyak neuron pada lapisan tersembunyi 4 neuron dengan lebih mempertimbangkan nilai MAPE baik pada data training dan data testing yang sangat kecil. Selanjutnya, dilakukan analisis residual untuk melihat apakah *error* bersifat acak dengan plot ACF berikut.



Gambar 2. Plot ACF residual model RNN1.

Dari Gambar 2 terlihat bahwa untuk semua lag, autokorelasi tidak signifikan atau *error* bersifat acak, sehingga dapat disimpulkan bahwa model RNN1 sudah memadai dan dapat digunakan untuk peramalan. Gambar 3 menyajikan arsitektur model RNN1 Elman dengan 4 neuron.

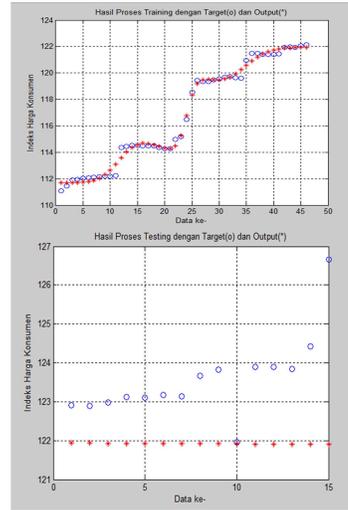


Gambar 3. Struktur jaringan model RNN1 Elman data IHK pendidikan, rekreasi, dan olah raga.

Persamaan matematis untuk model RNN1 jaringan Elman dapat dituliskan dalam persamaan berikut

$$y^* = \sum_{k=1}^4 v_{k1} \frac{1 - \exp\left(-\left(t_0 w_{1k(a)} + \sum_{j=1}^4 u_j w_{jk(b)} + b_k\right)\right)}{1 + \exp\left(-\left(t_0 w_{1k(a)} + \sum_{j=1}^4 u_j w_{jk(b)} + b_k\right)\right)} + b_0 \quad (1)$$

Hasil prediksi model baik pada data training maupun testing disajikan pada Gambar 4.



Gambar 4. Plot prediksi dan data asli training dan testing.

Model RNN1 dibentuk tanpa memperhatikan pola data yang cenderung naik dan membentuk seperti fungsi tangga dengan beberapa titik knot. Untuk itu, data IHK dimodelkan juga dengan input sebagaimana dijelaskan di atas dengan mempertimbangkan kecenderungan pola data dan titik knot. Tabel 2 memberikan hasil yang digunakan untuk menentukan banyak neuron pada lapisan tersembunyi.

Tabel 2. Nilai MSE dan MAPE hasil Pembelajaran RNN2 Elman.

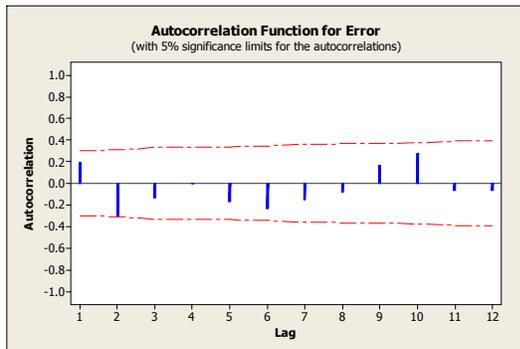
NEURON	Training		Testing	
	MSE	MAPE	MSE	MAPE
1	0.0812	0.191	4.2067	0.5856
2	0.1075	0.2221	4.9605	0.6813
3	0.0229	0.127	12.9287	0.3733
4	0.0298	0.142	2.3833	0.3742

Hasil pada Tabel 2 menunjukkan bahwa model dengan tingkat keakuratan tinggi adalah model RNN2 Elman dengan 3 neuron pada lapisan tersembunyi. Langkah berikutnya adalah penentuan input yang optimal, yang hasilnya disajikan pada table 3.

Tabel 3. Nilai MSE dan MAPE hasil pembelajaran RNN2 Elman (untuk mencari input yang optimal)

Input	Training		Testing	
	MSE	MAP	MSE	MAP
t0t1t2t3t4t5	0.053	0.148	0.950	0.710
t0t1t2	0.086	0.197	3.785	1.341
t0t1t2t3t4	0.064	0.162	1.279	3.523
t0t1t2t3t4t5	0.062	0.162	1.935	0.879

Berdasarkan nilai MAPE dan MSE baik untuk training maupun testing, dipilih model dengan input t0t1t2t3t4t5t6. Jika dibandingkan dengan model RNN1 Elman nilai MSE dan MAPE training maupun testing model RNN2 Elman ini cenderung lebih kecil, sehingga dalam penelitian ini model RNN2 Elman dipilih sebagai model peramalan nilai IHK pendidikan, rekreasi, dan olah raga. Untuk memastikan bahwa model RNN2 Elman ini layak, dilakukan uji keacakan *error*, yang hasilnya ditampilkan pada Gambar 5.

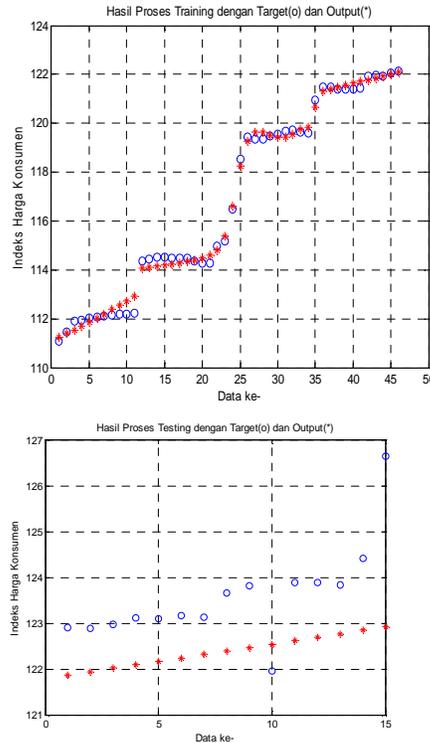


Gambar 5. Plot ACF residual model RNN2 Elman.

Plot ACF menunjukkan bahwa autokorelasi untuk semua lag tidak ada yang keluar dari garis signifikansi, sehingga *error* bersifat acak, yang berarti model sudah baik. Persamaan matematis untuk model RNN2 dengan arsitektur pada Gambar 6, adalah

$$y^* = \sum_{k=1}^3 v_{k1} \frac{1 - \exp\left(-\left(\sum_{i=0}^6 t_i w_{ik(a)} + \sum_{j=1}^3 u_j w_{jk(b)} + b_k\right)\right)}{1 + \exp\left(-\left(\sum_{i=0}^6 t_i w_{ik(a)} + \sum_{j=1}^3 u_j w_{jk(b)} + b_k\right)\right)} + b_0 \tag{2}$$

Hasil prediksi model baik pada data training maupun testing disajikan pada Gambar 7.



Gambar 7. Plot nilai prediksi dan data actual IHK pendidikan rekreasi dan olah raga di Yogyakarta.

Gambar 7 menunjukkan bahwa hasil prediksi model RNN2 Elman lebih baik dibandingkan RNN1 Eman.

Kesimpulan

Data IHK pendidikan, rekreasi, dan olah raga mempunyai pola seperti fungsi tangga, sehingga model RNN yang tepat adalah model dengan input sebagaimana pada model regresi spline polinomial *truncated*. Hasil prediksi model baik pada data training maupun testing menunjukkan tingkat keakuratan yang tinggi.

Daftar Pustaka

[1]. W. W. S. Wei, Time Series Analysis, Univariate and Multivariate Methods, 2nd ed., Pearson, New York, 2006.

- [2]. S. Makridakis, S. C. Wheelwright, dan V. E. McGee, Jilid 1 edisi kedua, terjemahan Ir. Hari Suminto, Metode dan Aplikasi Peramalan, Bina Rupa Aksara, Jakarta, 1999.
- [3]. D. C. Montgomery dan Weatherby, *AIIIE Transactions* December (1980) 289-307.
- [4]. E. Utami, Analisis Intervensi Krisis Ekonomi dan Travel Warning Terhadap Jumlah Kedatangan Wisman Melalui Bandara Juanda dan Ngurah Rai, Tugas Akhir S1 Statistika FMIPA ITS, Surabaya, 2001.
- [5]. Suhartono, *Jurnal MatStat* 7(2) (2007).
- [6]. G. Tkacz, *International Journal of Forecasting* 17 (2001) 57–69.
- [7]. N. Firdaus N., M. Shukor, A. Roselina, dan S. Nuradibah, Backpropagation Neural Network (BPNN) Model as a Solution of Short-Term Rainfall Prediction for Johor Catchment Area, IRCMSA Proceedings, Medan, Indonesia, 2005, pp. 365-375.
- [8]. M. Chan, C. Wong, dan C. Lam, Financial Time series Forecasting by using Conjugate Gradient Learning Algorithm and Multiple Linear Regression Weight Initialization (working paper), Department of Computing, The Hongkong Polytechnic University, Hongkong, 1999.
- [9]. X. Chen, J. Racine, dan N. R. Swanson, *IEEE Transaction on Neural Networks* 12 (4) (2001) 674-683.
- [10]. J. Moody, Economic Forecasting 5 Challenger and Neural Network Solutions, In Proceedings of the International Symposium on Artificial Neural Networks, Taiwan, 1995.
- [11]. G. Nikola G dan Jing Yang, The Application of Artificial Neural Networks to Exchange Rate Forecasting: The Role of Market Microstructure Variables (working paper), Financial Markets Department Bank of Canada, 2000.
- [12]. Suhartono, Subanar, Sri Rejeki, Feedforward Neural Networks Model for Forecasting Trend and Seasonal Time series, IRCMSA Proceedings, Sumatra Utara, Indonesia, 2005.
- [13]. D. U. Wutsqa, Comparison between the Neural Network (NN) and ARIMA Models for Forecasting the inflation in Yogyakarta, Proceeding ICAM, ITB, Bandung, 2005.
- [14]. D. U. Wutsqa dan A. M. Abadi, Modeling Islamic Lunar Calendar Effect in Tourism Data of Prambanan Templeby Using Neural Networks And Fuzzy Models, Presented in ICCT, Nidge, Turkey, 2005.