

MODEL REGRESI INTERVAL DENGAN NEURAL FUZZY UNTUK MEMPREDIKSI TAGIHAN AIR PDAM

Cucu Suhery¹⁾, Dedi Triyanto²⁾

^{1,2}Fakultas MIPA, Universitas Tanjungpura

e-mail: csuhery@siskom.untan.ac.id, email: dedi.triyanto@siskom.untan.ac.id

Abstract

Nilai interval tagihan air berlangganan melalui Perusahaan Daerah Air Minum (PDAM) dapat diprediksi melalui data tagihan sebelumnya dengan perhitungan menggunakan salah satu metode pada neuro-fuzzy, yaitu model regresi interval. Model regresi interval merupakan metode yang didasari oleh jaringan backpropagation. Pada model ini terdapat 2 jaringan backpropagation yang dilatih, satu model untuk mencari batas bawah dan model yang lain untuk mencari batas atas. Jaringan saraf tiruan merupakan salah satu metode soft computing yang banyak digunakan dan diterapkan pada analisis data runtun waktu. Sistem yang dibangun memiliki kemampuan untuk menerima input data berupa data tagihan PDAM beberapa periode sebelumnya, kemudian memprosesnya menggunakan model regresi interval dengan neural fuzzy, yang kemudian memberikan informasi hasilnya berupa interval nilai prediksi data tagihan langganan air pada PDAM.

Kata Kunci : regresi interval, neural fuzzy, peramalan, tagihan air

1. PENDAHULUAN

PDAM merupakan salah satu unit usaha milik daerah yang bergerak dalam distribusi air bersih bagi masyarakat. Tagihan PDAM merupakan tagihan rutin yang didapat oleh setiap warga yang berlangganan. Tagihan ini dibebankan kepada pelanggan setiap bulannya. Karena tagihan ini merupakan tagihan rutin setiap bulannya, maka masyarakat harus menyiapkan dana khusus untuk membayar tagihan tersebut. Namun, terkadang dana yang telah dialokasikan untuk membayar tagihan tidak sesuai dengan tagihan yang ada. Oleh karena itu diperlukan suatu sistem untuk memprediksi besarnya tagihan PDAM untuk periode yang akan datang.

Peramalan merupakan proses untuk memprediksi data atau kejadian yang mungkin akan terjadi di masa yang akan datang. Pemodelan data time series dilakukan ketika hanya terdapat satu variabel data yang dapat menjadi acuan, sehingga peramalan data masa yang akan datang dilakukan berdasarkan data pada masa sebelumnya. Model peramalan yang

memiliki akurasi tinggi dapat menghasilkan prediksi yang valid sehingga dapat menjadi alat bantu dalam mengambil keputusan.

Salah satu metode yang dapat digunakan untuk melakukan peramalan data *time series* adalah jaringan syaraf tiruan. Banyak peneliti yang telah melakukan penelitian terkait dengan penggunaan jaringan syaraf tiruan untuk meramalkan data time series. Pada umumnya penelitian dilakukan untuk mencari prediksi nilai-nilai tertentu. Selain itu terdapat juga penelitian yang menggunakan metode regresi interval yang menyelesaikan permasalahan kesehatan [1]

Pada penelitian ini akan dilakukan proses peramalan yang sedikit berbeda karena menggunakan model regresi interval, sehingga yang dihasilkan merupakan interval dari prediksi nilai mata uang.

Diharapkan dengan adanya penelitian ini maka diperoleh referensi model yang sedikit berbeda dari biasanya sehingga para pelanggan PDAM dapat memperoleh

interval prediksi tagihan PDAM pada masa yang akan datang. Pada penelitian ini akan digunakan data tagihan PDAM selama 5 tahun terakhir sebagai contoh kasusnya.

2. KAJIAN LITERATUR DAN PEGEMBANGAN HIPOTESIS

2.1 Model Regresi Interval dengan Neural Fuzzy

Model regresi interval merupakan model regresi yang menggunakan neural fuzzy. Misalkan diberikan pasangan input – output (x_k, d_k) , $k = 1, 2, \dots, p$ dengan $x_k = (x_{k1}, x_{k2}, \dots, x_{kn})$. Suatu model regresi fuzzy pada pola ke- k direpresentasikan sebagai [2]:

$$Y(x_k) = A_0 + A_1x_{k1} + \dots + A_nx_{kn} \quad (1)$$

Dengan A_i adalah bilangan fuzzy. Oleh karena itu, nilai estimasi output $Y(x_k)$ juga merupakan bilangan fuzzy. Analisis regresi fuzzy dapat disederhanakan menjadi analisis regresi interval dimana model regresi interval nantinya akan dibentuk [2]. Model regresi fuzzy merupakan pengembangan dari regresi klasik dimana beberapa elemen seperti masukan atau keluaran atau keduanya merupakan bilangan fuzzy. [3]

Konsep dasar dari analisis regresi interval yang didasarkan pada jaringan backpropagation, diperkenalkan oleh Ishibuchi dan Tanaka (1992). Model tersebut menggunakan 2 jaringan backpropagation. Satu jaringan digunakan untuk batas atas interval, sedangkan satu jaringan lainnya digunakan untuk batas bawah interval data. Kedua jaringan tersebut dilatih secara terpisah [4].

Misalkan $g^+(x_k)$ dan $g^-(x_k)$ adalah *output* dari kedua jaringan backpropagation tersebut (BPN⁺ dan BPN⁻) yang berhubungan dengan *input* vektor x_k , dimana setiap jaringan memiliki n neuron pada lapisan input dan 1 neuron

pada lapisan output. Proses pembelajaran dilakukan terhadap kedua jaringan (BPN⁺ dan BPN⁻) untuk mendapatkan *output* jaringan $g^+(x_k)$ dan $g^-(x_k)$ yang berkaitan dengan kondisi sebagai berikut [2]:

$$g^-(x) \leq d_k \leq g^+(x), k = 1, 2, \dots, p \quad (2)$$

Pada proses pembelajaran BPN⁺, fungsi biaya yang digunakan adalah:

$$E = \sum_{k=1}^p E_k = \frac{1}{2} \sum_{k=1}^p \alpha_k [d_k - g^+(x_k)]^2 \quad (3)$$

Dimana nilai α_k diberikan sebagai berikut:

$$\alpha_k = \begin{cases} 1; & d_k > g^+(x_k) \\ \alpha; & d_k \leq g^+(x_k) \end{cases} \quad (4)$$

Dimana α merupakan bilangan positif yang cukup kecil pada interval (0,1). Nilai α ini dapat diberikan dengan menggunakan fungsi penurunan, sebagai:

$$\alpha(t) = \frac{1}{[1 + (\frac{t}{2000})]^3} \quad (5)$$

Dimana nilai t merupakan nilai dari iterasi yang sedang dilakukan.

Aturan pembelajaran backpropagation digunakan untuk mendapatkan bobot-bobot w_j (bobot antara lapisan tersembunyi dengan lapisan output) dan w_{ji} (bobot antara lapisan input dengan lapisan tersembunyi) pada jaringan BP⁺, dan perbaikan bobot adalah sebagai berikut [5]:

$$\Delta w_j = \eta \left(-\frac{\partial E_k}{\partial w_j} \right) = \eta \delta_k y_{kj} \quad (6)$$

$$\Delta w_{ji} = \eta \left(-\frac{\partial E_k}{\partial w_{ji}} \right) = \eta \delta_{kj} x_{ki} \quad (7)$$

Dimana nilai dari δ_k dan δ_{kj} adalah [4]:

$$\delta_k = \alpha_k (d_k - y_k) y_k (1 - y_k) \quad (8)$$

$$\delta_{kj} = y_{kj} (1 - y_{kj}) \delta_k w_j \quad (9)$$

$Y_k = g^+(x_k)$ adalah output jaringan, dan y_{kj} adalah output neuron ke- j pada lapisan tersembunyi untuk input x_k [5].

Cara yang sama juga digunakan untuk melakukan pembelajaran pada jaringan BPN untuk mendapatkan output jaringan $g^-(x_k)$. Fungsi biaya yang digunakan untuk pembelajaran ini seperti terlihat pada persamaan (3), dengan α_k diberikan sebagai berikut [5]:

$$\alpha_k = \begin{cases} \alpha; & d_k \geq g^-(x_k) \\ 1; & d_k < g^-(x_k) \end{cases} \quad (10)$$

Dimana α adalah bilangan positif yang cukup kecil pada interval (0,1).

Kedua algoritma pembelajaran tersebut digunakan untuk menentukan 2 fungsi, $g^+(x)$ dan $g^-(x)$ dimana $g^-(x) \leq d_k \leq g^+(x)$, $k = 1, 2, \dots, p$. Dari sini, dapat diperoleh interval [5]:

$$G(x) = [g^-(x), g^+(x)] \quad (11)$$

Karena $g^+(x)$ dan $g^-(x)$ diperoleh dari pembelajaran yang terpisah, maka sangat dimungkinkan $g^+(x) < g^-(x)$. Sehingga, jika hal tersebut terjadi, maka dapat dilakukan modifikasi pada interval tersebut sebagai berikut [5]:

$$h^-(x) = \begin{cases} g^-(x); & g^-(x_k) \leq g^+(x) \\ \frac{1}{2}(g^-(x) + g^+(x)); & g^-(x_k) > g^+(x) \end{cases} \quad (12)$$

$$h^+(x) = \begin{cases} g^+(x); & g^-(x_k) \leq g^+(x) \\ \frac{1}{2}(g^-(x) + g^+(x)); & g^-(x_k) > g^+(x) \end{cases} \quad (13)$$

Berdasarkan persamaan (12) dan (13) maka interval yang terjadi adalah [5]:

$$G(x) = [h^-(x), h^+(x)] \quad (14)$$

2.2 Perusahaan daerah Air Minum (PDAM)

PDAM merupakan perusahaan milik negara yang merupakan pusat naungan masyarakat daerah dalam melakukan aktivitas sehari-hari dalam penyediaan air bersih. Setiap bulan terjadi transaksi pembayaran penggunaan air oleh pelanggan. PDAM melakukan pencatatan pemakaian yang dilakukan setiap bulannya, dan konsumen wajib melakukan pembayaran tagihan setiap

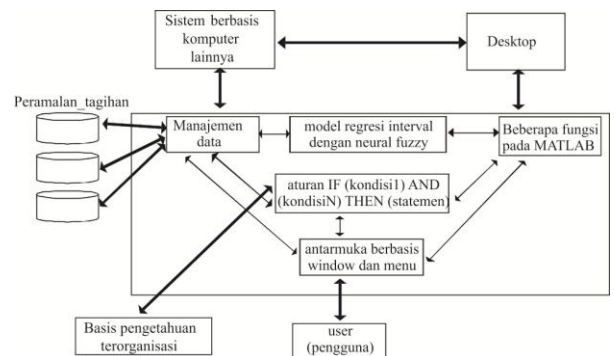
bulan. Tagihan PDAM merupakan tagihan rutin bulanan yang harus dibayar oleh masyarakat yang berlangganan.

3. METODE PENELITIAN

3.1 Gambaran Umum Model

Model yang diusulkan pada penelitian ini adalah sebuah sistem pendukung keputusan. Model ini menggunakan pendekatan kuantitatif dengan memanfaatkan data-data yang diambil dari tagihan listrik, khususnya selama bulan Januari 2010 sampai dengan bulan Desember 2014.

Komponen-komponen pembentuk sistem pendukung keputusan digunakan dengan pendekatan dari Turban. Komponen-komponen tersebut terdiri atas manajemen data, manajemen model, subsistem berbasis pengetahuan, dan antarmuka [6]. Diagram model yang diusulkan dapat dilihat pada gambar 1:



Gambar 1. Model yang diusulkan

Keterangan:

1. Peramalan_tagihan : merupakan nama basis data pada sistem.
2. Manajemen_data: berisi perintah-perintah yang digunakan untuk memanipulasi data yang ada pada basis data.
3. Model regresi interval dengan neural fuzzy: merupakan manajemen model yang digunakan pada sistem ini.
4. Aturan IF (kondisi) AND (kondisi) THEN (statemen): merupakan aturan umum yang terdapat di dalam basis pengetahuan pada sistem. Adapun aturan yang terdapat pada sistem adalah IF (data1=a) AND (data2=b) AND (data3=c) AND (data4=d) AND

(data5=e) AND (data6=f) THEN
 (data7=g).

5. Beberapa fungsi pada MATLAB: fungsi yang digunakan pada pada sistem yang diperoleh dari *tools* pemrograman MATLAB, antara lain: max, min, rand, size, dan load.
6. Antarmuka berbasis *window* dan menu: merupakan jenis antarmuka yang digunakan untuk membangun sistem ini sehingga dapat digunakan dengan mudah oleh *user*.
7. Desktop: merupakan basis pembangunan sistem ini.
8. Pengguna: semua *user* yang menggunakan sistem ini, yaitu administrator dan *end-user*.

Data sampel yang digunakan pada pelatihan merupakan pasangan data yang telah ditetapkan, dimana sampel yang pertama memiliki nilai *input* data ke-1 (satu) sampai data ke-5 (lima) dan nilai *output* adalah data ke-6 (enam). Kemudian sampel kedua memiliki nilai *input* data ke-2 (dua) sampai data ke-6 (enam) dan nilai *output* adalah data ke-7 (tujuh). Pola ini secara terus-menerus digunakan untuk membentuk data pelatihan hingga data ke-60, sehingga diperoleh 55 pasang data *input-output*.

Pada Tabel 1 dapat dilihat contoh pasangan data yang digunakan pada sistem prediksi yang dibangun.

Tabel 1. Contoh pasangan data sampel

Input 1	Input 2	Input 3	Input 4	Input 5	Output
154.526	183.952	183.862	177.474	179.179	160.888
183.952	183.862	177.474	179.179	160.888	154.071
183.862	177.474	179.179	160.888	154.071	159.047
177.474	179.179	160.888	154.071	159.047	194.509
179.179	160.888	154.071	159.047	194.509	155.479
160.888	154.071	159.047	194.509	155.479	159.929
154.071	159.047	194.509	155.479	159.929	163.948
159.047	194.509	155.479	159.929	163.948	163.441
194.509	155.479	159.929	163.948	163.441	168.463
155.479	159.929	163.948	163.441	168.463	167.860
159.929	163.948	163.441	168.463	167.860	154.412
163.948	163.441	168.463	167.860	154.412	162.070
163.441	168.463	167.860	154.412	162.070	178.231

168.463	167.860	154.412	162.070	178.231	152.792
167.860	154.412	162.070	178.231	152.792	186.458
154.412	162.070	178.231	152.792	186.458	157.323
162.070	178.231	152.792	186.458	157.323	153.749
178.231	152.792	186.458	157.323	153.749	188.742
152.792	186.458	157.323	153.749	188.742	179.526
186.458	157.323	153.749	188.742	179.526	153.675

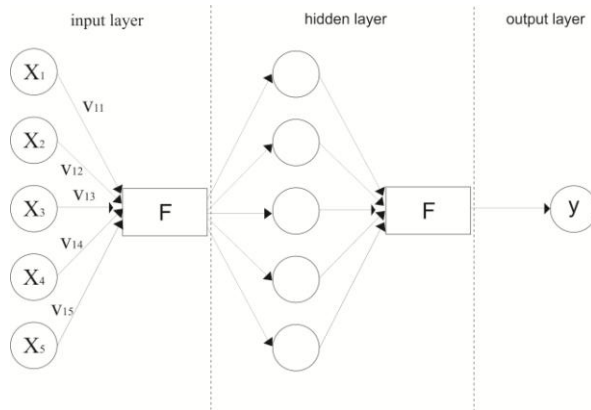
Sebelum melakukan proses prediksi, sistem yang ada harus dilatih terlebih dahulu agar dapat memperoleh nilai bobot-bobot hasil pelatihan yang nantinya akan digunakan untuk proses prediksi. Pada proses pelatihan, sistem memerlukan beberapa nilai parameter untuk pelatihan, yaitu nilai *goal*, laju pembelajaran, maksimum epoch, jumlah lapisan tersembunyi dan jumlah neuron pada setiap lapisan tersembunyi. Apabila pelatihan telah dilakukan, maka diperoleh bobot-bobot pelatihan yang kemudian disimpan dalam suatu file penyimpanan. Adapun arsitektur pelatihan yang digunakan dapat dilihat pada Gambar 2.

Setelah proses pelatihan dilakukan maka sistem dapat digunakan untuk melakukan prediksi. Untuk melakukan prediksi, sistem cukup diberikan masukan berupa data untuk tagihan PDAM selama 5 bulan terakhir. Setelah memasukkan data maka sistem dapat menghasilkan keluaran berupa nilai peramalan berupa interval untuk tagihan PDAM pada bulan ke-6.

4. HASIL DAN PEMBAHASAN

Berdasarkan langkah-langkah prediksi yang telah dijelaskan sebelumnya, metode regresi interval digunakan sebagai metode prediksi pada data tagihan PDAM. Jumlah data yang direkam pada bulan Januari 2011 sampai dengan Desember 2014 sebanyak 60 yang diubah menjadi 55 pasang data. Dari 55 pasang data tersebut, 45 data dijadikan sebagai data pelatihan dan 10 data menjadi data pengujian. Data pengujian ini yang nantinya digunakan untuk mengukur tingkat akurasi sistem.

Adapun arsitektur pelatihan yang digunakan dapat dilihat pada Gambar 2.

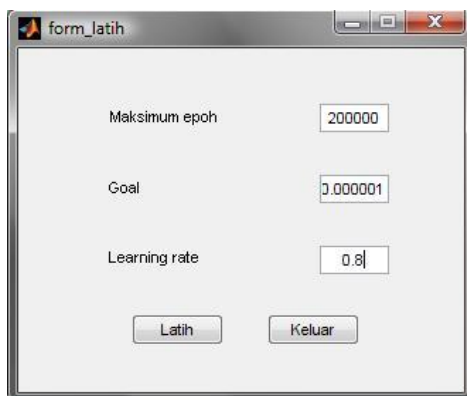


Gambar 2. Arsitektur jaringan untuk pelatihan

Setelah data melalui proses pelatihan, maka akan diperoleh bobot akhir dari pelatihan, dimana bobot akhir ini yang nantinya akan digunakan pada proses pengujian untuk memperoleh hasil akhir apabila ada data yang akan diujikan.

4.1 Tampilan Hasil

Untuk memperoleh hasil prediksi maka sistem harus dilatih terlebih dahulu agar dapat memperoleh bobot yang akan digunakan pada proses pengujian. Adapun halaman pengisian parameter pelatihan dapat dilihat pada Gambar 3.



Gambar 3. Tampilan halaman pelatihan

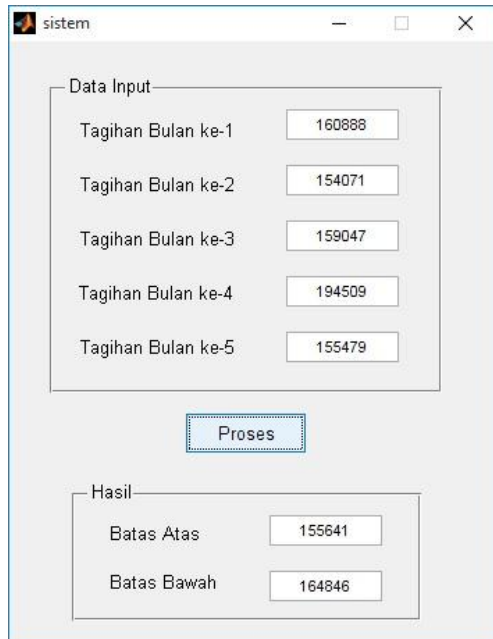
Pada proses pelatihan yang telah dilakukan digunakan parameter pelatihan maksimum perulangan 200.000, goal 0,000001 *learning rate* 0,8.

Setelah melakukan proses pelatihan, maka selanjutnya sistem dapat digunakan untuk proses pengujian. Proses pengujian dilakukan oleh pengguna dengan cara memasukkan data tagihan PDAM pada 5 bulan terakhir. Adapun tampilan dari halaman sistem untuk melakukan pengujian dapat dilihat pada Gambar 4.



Gambar 4. Tampilan halaman penggunaan sistem

Setelah pengguna memberikan semua data yang diperlukan, maka data tersebut akan diproses bersama nilai bobot yang telah diperoleh dari proses pelatihan. Setelah melalui proses perhitungan, maka sistem akan mengeluarkan hasil batas bawah dan batas atas dari nilai prediksi untuk tagihan PDAM pada bulan selanjutnya. Gambar 5 merupakan tampilan sistem yang telah melalui proses perhitungan dan memberikan hasil akhir kepada pengguna. Dapat dilihat pada Gambar 5 bahwa sistem memberikan interval nilai prediksi untuk tagihan PDAM di bulan berikutnya.



Gambar 5. Tampilan hasil sistem

4.2 Pengujian

Proses pengujian dilakukan setelah sistem siap digunakan. Pengujian dilakukan terhadap 10 data yang belum pernah dilatih sebelumnya. Hasil dari pengujian dapat dilihat pada Tabel 2.

Tabel 2. Hasil pengujian 10 data uji

Data ke-	Batas Atas	Batas Bawah	Target	Keterangan
1	185.811	193.703	190.284	T
2	186.824	192.693	193.930	F
3	151.821	158.712	153.075	T
4	158.821	167.307	161.062	T
5	186.717	196.711	195.950	T
6	168.164	176.738	173.204	T
7	174.543	183.745	179.584	T
8	168.745	175.128	173.080	T
9	150.823	156.704	153.872	T
10	176.538	185.718	182.793	T

Dari Tabel 2 dapat dilihat bahwa terdapat 1 data uji yang nilai targetnya tidak berada diantara batas bawah dan batas atas sistem, hal ini menandakan bahwa data tersebut mengalami kesalahan pada saat proses prediksi. Sedangkan 9 data yang lain berhasil diramal dimana nilai target berada diantara nilai batas bawah dan batas atas

5. KESIMPULAN

Berdasarkan hasil penelitian yang telah berhasil diperoleh, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Metode peramalan regresi interval yang menggunakan 2 jaringan backpropagation dapat digunakan untuk prediksi data tagihan PDAM dan memberikan hasil yang cukup baik.
2. Arsitektur jaringan backpropagation dengan 5 node input, 1 lapisan tersembunyi dengan 5 node hidden dapat memberikan hasil peramalan yang baik pada data yang diuji.
3. Pengujian terhadap 10 data baru memperoleh hasil 9 data berhasil diramal dengan benar dan 1 data tidak berhasil diramal

6. REFERENSI

- [1] Rismawan, Tedy dan Kusumadewi, Sri. 2008. "Sistem Pendukung Keputusan Penentu Nilai Interval Kadar Lemak Tubuh Menggunakan Regresi Interval dengan Neural Fuzzy". Makalah disampaikan pada Seminar Nasional Teknoin 2008. Yogyakarta
- [2] Lin, Chin-Teng; dan Lee, George. 1996. "Neural Fuzzy Systems". London: Prentice-Hall.
- [3] Ghodsi, R., Zakerinia, M. S., Jekar, M. 2011. *Neural Network and Fuzzy Regression Model for Forecasting Short Term Price in Ontario Electricity Market*. 41st International Conference on Computers & Industrial Engineering. 25 Oktober 2011. Los Angeles, California, USA. P. 954-959.
- [4] Jang, JSR; Sun, CT; dan Mizutani, E. 1997. "Neuro-Fuzzy and Soft Computing". London: Prentice-Hall.
- [5] Kusumadewi, Sri dan Hartati, Sri. 2006. "Neuro-Fuzzy Integrasi Sistem Fuzzy dan Jaringan Syaraf". Yogyakarta: Graha Ilmu.

[6] Turban, Efraim; Aronson, Jay, E.;
Liang, Ting-Peng. 2005. "Decision
Support Systems and Intelligent

Systems". International Edition, Edisi
7. New Jersey: Pearson Prentice-
Hall Education International