

Penerapan Jaringan Syaraf Tiruan Untuk Memprediksi Jumlah Pengangguran di Provinsi Kalimantan Timur Dengan Menggunakan Algoritma Pembelajaran Backpropagation

M.F. Andrijasa* dan Mistianingsih*

*Program Studi Ilmu Komputer, FMIPA Universitas Mulawarman
Jl. Barong Tongkok no.5 Kampus Unmul Gn. Kelua Sempaja Samarinda 75119*

Abstrak

Prediksi pengangguran sebelumnya menggunakan ilmu statistika sosial, perhitungan statistika modern banyak dilakukan oleh komputer, dan bahkan beberapa perhitungan hanya dapat dilakukan oleh komputer berkecepatan tinggi yaitu Jaringan Saraf Tiruan. Penerapan JST untuk memprediksi jumlah pengangguran di Provinsi Kalimantan Timur dengan menggunakan algoritma pembelajaran backpropagation berdasarkan pada data survey lapangan sebagai tolok ukur perlu tidaknya ditinjau kembali peningkatan lapangan pekerjaan yang bersesuaian. Dengan menggunakan beberapa variabel masukan yaitu: data tahun 2004 – 2008. Algoritma pembelajaran backpropagation diterapkan untuk pelatihan kedelapan variabel tersebut, sehingga diperoleh keluaran prediksi jumlah pengangguran ditahun berikutnya. Dari hasil pengujian metode JST- Backpropagation dapat memprediksi jumlah pengangguran dan target output yang diinginkan lebih mendekati ketepatan dalam melakukan pengujian, karena terjadi penyesuaian nilai bobot dan bias yang semakin baik pada proses pelatihan. Hasil pengujian di peroleh prediksi jumlah pengangguran di tahun 2009 adalah 133.104. Sedangkan hasil prediksi pengangguran tahun 2009 yang dilakukan oleh BPS Provinsi Kalimantan Timur adalah 139.830.

Kata kunci : JST, Prediksi Pengangguran, dan Algoritma Backpropagation.

* Jurusan Teknologi Informasi, Politeknik Negeri Samarinda

1. PENDAHULUAN

Banyak problem-problem menarik dalam ilmu pengetahuan yang salah satunya dapat digolongkan ke dalam peramalan (forecasting). Pada penerapannya, peramalan biasanya digunakan untuk aplikasi peramalan besarnya penjualan, prediksi nilai tukar uang, prediksi besarnya aliran air sungai, dan lain-lain. Peramalan dapat dilakukan dengan berbagai macam cara, salah satunya adalah dengan mengembangkan teknik kecerdasan buatan, yang dalam hal ini yang paling banyak digunakan untuk maksud di atas adalah menggunakan Artificial Neural Network (ANN), Jaringan Saraf Tiruan, (JST).

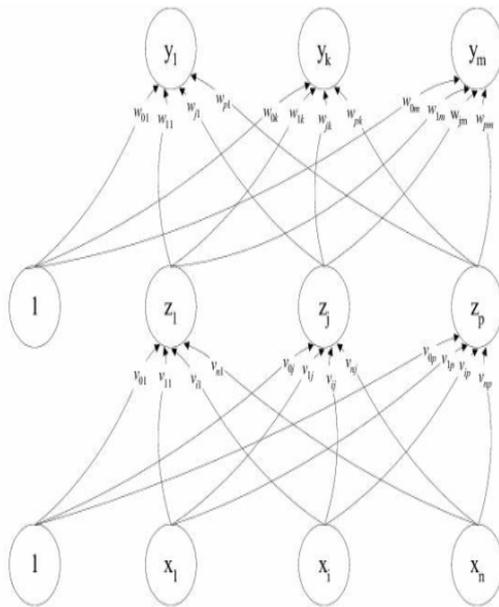
Konsep praktis dari JST untuk memprediksi jumlah pengangguran dengan cara pola data pengangguran periode masa lalu yang di masukkan kedalam sistem dilakukan proses pelatihan menggunakan Jaringan Saraf Tiruan (JST) dan algoritma pembelajaran Backpropagation. Setelah dilakukan proses pelatihan, sistem akan menghasilkan bobot-bobot yang akan digunakan untuk memprediksi jumlah pengangguran pada periode tahun-tahun selanjutnya.

2. JARINGAN SARAF TIRUAN

Jaringan saraf Tiruan adalah merupakan salah satu representasi buatan dari otak manusia yang selalu mencoba untuk mensimulasikan proses pembelajaran pada otak manusia tersebut. Istilah buatan disini digunakan karena jaringan saraf ini diimplementasikan dengan menggunakan program komputer yang mampu menyelesaikan sejumlah proses perhitungan selama proses pembelajaran. (Kusumadewi, 2003).

2.1 Algoritma Pembelajaran Back-propagation

Kunggulan yang utama dari sistem JST adalah kemampuan untuk "belajar" dari contoh yang diberikan. Backpropagation merupakan algoritma pembelajaran yang terawasi dan biasanya digunakan oleh perceptron dengan banyak layer lapisan untuk mengubah bobot-bobot yang ada pada lapisan tersembunyinya. Backpropagation adalah pelatihan jenis terkontrol (supervised) dimana menggunakan pola penyesuaian bobot untuk mencapai nilai kesalahan yang minimum antara keluaran hasil prediksi dengan keluaran yang nyata. (F.Suhandi, 2009).



Gambar 1. Arsitektur Jaringan Multilayer Backpropagation Dengan Satu Hidde Layer

Penulisan istilah:

x : Vektor input pembelajaran.

$X = (x_1, x_2, \dots, x_j, \dots, x_n)$.

t : Vektor target output.

$T = (t_1, t_2, \dots, t_k, \dots, t_m)$.

δ_k : Bagian koreksi error penyesuaian bobot w_{jk} berpedoman pada error output neuron y_k .

δ_j : Bagian koreksi error penyesuaian bobot v_{ij} berpedoman pada error output neuron z_j .

α : Learning rate.

θ : Toleransi error.

x_i : Input neuron.

v_{0j} : Bias pada Hidden neuron ke- j .

z_i : Hidden neuron ke- j .

Nilai input z_i ditunjukkan dengan

$$z_in_j = v_{0j} + \sum_{i=1}^n x_i v_{ij}$$

z_in_i :

Nilai output dengan menggunakan fungsi aktivasi yang dipilih ditunjukkan dengan z_i :

$$z_j = f(z_in_j)$$

w_{0k} : Bias pada output neuron ke- k .

Y_k : Output neuron ke- k .

Nilai input Y_k ditunjukkan dengan

y_in_k :

$$y_in_k = w_{0k} + \sum_{j=1}^p z_j w_{jk}$$

Nilai output dengan menggunakan fungsi aktivasi yang dipilih ditunjukkan dengan y_k :

$$y_k = f(y_in_k)$$

Algoritma pembelajaran untuk JST backpropagation adalah sebagai berikut : (F. Suhandi, 2009)

Langkah 0 : Inisialisasi bobot dengan nilai random atau acak yang cukup kecil.

Set *learning rate* α ($0 < \alpha \leq 1$)

Langkah 1 : Selama kondisi berhenti masih belum terpenuhi, laksanakan langkah sampai 9.

Langkah 2 : Untuk tiap pasangan pelatihan, kerjakan langkah 3 sampai 8.

Feedforward:

Langkah 3 : Untuk tiap input neuron (X_i , $i=1,2,3,\dots,n$) menerima input x_i dan menyebarkan sinyal tersebut ke seluruh neuron kepada lapisan atasnya (lapisan tersembunyi).

Langkah 4 : Untuk hidden neuron (Z_i , $j=1,2,3,\dots,p$) dihitung nilai input dengan menggunakan nilai bobotnya:

$$z_in_j = v_{0j} + \sum_{i=1}^n x_i v_{ij}$$

kemudian dihitung nilai output dengan menggunakan fungsi aktivasi yang digunakan:

$$z_j = f(z_in_j)$$

dimana fungsi aktivasi yang digunakan adalah fungsi sigmoid biner yang mempunyai persamaan:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Hasil fungsi tersebut dikirim kesemua neuron pada lapisan diatasnya.

Langkah 5 : Untuk tiap output neuron (Y_k , $k=1,2,3,\dots,m$) dihitung nilai input dengan nilai bobotnya: $y_in_k = w_{0k} + \sum_{j=1}^p z_j w_{jk}$

Kemudian dihitung nilai output dengan menggunakan fungsi aktivasi:

$$y_k = f(y_in_k)$$

Backpropagation:

(Perhitungan nilai kesalahan):

Langkah 6 : Untuk tiap output neuron (Y_k , $k=1,2,3,\dots,m$) menerima pola target yang bersesuaian dengan pola input dan kemudian dihitung informasi kesalahan:

$$\Delta_k = (t_k - y_k) f'(y_in_k)$$

Kemudian dihitung koreksi nilai bobot yang kemudian akan digunakan untuk memperbaharui nilai w_{jk} : $\Delta w_{jk} = \alpha \delta_k z_j$

Hitung koreksi nilai bias yang kemudian akan digunakan untuk memperbaharui nilai

$$w_{0k} : \Delta w_{0k} = \alpha \delta_k$$

Dan kemudian nilai δ_k dikirim ke neuron pada lapisan sebelumnya.

Langkah 7 : Untuk tiap hidden neuron (Z_j , $j=1,2,3,\dots,p$) dihitung delta input yang berasal dari neuron pada layer diatasnya:

$$\delta_in_j = \sum_{k=1}^m \delta_k w_{jk}$$

Kemudian nilai tersebut dikalikan dengan nilai turunan dari fungsi aktivasi untuk menghitung informasi kesalahan: $\delta_j = \delta_{in_j} \cdot f'(z_{in_j})$

Hitung koreksi nilai bobot yang kemudian digunakan untuk memperbaharui

v_{ij} :

$$\Delta v_{ij} = \alpha \delta_j x_i$$

Dan hitung nilai koreksi bias yang kemudian digunakan untuk memperbaharui nilai

$$v_{0j} : \Delta v_{0j} = \alpha \delta_j$$

Memperbaharui nilai bobot dan nilai bias:

Langkah 8 : Tiap nilai bias dan bobot ($j=0, \dots, p$) pada output neuron (Y_k , $k=1, 2, 3, \dots, m$) diperbaharui:

$$w_{jk} \text{ (baru)} = w_{jk} \text{ (lama)} + \Delta v_{ij}$$

$$v_{ij} \text{ (baru)} = v_{ij} \text{ (lama)} + \Delta v_{ij}$$

Langkah 9 : Menguji apakah kondisi berhenti sudah terpenuhi. Kondisi berhenti ini terpenuhi jika nilai kesalahan yang dihasilkan lebih kecil dari nilai kesalahan referensi.

3. PERANCANGAN SISTEM

3.1 Perancangan Data Input – Output

Data yang digunakan adalah data tahunan dari tahun 2004 sampai dengan 2008, yaitu:

Data yang masuk dengan 8 input yang berasal dari:

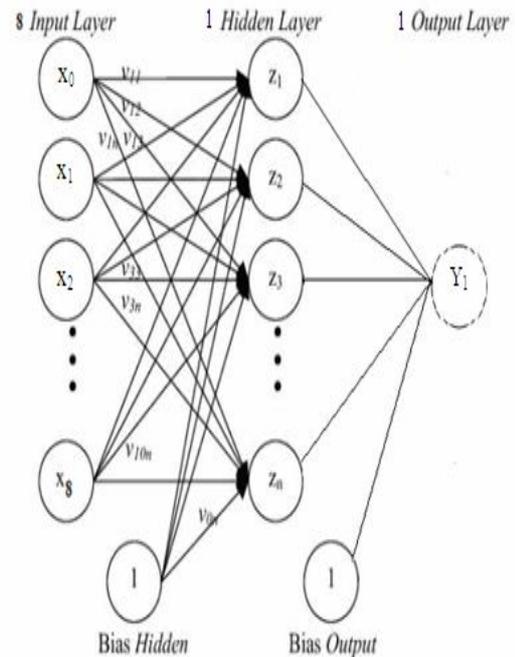
1. Data tahun 2004 – 2008.
2. Penduduk menurut kelompok umur dan jenis kelamin.
3. Data historis TPAK (Tingkat Partisipasi Angkatan Kerja) menurut kelompok umur.
4. Angkatan kerja.
5. Tingkat pengangguran terbuka (TPT.)
6. Penduduk yang bekerja menurut lapangan usaha.
7. PDRB (Produk Domestik Regional Bruto) atas dasar harga konstan 2000 menurut lapangan usaha.
8. Data historis penduduk usia 15 tahun keatas, status, pendidikan dan pekerjaan.

Data asli yang diperoleh dari BPS Provinsi Kalimantan Timur (Lihat Lampiran)

3.2 Perancangan Arsitektur Sistem

Jumlah *hidden* ditentukan dengan cara *trial and error*, dalam arti hasil pembelajaran yang tercepat dan terbaik itulah yang akan menentukan jumlah *hidden layer* tersebut. Mengenai jumlah banyaknya *hidden layer* yang dibutuhkan, tidak ada ketentuan khusus karena tidak ada teori yang dengan pasti dapat dipakai. Adapun perancangan jaringan saraf tiruan

backpropagation untuk memprediksi jumlah pengangguran adalah:



Gambar 2. Rancangan Arsitektur Jaringan Saraf Tiruan Untuk Prediksi Pengangguran

4. PENGUJIAN DAN ANALISIS

Dari hasil perancangan yang dibuat dihasilkan training yang terbaik yaitu mempunyai hasil ketepatan atau kebenaran yang cukup baik dan pembelajaran yang cukup baik. Dari hasil pengujian yang dilakukan didapat jumlah iterasi maksimal yang dicapai mencapai 500000 iterasi, dan diperlukan pembelajaran 13 jam terhadap seluruh pola yang diberikan. Selain hal tersebut juga ditetapkan konstanta pembelajaran terbaik adalah 0.01 dengan toleransi kesalahan $1e-05$ dan lapisan tersembunyi 1 layer dengan jumlah neuron 10.

Setelah dilakukan training dan testing pola-pola yang dilatih, diperoleh hasil bahwa test terhadap pola-pola tersebut benar (akurat).

Rata – rata *error* jaringan pada proses testing adalah:

$$RMSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{N}$$

$$RMSE = \frac{\sum_{i=1}^1 (0,74779 - 0,74779)^2}{8} = 13,7001931$$

Dimana :

y_i = Nilai aktual data 0,74779.

Y_i = Nilai hasil prediksi 0,74779.
 N = Jumlah nilai *testing* 8.

Target error $1e-05$, sedangkan $RMSE \geq$ target.

Denormalisasi testing:

$$x_i = y (x_{\max} - x_{\min}) + x_{\min}$$

$$\begin{aligned} 2009 &= 0,74779 (177.997 - 96236) + \\ 96236 &= 133.104 \end{aligned}$$

dimana:

x_i = Jumlah pengangguran tahun 2009
 y = Hasil output jaringan 0,74779
 x_{\min} = Data dengan nilai minimum 96236
 x_{\max} = Data dengan nilai maksimum 177.997

5 ANALISIS

Karena jumlah data yang besar mengakibatkan lamanya proses Training. Setelah dilakukan pelatihan dengan maksimum epoch 500000 dengan neuron pada hidden layer 10 neuron. Untuk training 3 neuron membutuhkan waktu kurang lebih 10 jam dan berhenti pada epoch 500000. Untuk training 10 neuron membutuhkan waktu kurang lebih 11 jam dan berhenti pada epoch 500000. Dari hasil training model jaringan saraf tiruan yang telah dibuat dengan konstanta belajar 0,1, keberhasilan dalam memprediksi jumlah pengangguran dengan RMSE sebesar 13.7001521. Jadi setelah proses training dan testing dapat diprediksi jumlah pengangguran pada tahun 2009 adalah 133.104

6. KESIMPULAN DAN SARAN

Dari hasil penelitian, dapat disimpulkan bahwa metode *Backpropagation* dapat digunakan untuk melakukan prediksi pengangguran, maupun kasus yang memiliki data masa lalu, dan dengan menggunakan metode *Backpropagation*, target output yang diinginkan lebih mendekati ketepatan dalam melakukan pengujian, karena terjadi penyesuaian nilai bobot dan bias yang semakin baik pada proses pelatihan. Diperoleh hasil prediksi jumlah pengangguran pada tahun 2009 adalah 133.104. Sedangkan hasil prediksi jumlah pengangguran tahun 2009 yang dilakukan oleh BPS Provinsi Kalimantan Timur adalah 139.830.

Walaupun memiliki segudang kelebihan, JST juga mempunyai sejumlah keterbatasan, antara lain kekurangan kemampuannya dalam melakukan operasi-operasi numerik dengan presisi tinggi, operasi algoritma aritmatik, operasi logika, dan operasi simbolis serta lamanya proses pelatihan

yang terkadang membutuhkan waktu sehari-hari untuk jumlah data yang besar. Karena sistem JST merupakan sistem yang baru maka hanya dapat berfungsi sebagai alat Bantu sehingga untuk mengambil keputusan masih perlu menggunakan tambahan kebijakan yang lain.

DAFTAR PUSTAKA

- Abdia Away, Gunaidi.2006, "The Shortcut of MATLAB Programming", Informatika Bandung, Bandung.
- Badan Pusat Statistik Provinsi Kalimantan Timur, 2008, "Kalimantan Timur Dalam Angka 2008" , BAPPEDA Provinsi Kalimantan Timur & BPS Provinsi Kalimantan Timur, Kalimantan Timur.
- Budiyanto, Alex. 2003, "Pengantar Algoritma Dan Pemrograman", Kuliah Berseri Ilmu Komputer.
- Desiani Anita, Arhami Muhammad. 2006, "Konsep Kecerdasan Buatan", Andi Yogyakarta, Yogyakarta.
- F. Suhandi, Krisna. 2009, Prediksi Harga Saham Dengan Pendekatan Artificial Neural Network Menggunakan Algoritma Backpropagation, viewed 26 Agustus 2009, <<http://krisnabebrianto.blog.upi.edu/2009/06/27/prediksi-harga-saham-dengan-menggunakan-pendekatan-artificial-neural-network-menggunakan-algoritma-backpropagation.html>>
- Hasan, Iqbal. 2004, "Analisis Data Penelitian Dengan Statistik", Bumi Aksara, Jakarta.
- Kusumadewi, Sri. 2003, "Artificial Intelligence (Teknik dan Aplikasinya)", Graha Ilmu, Yogyakarta.
- Kurniawan, Arief. 2006, "Jaringan Saraf Tiruan Teori Dan Aplikasi" , Andi Yogyakarta, Yogyakarta.
- M.Etter, Delores. 2003, "Pengantar MATLAB 6", PT Indeks kelompok Gramedia, Jakarta.
- Perancangan dan Pembuatan Perangkat Lunak Berbasis Jaringan Syaraf Tiruan Untuk Memprediksi Harga saham, Universitas Kristen Petra, viewed 23 Juli 2009, <<http://digilib.petra.ac.id/jiunkpe/s1/info/2005/jiunkpe-ns-s1-2005-26401114-4881-backpropagation-chapter1.pdf>>
- Puspitaningrum, Diah. 2006, "Pengantar Jaringan Syaraf Tiruan", Andi Yogyakarta, Yogyakarta.
- Siang Jong, Jek. 2005, "Jaringan Syaraf Tiruan & pemogramannya Menggunakan Matlab", Andi Yogyakarta, Yogyakarta.
- W.Panjaitan, Lanny.2007, "Dasar - Dasar Komputasi Cerdas", Andi Yogyakarta, Yogyakarta.

Lampiran 1

Tabel 1. Data Asli BPS Provinsi Kalimantan Timur

No	Input								Target
	X ₀	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈
1	2004	2.750,3	59,75	1.137.730	8,46	1.041.494	89.483.540	1.041.494	96.236
2	2005	2.887,1	62,36	1.213684	11,17	1.078.094	91.050.429	1.078.094	135.590
3	2006	2.955,5	67,27	1.324.878	13,43	1.146.881	93.938.002	1.146.881	177.997
4	2007	3.024,8	61,76	1.241.421	12,07	1.091.625	96.612.842	1.091.625	149.796
5	2008	3.094,7	64,31	1.416.963	11,11	1.259.587	97.803.248	1.259.587	157.376

Keterangan :

Input:

X₀ = Data tahun.

X₁ = Penduduk menurut kelompok umur dan jenis kelamin.

X₂ = Data historis TPAK (Tingkat Partisipasi Angkatan Kerja).

X₃ = Angkatan kerja.

X₄ = Tingkat pengangguran terbuka (TPT).

X₅ = Penduduk yang bekerja menurut lapangan usaha.

X₆ = PDRB (Penduduk Domestik Regional Bruto) atas dasar harga konstan 2000 menurut lapangan usaha.

X₇ = Data historis penduduk usia 15 tahun keatas, status, pendidikan dan jenis pekerjaan.

Target:

X₈ = Data pengangguran reality dilapangan tahun 2004 sampai dengan 2008.

Tabel 2. Data Setelah Normalisasi

Input								Target
X ₀	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈
0	0	0	0	0	0	0	0	0
0,25	0,39721254	0,34707447	0,2720094	0,54527163	0,16781832	0,16781832	0,16781832	0,48132973
0,5	0,59581882	1	0,67022164	1	0,48322046	0,53541086	0,48322046	1
0,75	0,79703833	0,26728723	0,37134221	0,72635815	0,22986066	0,85691733	0,22986066	0,65508005
1	1	0,60638298	1	0,5331992	1	1	1	0,74778929