

**JARINGAN SYARAF TIRUAN DAN ALGORITMA GENETIKA DALAM PEMODELAN KALIBRASI (STUDI KASUS : TANAMAN OBAT TEMULAWAK)**  
*(Artificial Neural Network and Genetic Algorithm on Calibration Spectroscopy (Case Study: Curcuma Medicinal Plant))*

Bartho Sihombing<sup>1</sup>, Erfiani<sup>2</sup>, Utami Dyah Syafitri<sup>3</sup>

<sup>1</sup>Mahasiswa Pasca Sarjana Statistika pada Departemen Statistika FMIPA IPB

<sup>2,3</sup>Departemen Statistika, FMIPA-IPB

E-mail : <sup>1</sup>[bartho.sihombing@yahoo.com](mailto:bartho.sihombing@yahoo.com), <sup>2</sup>[erfiani\\_ipb@yahoo.com](mailto:erfiani_ipb@yahoo.com), <sup>3</sup>[tam\\_ds@yahoo.com](mailto:tam_ds@yahoo.com)

**Abstract**

*The problems in prediction of calibration model are multicollinearity and the number of variables is larger than the number of observations. Principal Component Analysis-Artificial Neural Network-Genetic Algorithm (PCA-ANN-GA) models were applied for the relationship between sample of concentration which is limited and transmittance data which is in large dimensions. A large number of variables were compressed into principal components (PC's). From these PC's, the ANN was employed for prediction of concentration. The principal components computed by PCA were applied as inputs to a backpropagation neural network with one hidden layer. The models was evaluated using GA for the best network structure on hidden layer. Root Mean Square Error (RMSE) for 80% training set and 20% testing set are 0.0314 and 0.5225, respectively. Distribution of data according to the percentage of training data and testing data were also very influential to obtain the best network structure with the smallest RMSE achievement. The best model for these methods is two layers Neural Network with eight neuron in the hidden layer.*

**Keywords :** *Principal Component Analysis, Artificial Neural Network, Genetic Algorithm, RMSE*

**PENDAHULUAN**

Pada beberapa percobaan kimia, penelitian kandungan senyawa aktif suatu tanaman obat memerlukan tahapan yang panjang dan rumit. Selain itu tidak sedikit biaya yang diperlukan untuk pengadaan bahan selama persiapan contoh sampai pada pengukuran menggunakan HPLC (*High Performance Liquid Chromatography*). Pemodelan kalibrasi dikembangkan untuk mengatasi permasalahan waktu dan biaya.

Kandungan senyawa aktif suatu tanaman obat dapat diketahui dengan metode HPLC (*High Performance Liquid Chromatography*) dan FTIR (*Fourier Transform Infrared*). Penggunaan HPLC membutuhkan proses yang lama dan biaya yang mahal untuk memperoleh konsentrasi suatu unsur atau senyawa dari tanaman obat sedangkan FTIR yang relatif lebih sederhana dan murah namun keluaran yang dihasilkannya hanya berupa spektrum yang menunjukkan besarnya nilai serapan saat contoh disinari inframerah (persen transmitan) dengan alat spektrometer FTIR (Sunaryo 2005). Dengan metode HPLC, suatu senyawa dapat diketahui secara kualitatif dan kuantitatif yaitu dengan mengetahui pola kromatogram dan membandingkan luas area

terhadap suatu standar senyawa yang diketahui, sedangkan spektroskopi FTIR memberikan informasi yang mencerminkan gugus fungsi yang terdapat pada suatu senyawa aktif dan kuantitatif melalui nilai absorbansinya. Data keluaran FTIR merupakan data kontinu terhadap bilangan gelombang. Oleh karena itu diperlukan suatu metode alternatif yang lebih murah, mudah dan cepat untuk memperoleh dugaan kandungan senyawa aktif dalam tanaman obat.

Salah satu metode alternatif yang dapat menyatakan hubungan antara konsentrasi senyawa aktif hasil pengukuran HPLC dengan persen transmitan yang diukur dengan menggunakan FTIR adalah model kalibrasi. Model kalibrasi merupakan suatu bagian dari kemometrik yang fokus untuk mencari hubungan antara suatu himpunan pengukuran yang diperoleh melalui proses yang relatif mudah atau murah dan himpunan pengukuran lain yang memerlukan waktu lama dan biaya mahal dalam memperolehnya (Naes *et al.* 2002). Tujuan pemodelan kalibrasi adalah menemukan model yang dapat digunakan untuk memprediksi konsentrasi senyawa secara akurat berdasarkan informasi persen transmitan senyawa yang dianalisis (Erfiani 2005). Dalam penyusunan

model kalibrasi masalah yang sering timbul adalah adanya multikolinieritas antara peubah bebas. Selain itu muncul juga masalah bahwa banyaknya peubah bebas jauh lebih besar dari banyaknya pengamatan.

Terdapat beberapa metode pendekatan untuk menyusun model kalibrasi peubah ganda antara lain metode *Partial Least Square* (PLS) dan Jaringan Syaraf Tiruan (JST). Beberapa peneliti di Institut Pertanian Bogor (IPB) telah mengembangkan model kalibrasi untuk kasus yang berbeda. Arnita (2005) melakukan koreksi pencaran pada senyawa aktif *gingerol* serbuk rimpang jahe, model kalibrasi yang dibentuk dari data yang dikoreksi pencarannya mampu memberikan nilai RMSEP yang lebih kecil dibanding model kalibrasi yang dibentuk dari data yang tidak dikoreksi pencarannya. Atok (2005) menggunakan Jaringan Syaraf Tiruan (JST) dengan metode pra-pemrosesan Analisis Komponen Utama (AKU) untuk mencari pemodelan kalibrasi untuk data HPLC dan FTIR dari zat aktif serbuk jahe, kajian dilakukan hanya didasarkan pada bangkitan data yang konvergen sehingga hasil yang diperoleh dapat diterapkan jika algoritma pembelajaran pada JST menjamin kekonvergenan tersebut. Erfiani (2005) mengembangkan model kalibrasi dengan pendekatan Bayes dengan menggunakan pendekatan regresi terpenggal untuk mereduksi dimensi data, sedangkan Sunaryo (2005) menggunakan Regresi Komponen Utama dengan metode Wavelet digunakan untuk pra-pemrosesan, hasil yang diperoleh memperbaiki nilai RMSEP yang diperoleh oleh Arnita (2005). Mukid (2009) menerapkan regresi proses Gaussian dengan kajian terhadap penggunaan berbagai fungsi peragam, Santi (2010) menggunakan metoda GA-PLS untuk pendugaan model kalibrasi. Peneliti lain yang telah mempublikasikan metode kalibrasi adalah Habibi *et.al.* yang dalam penelitiannya membandingkan aplikasi AKU-Algoritma Genetika (AG)-JST dengan AKU-AG-Multiple Linear Regression dengan kesimpulan bahwa pemodelan dengan menggunakan AKU-AG-JST relatif lebih baik.

Dengan memperhatikan berbagai penelitian pemodelan kalibrasi yang telah dilakukan, penulis kemudian akan menggunakan aplikasi AKU-JST-AG untuk menduga model kalibrasi pada pengukuran konsentrasi kurkumin pada tanaman temulawak berdasarkan data persen transmisinya. Prapemrosesan akan dilakukan dengan AKU, kemudian dimodelkan dengan JST. Arsitektur dari JST akan dievaluasi oleh AG dengan harapan diperoleh arsitektur yang baik dengan RMSE minimum. Penelitian ini bertujuan untuk membangun model kalibrasi dengan menggunakan JST yang dioptimasi menggunakan Algoritma Genetika pada pengukuran konsentrasi kurkumin temulawak berdasarkan persen transmisinya.

## SENYAWA AKTIF PADA TEMULAWAK

Menurut Sinambela (1985), komposisi rimpang temulawak dapat dibagi menjadi dua fraksi utama yaitu zat warna kurkuminoid dan minyak atsiri. Warna kekuningan temulawak disebabkan adanya kurkuminoid. Kandungan utama kurkuminoid terdiri dari senyawa kurkumin, desmetoksi-kurkumin dan bisdesmetoksikurkumin. Rimpang temulawak segar, selain terdiri dari senyawa kurkuminoid dan minyak atsiri juga mengandung lemak, protein, selulosa, pati, dan mineral. Kadar masing-masing zat tersebut tergantung pada umur rimpang yang dipanen serta juga dipengaruhi oleh letak dan ketinggian tempat temulawak berada.

### Spektroskopi FTIR (*Fourier Transform Infrared*)

FTIR merupakan salah satu teknik spektroskopi inframerah. Instrumentasi spektrum inframerah dibagi kedalam tiga jenis radiasi yaitu inframerah dekat (bilangan gelombang 12800-4000  $\text{cm}^{-1}$ ), inframerah pertengahan (bilangan gelombang 4000-200  $\text{cm}^{-1}$ ), dan inframerah jauh (bilangan gelombang 200-10  $\text{cm}^{-1}$ ) (Nur & Adijuwana 1989). FTIR termasuk dalam kategori radiasi inframerah pertengahan.

Spektrum inframerah senyawa tumbuhan dapat diukur dengan spektrofotometri inframerah yang merekam secara otomatis dalam bentuk larutan (dalam kloroform, karbontetraklorida, 1-5%), bentuk gerusan dalam minyak nujol, atau bentuk padat yang dicampur dengan kalium bromida. Daerah pada spektrum inframerah di atas 1200  $\text{cm}^{-1}$  menunjukkan pita spektrum atau puncak yang disebabkan oleh getaran ikatan kimia atau gugus fungsi dalam molekul yang ditelaah (Harbone 1996).

### High Performance Liquid Chromatography (HPLC)

Kromatografi adalah suatu metode pemisahan komponen-komponen suatu campuran, komponen-komponen tersebut akan terdistribusi diantara dua fase. Salah satu fase dibuat diam dan dinamakan fase diam atau fase stasioner, fase lainnya disebut fase gerak atau fase mobil yang bergerak diantara celah-celah atau pada permukaan fase stasioner. Pergerakan fase mobil ini mengakibatkan pergerakan diferensial dari komponen-komponen contoh (Nur dan Adijuwana 1989). Fase diam pada kromatografi dapat berupa cair atau padatan sedangkan fase gerak dapat berupa cair atau gas. Berdasarkan jenis fasenya kromatografi dapat digolongkan menjadi empat jenis yaitu: cair-padatan, gas-padatan, cair-cair, dan gas-cair.

### Analisis Komponen Utama (AKU)

Analisis Komponen Utama (AKU) adalah teknik yang digunakan untuk menyederhanakan

suatu data, dengan cara mentransformasi data secara linier sehingga terbentuk sistem koordinat baru dengan keragaman maksimum. AKU dapat digunakan untuk mereduksi dimensi suatu data tanpa mengurangi karakteristik data tersebut secara signifikan. AKU juga sering digunakan untuk menghindari masalah multikolinearitas antar peubah bebas dalam model regresi berganda. Dalam AKU peubah-peubah yang masih saling berkorelasi ditransformasi menjadi satu set peubah baru yang tidak berkorelasi lagi, peubah-peubah baru itu disebut sebagai Komponen Utama (KU) (Johnson & Wichren 1982).

### Algoritma Genetika (AG)

AG merupakan metode adaptif yang biasa digunakan untuk memecahkan suatu pencarian nilai dalam sebuah masalah optimasi (Suyanto 2005). Algoritma ini didasarkan pada proses genetik yang ada dalam makhluk hidup; yaitu perkembangan generasi dalam sebuah populasi yang alami, secara lambat laun mengikuti proses seleksi alam atau “siapa yang kuat, dia yang bertahan (*survive*)”. Dengan meniru teori evolusi ini, AG dapat digunakan untuk mencari permasalahan-permasalahan dalam dunia nyata. Algoritma ini bekerja dengan sebuah populasi yang terdiri dari individu-individu, masing-masing individu mempresentasikan sebuah solusi yang mungkin bagi persoalan yang ada. Dalam kaitan ini individu dilambangkan dengan sebuah nilai *fitness* yang akan digunakan untuk mencari solusi terbaik dari persoalan yang ada. AG pertama kali dikembangkan oleh John Holland pada tahun 1970 di New York.

### Jaringan Syaraf Tiruan (JST)

JST adalah sistem komputasi dimana arsitektur dan operasi diilhami dari pengetahuan tentang sel syaraf biologis dalam otak. Istilah JST digunakan karena jaringan syaraf ini diimplementasikan dengan menggunakan program komputer yang mampu menyelesaikan sejumlah proses perhitungan selama proses pembelajaran, cara kerja jaringan syaraf tiruan meniru cara kerja otak manusia (Siang 2009). Salah satu contoh pengambilan ide dari jaringan syaraf biologis adalah adanya elemen-elemen pemrosesan pada JST yang saling terhubung dan beroperasi secara paralel. Ini meniru jaringan syaraf biologis yang tersusun dari sel-sel syaraf (*neuron*).

## METODOLOGI PENELITIAN

### Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang merupakan bagian dari data penelitian Hibah Pascasarjana tahun 2003-2005 hasil kerjasama antara Departemen Statistika IPB dengan Pusat Studi Biofarmaka LPPM IPB.

Penelitian tersebut didanai oleh Dirjen Pendidikan Tinggi, Departemen Pendidikan Nasional. Data yang digunakan adalah persen transmittan kurkumin dari serbuk temulawak hasil pengukuran spektrometer FTIR dan data konsentrasi senyawa aktif kurkumin yang diukur dengan menggunakan HPLC. Temulawak yang dijadikan contoh diambil dari beberapa daerah sentra tanaman obat, yaitu Bogor, Sukabumi, Kulon Progo, Karanganyar, Cianjur dan Balitro. Data-data tersebut diperoleh dari Pusat Studi Biofarmaka Institut pertanian Bogor.

### Metode Analisis

Penggunaan algoritma genetika dalam optimasi jaringan syaraf tiruan dilakukan untuk mendapatkan struktur *neuron* pada lapis tersembunyi yang mendekati optimal. Tingkat pengenalan JST dalam pendugaan yang tinggi akan didapat apabila seluruh *neuron* pada lapis tersembunyi memberikan nilai kontribusi objektif yang tinggi, dalam hal ini penulis akan menekankan nilai  $R^2$ . Apabila *neuron* yang memberikan kontribusi  $R^2$  yang kecil dapat dihilangkan, sedangkan yang memberikan kontribusi  $R^2$  besar dapat dipertahankan, maka jaringan syaraf tiruan ini dapat diharapkan memberikan nilai  $R^2$  yang lebih tinggi.

Penghilangan *neuron* yang kurang bermanfaat ini dapat dilakukan dengan dua cara, yaitu dengan membuang sejumlah bobot dari setiap *neuron* yang memberi kontribusi  $R^2$  kecil atau dengan membuang sejumlah *neuron* yang berarti membuang seluruh bobot keterhubungan dari *neuron* yang kurang bermanfaat (Apriyanti 2005). Pada penelitian ini digunakan pendekatan kedua yaitu membuang bobot keterhubungan dari *neuron* yang memberikan kontribusi  $R^2$  kecil.

Pada proses AG pengkodean yang akan dipakai adalah *string* biner, dengan tiap bit dalam *string* kromosom merepresentasikan sebuah *neuron*. Bit yang bernilai 1 merepresentasikan *neuron* yang dipertahankan dan bit yang bernilai 0 merepresentasikan *neuron* yang dibuang. Penggunaan parameter seperti dijelaskan diatas diharapkan mencukupi bagi AG untuk melakukan pencarian solusi optimal bagi jumlah *neuron* lapis tersembunyi.

Pada penelitian ini langkah pertama yang akan dikerjakan adalah melakukan prapemrosesan dengan PCA untuk mereduksi dimensi dari peubah bebas yang dalam hal ini adalah persen transmittan yang dihasilkan oleh FTIR. Pada prapemrosesan AKU pengambilan Komponen Utama dilakukan untuk berbagai keragaman kumulatif komponen tersebut menerangkan keragaman data asli. Data dari hasil prapemrosesan dibagi 2 bagian : bagian pertama terdiri dari beberapa pengamatan untuk pemodelan dalam tahap pelatihan (*training*) dan bagian kedua untuk *testing*. Pembagian data

pengamatan ke dalam kelompok data *training* dan data *testing* dicobakan dalam berbagai komposisi yaitu 60%, 70% dan 80% pada data *training*.

Selanjutnya, JST dengan algoritma *backpropagation* digunakan untuk memproses hubungan antara peubah-peubah baru hasil PCA dengan respon. Dilakukan pendugaan terhadap nilai target (HPLC). Pendugaan untuk memperoleh nilai dugaan mendekati nilai target dilakukan dengan cara menyesuaikan bobot pada masing-masing *neuron*. Spesifikasi JST yang digunakan :

- Arsitektur yang digunakan adalah *Feed Forward Neural Network* (FFNN) banyak lapisan, dalam hal ini *neuron-neuron* disusun dalam lapisan-lapisan dan sinyal-sinyal mengalir dari input ke lapisan pertama, kemudian ke layer kedua, dan seterusnya.
- Masukan merupakan hasil prapemrosesan AKU. Pada JST dua lapis, pengkodean *string biner* sepanjang 16 bit diambil dari jumlah *neuron* maksimum lapisan tersembunyi. Hal ini berdasarkan penelitian Kusumoputro (2004). Lapisan keluaran menggunakan satu *neuron* sesuai dengan target, yaitu setiap observasi terhubung dengan satu nilai target (HPLC).
- Algoritma pembelajaran yang digunakan adalah *Backpropagation* dengan fungsi transfer/aktivasi untuk lapis tersembunyi adalah *sigmoid biner (logsig)* dan fungsi transfer linear untuk lapis keluaran.
- Teknik inisialisasi yang digunakan adalah inisialisasi Nguyen-Widrow. Algoritma ini memberikan bobot awal pada neuron dengan nilai antara -0,5 sampai 0,5, sedangkan bobot-bobot dari masukan ke lapis tersembunyi dirancang sedemikian rupa sehingga mempercepat proses pembelajaran (Egan *et al.* 2001). Nguyen-Widrow didefinisikan sebagai persamaan berikut:

- Hitung harga faktor pengali  $\beta = 0.7 p^{1/n}$  dengan p banyaknya jumlah *neuron* lapisan tersembunyi dan n banyaknya neuron pada lapisan *input*.
- Untuk setiap unit tersembunyi ( $j=1, 2, \dots, p$ ), dihitung  $v_{ij}$  (lama) yaitu bilangan acak diantara -0.5 dan 0.5 (atau di antara  $-\gamma$  dan  $+\gamma$ ). Pembaharuan bobot  $v_{ij}$  (lama) menjadi  $v_{ij}$  baru yaitu:  $v_{ij}(\text{baru}) = \frac{\beta v_{ij}(\text{lama})}{\|v_{ij}(\text{lama})\|}$
- Tetapkan Bias,  $v_{oj} = \text{Bobot pada bias bernilai antara } -\beta \text{ dan } \beta$ .

Respon yang diambil dalam penelitian ini adalah nilai  $R^2$  yang dicapai oleh JST setelah dilakukan optimasi oleh algoritma genetika berdasarkan presentase data training, persentase keragaman AKU yang digunakan dan nilai pembelajaran. Sedang fungsi *fitness* yang dipakai adalah memaksimalkan nilai  $R^2$ . Dalam

penggunaan data *training* dan data *testing* dievaluasi dengan mencari *Root Mean Square Error* (RMSE) yang didefinisikan dengan :

$$RMSE = \sqrt{\frac{1}{N_{cal}} \sum_{n=1}^{N_{cal}} (y_{cal,n} - \hat{y}_{cal,n})^2}$$

dengan

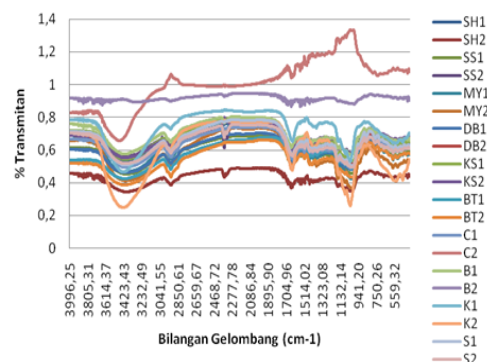
$y_{cal,n}$  adalah nilai pengamatan berdasarkan n contoh kalibrasi,  $\hat{y}_{cal,n}$  adalah nilai dugaan pengamatan dengan menghilangkan sebanyak n contoh dari sekumpulan N contoh kalibrasi, dan  $n_p$  adalah banyak contoh yang digunakan dalam model validasi.

## HASIL DAN PEMBAHASAN

### Deskripsi Spektrum Kurkumin

Spektrum kurkumin diambil dari 20 contoh serbuk temulawak yang berasal dari berbagai daerah sentra tanaman obat dapat dilihat pada Gambar 1. Spektrum kurkumin dari berbagai daerah tersebut sebagian besar memiliki pola yang hampir sama kecuali untuk beberapa spektrum yang menunjukkan pola yang agak berbeda. Terlihat bahwa spektrum kurkumin dari contoh serbuk temulawak yang diambil dari daerah Cianjur (contoh C2) dan Bogor (contoh B2) agak berbeda.

Data persen transmittan diperoleh dari pengukuran dengan menggunakan FTIR pada 1866 bilangan gelombang yang berkisar antara 4000 – 400  $\text{cm}^{-1}$ . Pada indeks bilangan gelombang disekitar 1500  $\text{cm}^{-1}$  ketika spektrum kurkumin serbuk temulawak dari sebagian besar contoh memiliki pola grafik yang cekung ke atas, tetapi temulawak yang diambil dari daerah Cianjur menunjukkan pola grafik yang cekung kebawah. Tampak juga bahwa spektrum kurkumin serbuk temulawak yang diambil dari daerah Bogor (contoh B2) menunjukkan pola yang cenderung konstan di setiap bilangan gelombang.



Gambar 1 Spektra kurkumin serbuk temulawak

**Reduksi Peubah Penjelas**

Data persen transmitan diukur pada 1866 bilangan gelombang yang dalam pemodelan kalibrasi ini berperan sebagai peubah penjelas. Ada beberapa alasan utama mengapa reduksi jumlah peubah penjelas ini dilakukan. Pertama, besar kemungkinan antara peubah penjelas satu dengan lainnya tidak saling bebas. Kedua, bekerja dengan sedikit peubah penjelas akan menyederhanakan proses komputasi.

Analisis Komponen Utama (AKU) digunakan untuk mereduksi banyaknya peubah penjelas dengan persentase keragaman kumulatif yang mampu dijelaskan digunakan sebagai kriteria untuk menentukan banyaknya komponen utama. Tabel 1 menjelaskan bahwa pada bilangan gelombang 4000–400 cm<sup>-1</sup> dengan menggunakan 1 komponen utama, keragaman yang dapat dijelaskan sebesar 89,7592% dan apabila menggunakan 2 komponen utama keragaman yang dapat dijelaskan sebesar 95,1612% sedangkan apabila menggunakan 3 komponen utama keragaman yang dapat dijelaskan sebesar 99,2061% dari keragaman pada data asal.

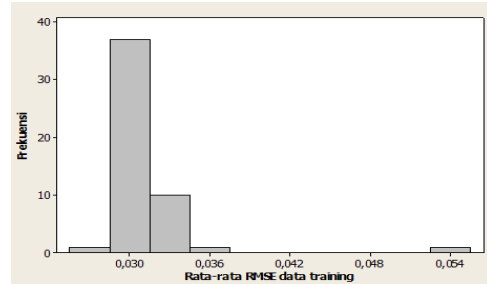
Tabel 1 Ragam kumulatif komponen utama

Komponen Utama	Ragam yang dijelaskan (%)	Ragam Kumulatif (%)
1	89,7592	89,7592
2	5,4019	95,1612
3	4,0449	99,2061
4	0,5447	99,7508
5	0,1047	99,8555
6	0,0624	99,9179
7	0,0309	99,9488
8	0,0304	99,9792
9	0,0099	99,9891
10	0,0033	99,9924
11	0,0030	99,9954
12	0,0019	99,9973
13	0,0011	99,9984
14	0,0007	99,9991
15	0,0004	99,9995
16	0,0002	99,9996
17	0,0002	99,9998
18	0,0001	99,9999

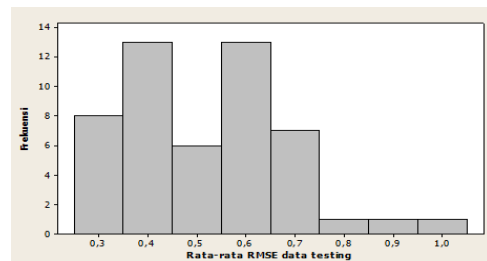
**Nilai Dugaan Terhadap Nilai HPLC**

Pemodelan menggunakan JST dengan optimasi AG dilakukan dalam berbagai kelompok data. Pada percobaan JST dengan optimasi AG, kelompok data dengan jumlah data *training* sebanyak 80% dan komponen utama yang digunakan sebagai *input* pada jaringan menjelaskan 99,8555% data

asal, diperoleh model terbaik dengan nilai rata-rata  $RMSE_{training}$  0,0314 dan nilai  $RMSE_{testing}$  0,5225, dengan nilai rata-rata  $R^2$  yang diperoleh 49,93%. Gambar 2 dan Gambar 3 merupakan histogram frekuensi rata-rata RMSE yang diperoleh pada kelompok 80% data *training* dan data *testing*.



Gambar 2 Histogram frekuensi rata-rata nilai RMSE pada data *training*



Gambar 3 Histogram frekuensi rata-rata nilai RMSE pada data *testing*

Percobaan dilakukan dengan mengamati pembagian kelompok data menurut data *training* dan data *testing* lainnya. Pada kelompok data *training* 70%, diperoleh nilai  $RMSE_{training}$  0,0297 dan  $RMSEP$  0,5556, nilai  $R^2$  yang diperoleh adalah 0,3964 (39,644%). Demikian juga dengan menggunakan kelompok data pada data *training* 60% diperoleh rata-rata nilai  $RMSE_{training}$  sebesar 0,0326 dan  $RMSEP$  0,5844, dengan nilai  $R^2$  yang diperoleh sebesar 0,2789 (27,89%). Pembagian kelompok data menjadi data *training* relatif berpengaruh terhadap pencapaian  $R^2$ , sesuai dengan penelitian Apriyanti (2005) yang menyimpulkan bahwa pembagian data menjadi data *training* dan *testing* cukup berpengaruh terhadap peningkatan nilai  $R^2$ . Peneliti sebelumnya yang dilakukan oleh Mukid (2009) menghasilkan nilai  $RMSEP$  terbaik untuk berbagai jenis fungsi peragam dengan menggunakan seluruh gugus data adalah 0,5913. Tabel 2 merupakan tabel perbandingan nilai rata-rata RMSE yang diperoleh sesuai dengan pengelompokan data.

Algoritma genetika sangat baik dilakukan untuk memperoleh struktur jaringan yang baik dengan nilai  $RMSE_{testing}$  minimum. Jika dibandingkan dengan jaringan tanpa menggunakan AG, hasil yang diperoleh JST optimasi AG relatif lebih baik. Pada kelompok data *training* 80%, dilakukan

percobaan AKU-JST tanpa menggunakan AG, diperoleh rata-rata nilai  $RMSE_{training}$  0,0282 dan  $RMSE_{testing}$  0,6117 dengan rata-rata  $R^2$  sebesar 0,4064. Hasil ini tidak lebih baik dibanding dengan program AKU-JST dengan menggunakan AG. Tabel 3 merupakan tabel perbandingan nilai rata-rata RMSE yang diperoleh sesuai pengelompokan data tanpa menggunakan AG.

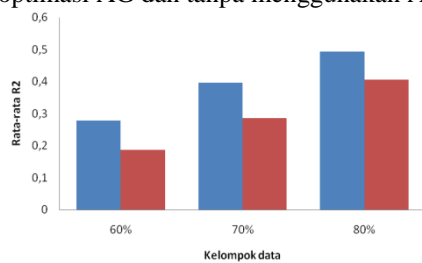
Tabel 2 Perbandingan rata-rata RMSE AKU-JST-AG untuk data *training* dan data *testing*

Data Training	$RMSE_{training}$	$RMSE_{testing}$	$R^2$
60%	0,0326	0,5844	0,2789
70%	0,0297	0,5556	0,3964
80%	0,0304	0,5225	0,4993

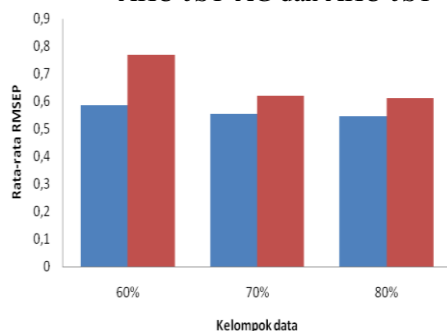
Tabel 3 Perbandingan rata-rata RMSE GA-JST tanpa AG untuk data *training* dan data *testing*

Data Training	$RMSE_{training}$	$RMSE_{testing}$	$R^2$
60%	0,0304	0,7698	0,1867
70%	0,0282	0,6198	0,2860
80%	0,0282	0,6117	0,4065

Gambar 4 dan Gambar 5 merupakan plot perbandingan nilai  $R^2$  dan RMSEP sesuai pengelompokan data untuk jaringan menggunakan optimasi AG dan tanpa menggunakan AG.



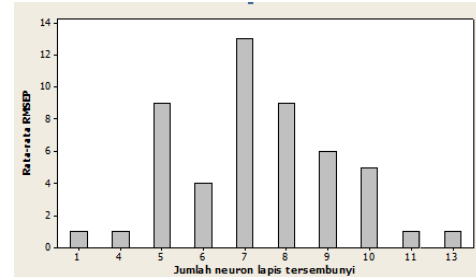
Gambar 4 Perbandingan nilai rata-rata  $R^2$  metode AKU-JST-AG dan AKU-JST



Gambar 5 Perbandingan nilai  $RMSE_{testing}$  metode AKU-JST-AG dan AKU-JST

Sebaran jumlah *neuron* lapis tersembunyi metode AKU-JST-AG yang menghasilkan rata-rata  $RMSE_{testing}$  minimum untuk melihat kecenderungan jumlah *neuron* yang memberikan hasil dugaan

paling mendekati nilai target diperlihatkan pada Gambar 6. Secara umum metode AKU-JST-AG lapis banyak rata-rata menghasilkan delapan *neuron* pada lapis tersembunyi dalam arsitektur JST yang menghasilkan rata-rata  $RMSE_{testing}$  paling minimum.



Gambar 6 Sebaran jumlah neuron lapisan tersembunyi dan rata-rata RMSEP

Pemilihan n-komponen utama pertama yang dijadikan *input* pada jaringan memberikan hasil yang berbeda. Tabel berikut menyajikan sebelas model yang disusun dengan mencoba setiap komponen utama pada sebagai *input* pada model.

Tabel 4 Nilai RMSEP dan  $R^2$  berdasarkan jumlah KU

Model	Peubah	$RMSE_{testing}$	$R^2$
1	2	0,7392	0,3342
2	3	0,5440	0,4672
3	4	0,5162	0,4651
4	5	0,5225	0,4994
5	6	0,6246	0,4226
6	7	0,6003	0,3687
7	8	0,5944	0,4873
8	9	0,5538	0,3979
9	10	0,5515	0,4223
10	11	0,5495	0,4519
11	12	0,5260	0,4276
12	13	0,6046	0,4367
13	14	0,5459	0,4941
14	15	0,5754	0,4149
15	16	0,6209	0,3878
16	17	0,5715	0,4216
17	18	0,7560	0,3684

Tabel 4 memperlihatkan bahwa model ketiga dan keempat adalah model yang baik. Model keempat memberikan nilai  $R^2$  yang lebih tinggi dengan  $RMSE_{testing}$  0,5225 yang memberikan sumber keragaman terbesar dengan kumulatif keragaman sebesar 99,8555%.

## SIMPULAN DAN SARAN

### Simpulan

Algoritma Genetika dapat digunakan untuk mengoptimalkan kerja JST, dalam hal ini dilakukan pada struktur *neuron* pada lapis tersembunyi untuk



memperoleh model terbaik yang diukur dengan nilai RMSE yang diperoleh. Kajian dilakukan terhadap pembagian kelompok data yang dibagi menjadi data *training* dan data *testing* memberikan hasil yang relatif berbeda.

#### Saran

Penelitian ini menggunakan metode AKU-JST-AG dengan harapan untuk mendapatkan pemodelan kalibrasi terbaik. Banyak metode lain yang perlu untuk dikembangkan lagi sehingga diperoleh pemodelan kalibrasi terbaik. Pemilihan prapemrosesan lain untuk mereduksi dimensi data yang besar sangat diperlukan untuk mendapatkan pendugaan yang terbaik. AG perlu dipertimbangkan dalam pemilihan komponen utama yang paling berpengaruh sebelum dilakukan pendugaan oleh JST.

#### DAFTAR PUSTAKA

- Apriyanti N. 2005. Optimasi Jaringan Syaraf Tiruan Dengan Algoritma Genetika Untuk Peramalan Curah Hujan [Skripsi]. Bogor : Program Sarjana Institut Pertanian Bogor.
- Arnita. 2005. Koreksi Pencaran dalam Model Kalibrasi Peubah Ganda pada Data Senyawa Aktif *Gingerol* Serbuk Rimpang Jahe (*Zingiber Officinale Roscoe*) [Tesis]. Bogor: Program Pascasarjana, Institut Pertanian Bogor.
- Atok. 2005. Jaringan Syaraf Tiruan Untuk Pemodelan Kalibrasi Dengan Prapemrosesan Analisis Komponen Utama dan Transformasi Fourier Diskret [Tesis]. Bogor: Program Pascasarjana, Institut Pertanian Bogor.
- Erfiani. 2005. Pengembangan Model Kalibrasi dengan Pendekatan bayes (Kasus Tanaman Obat [Disertasi]. Bogor: Program Pascasarjana, Institut Pertanian Bogor.
- Habibi-Yangjeh A, Pourbasheer E, Danandeh-Jenagharad M. 2009. Application of Principal Component-Genetic Algorithm-Atificial Neural Network for Prediction Acidity Constant of Various Nitrogen-Containing Compounds in Water. *Monatsh Chem*.
- Harbone JB. 1996. Metode Fitokimia "Penuntun Cara Modern Menganalisis Tumbuhan". Terbitan Kedua. Penerbit ITB. Bandung.
- Kusumoputro B. 2004. Pengembangan Sistem Pengenalan Wajah Secara 3 Dimensi Menggunakan Hemisphere Structure of Neural Networks dan Optimasi Struktur Menggunakan Algoritma Genetika. Makalah Seminar Nasional dan Ilmu Komputer dan Teknologi Informasi V. Fakultas Ilmu Komputer UI. Depok
- Naes T, Issackson T, Fearn T, Davies T. 2002. User Friendly Guide to Multivariate Calibration and Classification. United Kingdom: NIR Publication Chichester.
- Nur MA, Adijuwana H. 1989. Teknik Spektroskopi dalam Analisis Biologi. Bogor: Institut Pertanian Bogor.
- Santi VM. 2010. Model Kalibrasi Spektroskopi Dengan Preprocessing Genetich Algorithm (GA) (Studi Kasus : Tanaman Obat Temulawak) [Tesis]. Bogor: Program Pascasarjana, Institut Pertanian Bogor.
- Siang JJ. 2009. Jaringan Syaraf Tiruan dan Pemrogramannya menggunakan Matlab. Penerbit Andi. Yogyakarta.
- Sinambela JM. 1985. Fitoterapi, Fitostandar dan Temulawak. Prosiding Simposium Nasional temulawak. Bandung, 17 September 1985. Lembaga Penelitian Universitas Padjajaran.
- Socrates G. 1994. Infrared Characteristic Group Frequencies Tables and Charts. New York: John Wiley and Sons.
- Sunaryo S. 2005. Model Kalibrasi dengan Transformasi *Wavelet* Sebagai Metode Prapemrosesan [Disertasi]. Bogor: Program Pascasarjana, Institut Pertanian Bogor.
- Suyanto. 2005. Algoritma Genetika dalam Matlab. Penerbit Andi. Yogyakarta.