

PENINGKATAN NEURAL NETWORK DENGAN FEATURE SELECTION UNTUK PREDIKSI KANKER PAYUDARA

Evy Priyanti

Program Studi Komputerisasi Akuntansi
AMIK BSI Jakarta
Jl. Margonda Raya No.8, Depok
evy.evp@bsi.ac.id

ABSTRACT

Breast cancer is increasing in every countries in the world, especially in developing countries like Indonesia. Neural network is able to solve problems with the accuracy of data and not linear. Neural network optimization tested weeks to produce the best accuracy value, applying neural network with feature selection methods such as Wrapper with Backward Elimination to raise the accuracy produced by Neural Network. Experiments conducted to obtain optimal architecture and to increase the value of accuracy. Results of the research is a confusion matrix to prove the accuracy of Neural network before optimized by Backward Elimination was 96.42% and 96.71% after becoming optimized. This proves the estimation of feature selection trials using neural network-based method Backward Elimination more accurate than the individual neural network method.

Kata Kunci: Kanker Payudara, Data Mining, Neural Network

1. PENDAHULUAN

Di seluruh dunia, kanker payudara terdiri 22,9% dari semua kanker di dunia. kanker payudara meningkat di setiap negara di dunia terutama di negara-negara berkembang seperti Indonesia. Pada tahun 1940, peningkatan risiko kanker payudara pada wanita adalah 5%, atau satu dari 20 orang. *The American Cancer Society* memperkirakan bahwa risiko menjadi 13% pada tahun 2012, atau satu dari 8 orang akan terinfeksi. Penelitian telah menunjukkan bahwa faktor-faktor tertentu, yang disebut risiko faktor, meningkatkan kemungkinan bahwa seorang wanita akan lebih mudah terkena kanker payudara. Banyak faktor risiko yang paling penting untuk kanker payudara berada di luar kendali seperti usia, riwayat keluarga, dan riwayat kesehatan. Namun, ada beberapa faktor risiko yang dapat dicegah, seperti berat badan, aktivitas fisik, dan mengkonsumsi alkohol [13].

Neural network (NN) adalah teknik peramalan yang paling umum digunakan, karena NN bisa cepat dan akurat, banyak peneliti menggunakan NN untuk memecahkan masalah peramalan [15]. *Neural network* mempunyai kelebihan yaitu jaringan syaraf mampu menyelesaikan problem nonlinear, mempunyai toleransi yang cukup tinggi terhadap data yang mengandung *noise* dan mampu menangkap hubungan yang sangat kompleks antara variabel-variabel *prediktor* dan *output* nya, namun *Neural network* juga mempunyai kekurangan yaitu adanya *over generalisasi*, dimana penggunaan jumlah data yang banyak, *Neural network*

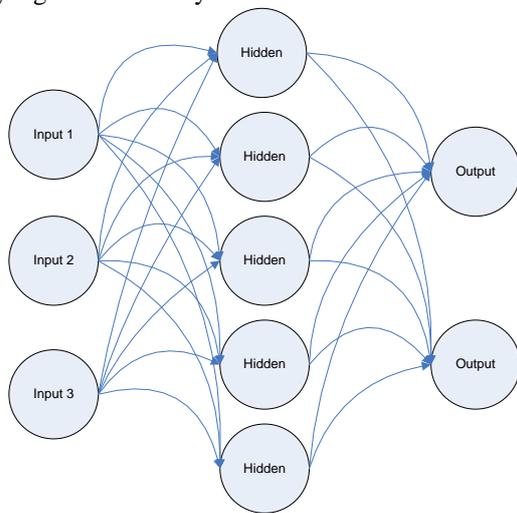
melakukan dengan sangat baik untuk data pelatihan, tetapi gagal untuk melakukan dengan baik untuk data validasi. Masalah ini mungkin disebabkan oleh jumlah *neuron* yang berlebihan, durasi pelatihan yang berlebihan, atau alasan lain yang serupa. Generalisasi yang berlebihan terjadi ketika *Neural Network*, dengan algoritma pelatihan, mulai meniru fungsi yang sangat kompleks di tempat yang sederhana [21].

Feature Selection adalah masalah yang berkaitan erat dengan pengurangan dimensi. Tujuan *Feature Selection* adalah untuk mengidentifikasi fitur dalam kumpulan data yang sama pentingnya, dan membuang semua fitur lain seperti informasi yang tidak relevan dan berlebihan. Karena *feature selection* mengurangi dimensi dari data, sehingga memungkinkan operasi yang lebih efektif & algoritma data mining yang cepat (yaitu data algoritma *mining* dapat dioperasikan lebih cepat dan lebih efektif dengan menggunakan *feature selection*) [19]. Penelitian menggunakan *feature selection* yang pernah dilakukan oleh Yuanning Liu, Gang Wang, Huiling Chen, Hao Dong, Xiaodong Zhu & Sujing Wang tahun 2011, dimana hasil penelitian menunjukkan bahwa penggunaan metode *Improved feature selection* (IFS) menunjukkan hasil yang signifikan lebih baik daripada tiga metode lain yang digunakan dalam penelitian tersebut (GA, SVM, PSO) dalam hal akurasi prediksi dengan *subset feature* yang lebih kecil.

2. LANDASAN TEORI

Data mining atau disebut juga *Knowledge Discovery in Database* (KDD) adalah ekstraksi pola secara otomatis mewakili pengetahuan yang disimpan atau ditangkap secara tersembunyi di dalam sebuah *database* besar [7].

Berdasarkan tugasnya, data mining dikelompokkan menjadi [14]:a). Klasifikasi, b). Estimasi, c). Prediksi, d). *Clustering*, e). Asosiasi.Klasifikasi (Han & Kamber, 2007:p24) adalah proses penemuan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui.



Sumber :(Maimon & Rokach, 2010:p423).

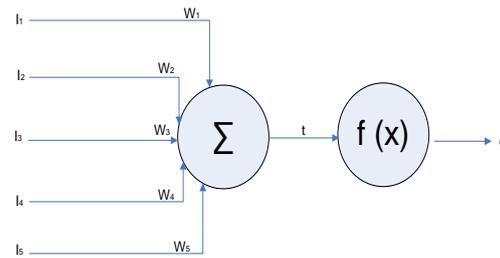
Gambar 2.1. Struktur Neural Network

Neural network terdiri dari *input layer*, satu atau lebih *hidden layer*, dan *output layer*. Berikut penjelasan masing-masing *layer* : [23]

Input layer digunakan untuk menerima nilai masukan dari tiap *record* pada data. Jumlah simpul input sama dengan jumlah variabel prediktor.

Hidden layer mentransformasikan nilai *input* di dalam *network*. Tiap simpul pada *hidden layer* terhubung dengan simpul-simpul pada *hidden layer* sebelumnya atau dari simpul-simpul pada *input layer* dan ke simpul-simpul pada *hidden layer* berikutnya atau ke simpul-simpul pada *output layer*. Jumlah *hidden layer* bisa berapa saja.

Garis yang terhubung dengan *Output Layer* berasal dari *Hidden Layer* atau *Input Layer* dan mengembalikan nilai keluaran yang bersesuaian dengan variabel prediksi.



Sumber : (Shukla,Tiwari & Kala, 2010: p45)

Gambar 2.2. Proses Neural Network

Korelasi antara ketiga komponen pada persamaan di atas yaitu:

Signal x berupa vektor berdimensi n (x_1, x_2, \dots, x_n) T akan mengalami penguatan oleh *synapse* w (w_1, w_2, \dots, w_n) T . Selanjutnya akumulasi dari penguatan tersebut akan mengalami transformasi oleh fungsi aktivasi f . Fungsi f ini akan memonitor, bila akumulasi penguatan signal itu telah melebihi batas tertentu, maka sel *neuron* yang semula berada dalam kondisi “0”, akan mengeluarkan signal “1”. Berdasarkan nilai *output* tersebut (y), sebuah *neuron* dapat berada dalam dua status: “0” atau “1”. *Neuron* disebut dalam kondisi *firing* bila menghasilkan *output* bernilai “1”.

Neural network adalah jaringan saraf yang mensimulasikan jaringan saraf biologis manusia kedalam arsitektur komputer dan arsitektur algoritma baru terhadap komputer konvensional [8]. Hal ini memungkinkan penggunaan operasi komputasi (penambahan, pengurangan, dan elemen logika *fundamental*) yang sangat sederhana untuk memecahkan masalah yang kompleks, matematis yang tidak jelas, masalah nonlinear atau masalah *stokastik*.

Neural network adalah satu set unit input/output yang terhubung dimana tiap relasinya memiliki bobot [7]. *Neural network* dimaksudkan untuk mensimulasikan perilaku sistem biologi susunan syaraf manusia, yang terdiri dari sejumlah besar unit pemroses yang disebut *neuron*, yang beroperasi secara paralel. *Neuron* mempunyai relasi dengan *synapse* yang mengelilingi *neuron-neuron* lainnya. Susunan syaraf tersebut dipresentasikan dalam *neural network* berupa grafik yang terdiri dari simpul (*neuron*) yang dihubungkan dengan busur, yang berkorespondensi dengan *synapse* [1]. Sejak tahun 1950-an [23] *Neural network* telah digunakan untuk tujuan prediksi, bukan hanya klasifikasi tapi juga untuk regresi dengan atribut target kontinu.

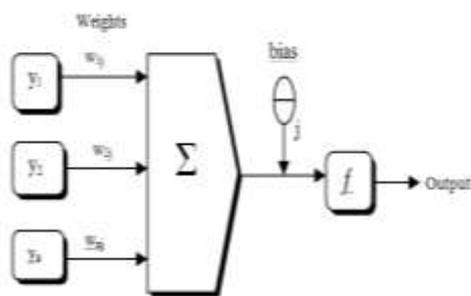
Menurut Kohonen, *Neural network* dibuat pertama kali oleh Warren McCulloh dan Walter Pitts pada tahun 1943 dan

dianggap sebagai basis *neural network modern* saat ini [5]. Pada tahun 1950, Rosenblatt dan peneliti lainnya membuat sebuah kelas *Neural network* yang disebut perceptron yang merupakan model *neuron* biologi tapi hanya bisa menyelesaikan masalah *linearly separable*. Kemudian perceptron yang dihasilkan Minsky dan papert pada tahun 1969 memecahkan masalah *nonlinearly separable* dengan membuat *multilayer network* dengan unit *hidden*, tapi *training* terhadap network tidak bisa dilakukan.

Multilayer perceptron (MLP) disebut juga *multilayer feedforward neural network* merupakan algoritma yang paling luas digunakan. Menurut Wong, Bodnovich dan Selvi pada tahun 1997, sekitar 95% aplikasi bisnis yang menggunakan *neural network*, memakai algoritma ini [23].

Salah satu algoritma yang dipakai dalam proses peramalan adalah *Neural network* yang merupakan satu set unit *input/output* yang terhubung dimana tiap relasinya memiliki bobot [7]. Bobot relasi dalam jaringan dimodifikasi untuk meminimalkan nilai Mean Square Error (MSE) antara nilai prediksi dari jaringan syaraf tersebut dilakukan dalam arah mundur, dari *output layer* hingga *layer* pertama dari *hidden layer* sehingga metode ini disebut *backpropagation* [7].

Penemuan algoritma *backpropagation* untuk *multilayer perceptron*, merupakan metode yang sistematis untuk *training* sehingga bisa dilakukan dan lebih efisien. Algoritma *backpropagation* berasal dari *learning rule* Window dan Hoff, disusun oleh Werbos pada tahun 1974, dibuat oleh Parker pada tahun 1985, Rumelhart Hinton dan Williams pada tahun 1986 dan peneliti lainnya [19].



Sumber : (Han & Kamber, 2007: p331)

Gambar 2.3. Arsitektur Jaringan Backpropagation

Pada gambar 2.3 terdiri dari lapisan input yaitu Y_1, Y_2 sampai dengan Y_n ; seluruhnya dijumlahkan dengan bobot masing-masing dan ditambahkan bias dengan satuan j . Fungsi aktivasi yang digunakan, antara lapisan input dan lapisan

tersembunyi, dan antara lapisan tersembunyi dengan lapisan *output* adalah fungsi aktivasi sigmoid biner.

Dalam *backpropagation*, fungsi aktivasi yang dipakai harus memenuhi beberapa syarat yaitu: kontinu, terdiferensial dengan mudah dan merupakan fungsi tidak turun, salah satu fungsi yang memenuhi ketiga syarat tersebut sehingga sering dipakai adalah fungsi sigmoid biner yang memiliki batasan (0,1)

Pelatihan *backpropagation* meliputi tiga fase, yaitu:

Fase pertama adalah fase maju dimana pola masukan dihitung maju mulai dari layar masukan hingga layar keluaran menggunakan fungsi aktivasi yang ditentukan.

Fase kedua adalah fase mundur, dimana selisih antara keluaran jaringan dengan target yang diinginkan merupakan kesalahan yang terjadi. Kesalahan tersebut dipropagasikan mundur, dimulai dari garis yang berhubungan langsung dengan unit-unit diayar keluaran.

Fase ketiga adalah modifikasi bobot untuk menurunkan kesalahan yang terjadi.

Ketiga fase diatas diulang-ulang terus hingga kondisi penghentian dipenuhi (jumlah iterasi atau kesalahan).

Peningkatan akurasi dan performa yang ada didalam *Neural network* dapat dilakukan dengan pengurangan fitur yang tidak relevan [19]. Feature selection yang digunakan dengan metode wrapper dalam proses mundur atau Backward. Setelah dilakukan pengujian akan dianalisa dan dievaluasi dengan menggunakan Confusion matrix.

3. Metode Penelitian

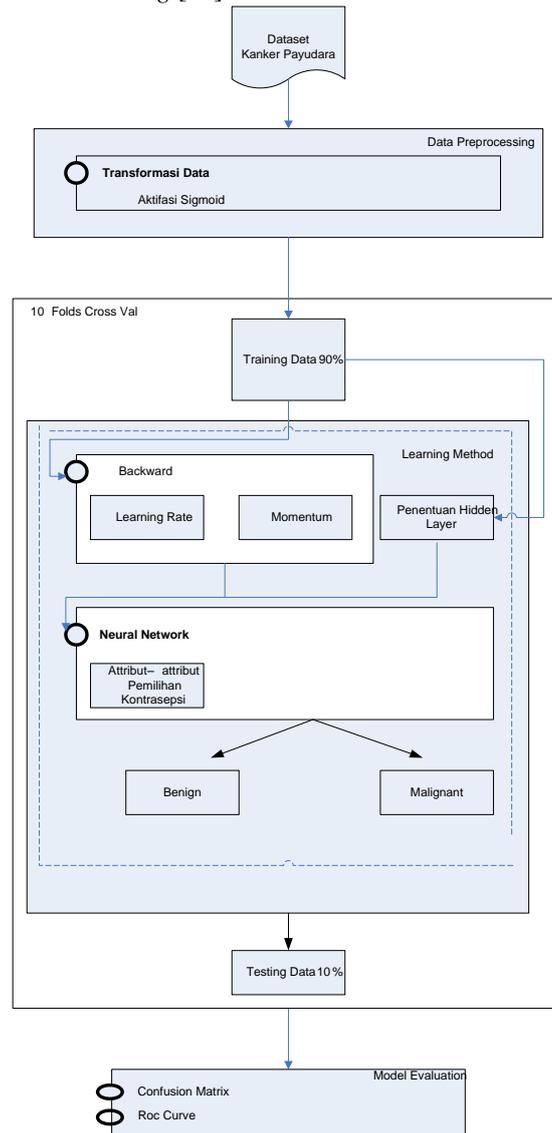
Dataset yang digunakan oleh peneliti diperoleh dari UCI Machine Learning Repository yang dibuat oleh Dr. William H. Wolberg pada 15 Juli 1992. Survei dataset dari Wisconsin Hospitals dengan total 699 data dan atribut berjumlah 11 atribut yang terdiri dari: *Sample code number, Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli, Mitoses, Class (class attribute)*.

Dataset ini akan dilatih dan diuji dengan Backpropagation.

Parameter yang berpengaruh selanjutnya adalah penentuan *Hidden layer*. Pada dasarnya Backpropagation dengan satu *Hidden layer* sudah cukup untuk mampu mengenali sembarang pasangan antara input dan target. Akan tetapi, penambahan jumlah layar tersembunyi kadangkala membuat pelatihan lebih mudah [12].

Parameter-parameter yang ada akan diukur seberapa tingkat akurasi dalam prediksi penentuan kanker payudara. Beberapa metode untuk mengukur performa dari hasil suatu prediksi dalam bentuk perhitungan kesalahan. Perhitungan kesalahan merupakan pengukuran bagaimana jaringan dapat belajar dengan baik sehingga jika dibandingkan dengan pola yang baru akan mudah dikenali. Kesalahan pada keluaran jaringan merupakan selisih antara keluaran sebenarnya dan target atau disebut metode pembelajaran *Supervised learning* [19].

Evaluasi terhadap model yang terbentuk akan dilakukan dengan pengukuran akurasi. Akurasi diukur dengan menggunakan confusion matrix. Confusion matrix akan menggambarkan hasil akurasi mulai dari prediksi positif yang benar, prediksi positif yang salah, prediksi negatif yang benar, dan prediksi negatif yang salah [7]. Nilai akurasi dari model akan dibandingkan antara model yang terbentuk dengan algoritma *neural network* dan algoritma *neural network* yang sudah dioptimasi.



Gambar 3.1. Model Pemikiran Penelitian

Gambar 3.1 merupakan model pemikiran penelitian dimana data yang diperoleh dari UCI Machine Learning Repository tentang *Breast Cancer* pada *data processing* akan ditransformasi kedalam bentuk biner dalam fungsi aktivasi sigmoid, untuk dapat diproses kedalam algoritma

Neural network untuk mendapatkan jaringan terbaik dari *Neural network* yang nantinya akan dilakukan proses feature selection dengan metode. Hasil pengujian berupa Confusion matrix dan kurva ROC. Dikarenakan fungsi aktivasi yang dipakai fungsi aktivasi sigmoid (biner), data harus

ditransformasikan dulu karena batasan keluaran fungsi aktivasi sigmoid adalah [0,1], tapi akan lebih baik jika ditransformasikan ke interval yang lebih kecil, misal pada interval [0.1,0.9] [22]. Maka, pada data penentuan kanker payudara yang ada dilakukan transform data dengan interval [0.1,0.9], dengan rumus sebagai berikut:

$$x^1 = \frac{0.8(x - a)}{b - a} + 0.1$$

$$x^1 = 0.8(x - a)$$

Berikut perhitungan transform dengan menggunakan fungsi aktivasi sigmoid: Berdasarkan data penentuan kanker payudara pada atribut *Clump Thickness* diketahui nilai minimum (a) = 1

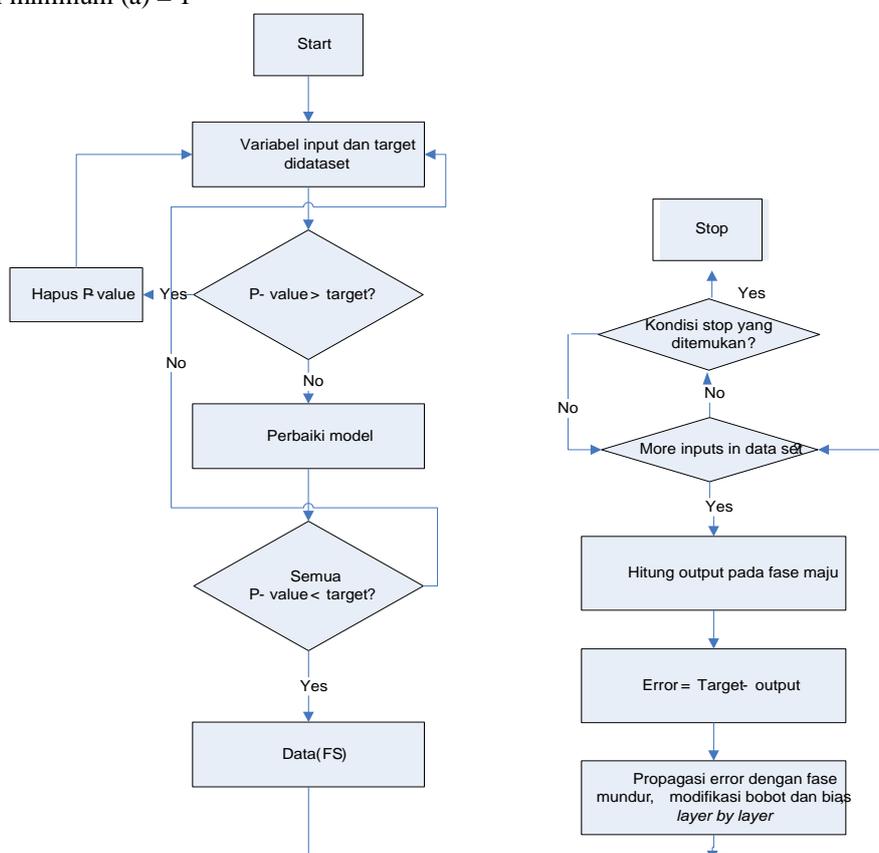
nilai maximum (b) = 10
 data kanker payudara atribut *Clump Thickness* yang akan di transform $x=5$

$$x^1 = \frac{0.8(5 - 1)}{10 - 1} + 0.1$$

$$x^1 = 0.456$$

Selanjutnya seluruh data ditransformasi dari bentuk tabel 1 menjadi tabel 2, maka didapat pola data pelatihan yang dipakai dalam bentuk aktivasi sigmoid.

Aktivasi sigmoid kemudian akan diolah dengan penggabungan antara algoritma *neural network*, *backward elimination*.



Gambar 3.2. Algoritma Penggabungan Neural Network, Backward Elimination

Gambar 3.2 menggambarkan metode algoritma yang diusulkan dalam penelitian ini yaitu *Neural network* berbasis backward elimination, pada tahap preprocessing menggunakan algoritma Backward Elimination dengan memproses variabel input dan target dataset, setiap dihasilkan sebuah bagian maka akan diseleksi dengan algoritma *Neural network* yang digunakan dan dibandingkan dengan bagian terbaik sebelumnya. Salah satu parameter yang

paling berpengaruh adalah penentuan bobot yang ada didalam jaringan. Bobot yang menghasilkan nilai turunan aktivasi kecil sedapat mungkin dihindari karena akan menyebabkan perubahan bobot menjadi sangat kecil, demikian pula dengan nilai bobot awal yang terlalu besar karena akan menyebabkan nilai subset menjadi sangat kecil juga. Oleh karena itu dalam standar Backpropagation bobot dan bias diisi dengan bilangan acak kecil. Jika ditemukan subset terbaik

maka itulah adalah subset terbaik saat ini, perulangan evaluasi akan berakhir sampai ditemukan subset terbaik dari hasil sebelumnya.

Agar jaringan dapat belajar dengan baik maka dibuatlah pola pembelajaran sehingga jika dibandingkan dengan pola yang baru akan mudah dikenali. Beberapa metode untuk mengukur performa dari hasil suatu prediksi dalam bentuk perhitungan kesalahan. Perhitungan kesalahan merupakan pengukuran bagaimana jaringan dapat belajar dengan baik. Kesalahan pada keluaran jaringan merupakan selisih antara keluaran sebenarnya dan target atau disebut metode pembelajaran *Supervised learning* [19].

4. HASIL DAN PEMBAHASAN

Eksperimen yang dilakukan dengan penentuan beberapa parameter penunjang seperti training cycle, learning rate, momentum dan hidden layer perlu dilakukan untuk mendapatkan hasil yang lebih akurat.

Setelah itu akan muncul nilai akurasi, precision, recall, AUC (Optimistic), AUC (Pessimistic) yang didapat dari hasil eksperimen yang telah dilakukan.

Penggabungan antara *Neural network* dan Feature selection akan menghasilkan nilai akurasi sebesar 96.71 %, Precision 94.31%, recall 96.22%, AUC (Optimistic) 0.990%, AUC (Pessimistic) 0.990%.

Tabel 1. Data Kanker Payudara yang Digunakan

<i>Sample code number</i>	<i>Clump Thickness</i>	<i>Uniformity of Cell Size</i>	<i>Uniformity of Cell Shape</i>	<i>Marginal Adhesion</i>
1000025	5	1	1	1
<i>Single Epithelial Cell Size</i>	<i>Bare Nuclei</i>	<i>Bland Chromatin</i>	<i>Normal Nucleoli</i>	<i>Mitoses</i>
2	1	3	1	1

Tabel 2. Data Uji Pemilihan Kanker Payudara yang Digunakan

<i>Sample code number</i>	<i>Clump Thickness</i>	<i>Uniformity of Cell Size</i>	<i>Uniformity of Cell Shape</i>	<i>Marginal Adhesion</i>
1000025	0.4555556	0.1	0.1	0.1
<i>Single Epithelial Cell Size</i>	<i>Bare Nuclei</i>	<i>Bland Chromatin</i>	<i>Normal Nucleoli</i>	<i>Mitoses</i>
0.1888889	0.1	0.2777778	0.1	0.1

Tabel 3. Nilai Bobot Akhir Untuk Hidden Layer

Node	Hidden Layer (Sigmoid)										Thres-hold
	A	B	C	D	E	F	G	H	I	J	
1	-1.338	2.155	2.355	1.715	5.319	3.107	-4.333	0.788	2.588	2.832	1.908
2	2.886	4.689	-1.701	2.623	3.202	0.471	4.189	0.206	8.936	2.201	-3.685
3	3.648	-4.263	-1.780	3.020	2.376	0.084	0.696	1.638	0.064	0.875	-4.505
4	3.886	-4.775	-0.954	2.691	3.250	0.432	1.931	1.888	0.809	0.943	-4.782
5	0.825	6.655	-0.010	0.848	4.077	0.940	0.106	4.003	3.683	3.961	0.680
6	0.680	-4.551	5.274	2.734	2.742	8.417	7.045	5.556	5.835	2.691	-1.604
7	6.085	-1.607	-3.326	2.813	0.889	4.007	2.446	1.198	5.685	0.532	-6.238

Tabel 4. Nilai Bobot Akhir Untuk Output Layer

CLASS	OUTPUT SIGMOID							THRESHOLD
	1	2	3	4	5	6	7	
Benign	-6.975	8.313	3.507	3.995	-6.988	8.046	3.258	-5.837
Malignant	6.975	-8.314	-3.532	-3.970	6.989	-8.047	-3.260	5.838

Tabel 1 menunjukkan data awal dari UCI Repository tentang Breast Cancer yang kemudian diubah ke bentuk aktivasi Tabel 2. dengan aktivasi sigmoid supaya data dapat diuji dan dilatih. Setelah data diubah menjadi bentuk aktivasi sigmoid maka akan didapat nilai untuk tiap bobot pada setiap node. Untuk output didapat nilai bobot sesuai dengan output layer yang ada seperti output untuk benign pada hidden layer satu bernilai -6.975, pada hidden layer dua bernilai 8.313, pada hidden layer tiga 3.507 dengan nilai bias -5.837. Bobot untuk output Malignant pada hidden layer satu bernilai 6.975, pada hidden layer dua bernilai -8.314, pada hidden layer tiga -3.532 dengan nilai bias 5.838. Setelah semua nilai bobot diketahui maka akan diuji nilai akurasi pada data penentuan jenis kanker payudara dengan menggunakan Rapidminer, berikut nilai akurasi terbaik yang didapat:

Tabel 5. Nilai Akurasi Terbaik

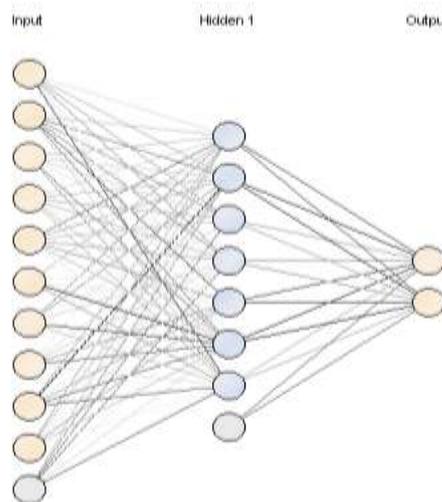
Trainin g Cycle	Learnin g Rate	Momentu m	Siz e	Akuras i
500	0.3	0.2	7	96.42

Nilai akurasi yang terbaik yang terbentuk adalah 96.42% dengan nilai Training Cycle 500, Learning Rate 0.3, Momentum 0.2 dan Size hidden layer sebanyak 7.

	true 2	true 4
pred 2	442	9
pred 4	16	232
class recall	96.51%	96.27%

Gambar 4.1. Hasil Akurasi Neural network Pada RapidMiner

Setelah data diolah dengan beberapa parameter yang ada didapatkan hasil arsitektur jaringan yang dianggap terbaik yaitu dengan nilai akurasi yang paling besar, sebagai berikut:



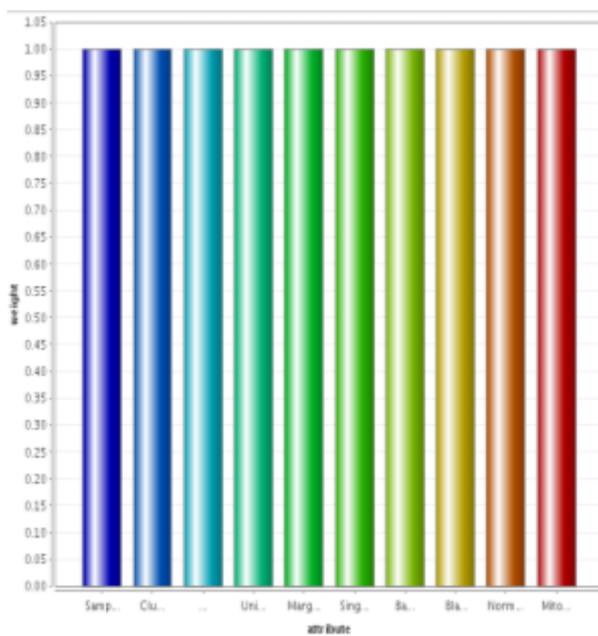
Gambar 4.2. Arsitektur Jaringan yang didapat dari Hasil Eksperimen Neural Network

Arsitektur jaringan hasil eksperimen seperti yang terlihat pada gambar 4.2 dimana pada jaringan tersebut terdiri dari input layer dengan jumlah neuron 10, hidden layer dengan jumlah neuron 7, dan output layer dengan 2 neuron. Node bias (threshold) terdiri dari 2, yaitu terdiri dari 1 node bias pada input layer dan 1 node bias hidden layer. Inisialisasi bobot secara acak untuk input, hidden dan bias. Untuk setiap data pada data training pada tabel 1, dihitung input untuk simpul berdasarkan nilai input dan bobot jaringan saat itu, kemudian berdasarkan nilai input yang didapat untuk membangkitkan output untuk simpul dengan fungsi aktivasi. Hitung nilai error antara nilai yang diprediksi dengan nilai yang sesungguhnya. Selanjutnya backpropagated yaitu balik ke layer sebelumnya untuk menghitung error pada hidden layer (hal ini untuk memperbaharui bobot pada relasi). Hasil perhitungan akhir backpropagation fungsi aktivasi untuk simpul pada hidden layer terdapat pada tabel 3. Nilai akhir pada output layer dihitung menggunakan fungsi aktivasi linear terdapat pada tabel 2.

Eksperimen neural network berbasis backward elimination spesifikasi pada backward elimination yaitu pemilihan pada feature selection (optimize selection) dengan selection direction

terdiri dari forward elimination dan backward elimination yang akan dipilih kemudian ditentukan *limit generation without improval* dan *generation without improval* sebanyak 1, tanpa disertai *limit number of generation* dan *keep best* sebanyak 1.

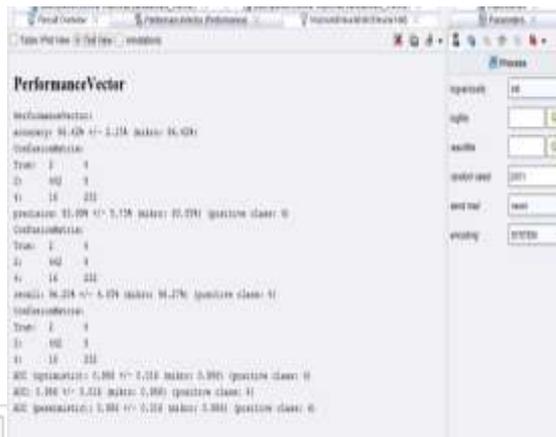
Hasil eksperimen untuk menentukan hasil akurasi terbesar pada hidden layer 1 penggabungan antara *neural network* dan backward elimination terdapat pada training cycle 500, learning rate 0.3, momentum 0.2, Hidden layer 7 dengan nilai akurasi 96.71%.



Gambar 4.3. Plot Nilai Bobot Setiap Atribut Pada Hidden Layer 1 ukuran 3

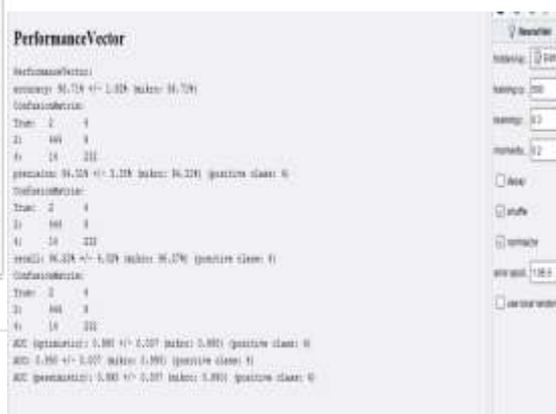
Pada gambar 4.3. menunjukkan, atribut yang dipakai setelah *Neural network* berbasis backward Elimination di jalankan dengan hidden layer 1 dengan ukuran 7 maka, setiap atribut terlihat memiliki ukuran weight yang berbeda-beda

Implementasi dataset hasil Backward Elimination kedalam metode *Neural Network*, hal ini guna mendapatkan arsitektur jaringan yang terbaik. Setelah didapatkan arsitektur jaringan yang terbaik dari hasil eksperimen adalah sebagai berikut:



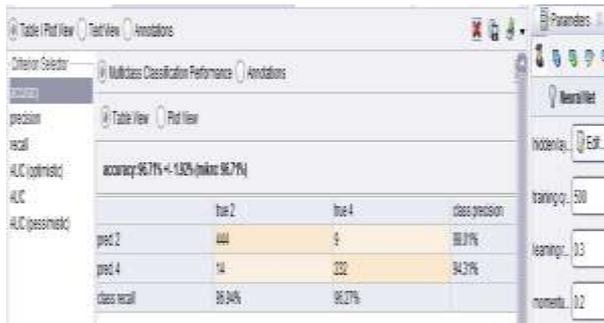
Gambar 4.4. Performance Vector yang Didapat Dari Hasil Eksperimen Neural Network

Gambar 4.4. menunjukkan Performance vector yang didapat dari hasil eksperimen neural network, dan mendapatkan hasil akurasi sebesar 96.42%.



Gambar 4.5. Performance Vector yang Didapat Dari Hasil Eksperimen Neural network Dengan Backward Elimination

Performance Vector yang ditampilkan pada gambar 4.5 menunjukkan bahwa nilai akurasi dataset kanker payudara mengalami peningkatan dengan feature selection berbasis backward elimination, dimana nilai akurasi naik menjadi 96.71%, precision 94.31% dan recall 96.22%.



Gambar 4.6. Hasil Akurasi untuk *Neural network* dan Backward Elimination

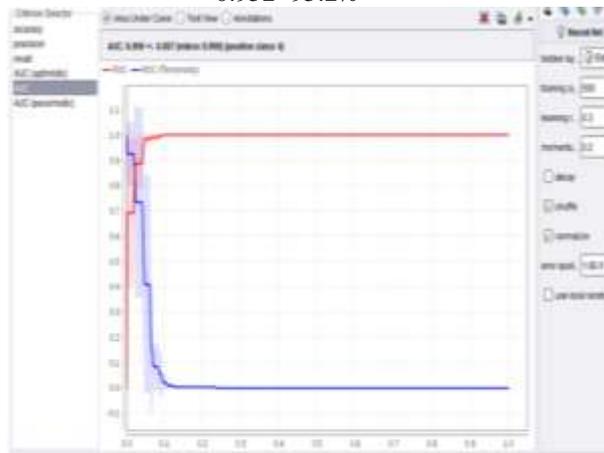
Perhitungan Akurasi, Precision, Recall dan F-Measure sebagai berikut :

$$\text{Akurasi} = \frac{444+232}{444+9+232+14} =$$

$$\text{Precision} = \frac{232}{14+232} =$$

$$\text{Recall} = \frac{232}{9+232} =$$

$$\text{F-Measure} = \frac{2 \times \text{recall} \times \text{precision}}{(\text{recall} + \text{precision})} = \frac{2 \times 0.96}{0.962} = 0.952 = 95.2\%$$



Gambar 4.7. AUC yang Didapat Dari Hasil Eksperimen *Neural network* Dengan Backward Elimination

Berdasarkan gambar 4.6 dan gambar 4.7, model yang didapat adalah model dengan pola 10-7-2 akurasi 96.71% dan terlihat kurva ROC yang ada pada AUC dan pada confusion matrix precision dan recall.



Gambar 4.8. Process Log yang Didapat Dari Hasil Eksperimen *Neural network* Dengan Backward Elimination

Perbandingan eksperimen yang telah dilakukan dengan metode *Neural network* dan metode *Neural network* berbasis Backward Elimination maka didapatkan hasil sebagai berikut:

Tabel 6. Perbandingan Akurasi

Met hod	Trai ning Cycle	Lear ning Rate	Mome ntum	Hid den Layer	Si ze	Aku rasi
NN +BE	500	0.3	0.2	1	7	96.71%
NN	500	0.3	0.2	1	7	96.42%

Berdasarkan tabel 6, Hasil penelitian menunjukkan metode jaringan syaraf tiruan berbasis backward elimination menghasilkan peningkatan akurasi 0.29% dibandingkan hanya dengan menggunakan metode jaringan syaraf tiruan saja. Akurasi yang dihasilkan dengan metode *Neural network* berbasis Backward Elimination dapat meningkatkan hasil akurasi lebih tinggi yaitu 96.71% jika dibandingkan dengan Akurasi yang dihasilkan dengan metode *Neural network* yaitu 96.42%. Hal ini menunjukkan bahwa penentuan jenis kanker payudara dengan menggunakan *Neural network* berbasis Backward Elimination lebih akurat.

5. KESIMPULAN DAN SARAN

Penelitian ini dilakukan pengujian model dengan menggunakan *neural network* dan *neural network* dengan feature selection berupa metode Wrapper dengan penyeleksian berupa Backward Elimination dengan menggunakan data dari UCI yaitu *Breast Cancer*.

Pada penelitian ini secara umum penerapan model NN dan BE dapat meningkatkan akurasi

pada prediksi kanker payudara, akan tetapi karena keterbatasan penelitian ini perlu disarankan untuk melakukan penelitian lanjutan yang berkaitan dengan prediksi untuk mendapatkan akurasi yang lebih baik. Adapun saran-saran yang perlu diberikan yaitu:

1. Perlu dilakukan penelitian yang sejenis dengan variasi proses model misalnya dengan penambahan Cross validation.
2. Perlu dilakukan penelitian yang sejenis dengan variasi metode seleksi dalam pengklasifikasian misalnya feature selection dengan metode embedded atau filter sehingga dapat digunakan untuk perbandingan dari hasil penelitian ini.

6. DAFTAR PUSTAKA

- [1] Alpaydin, Ethem. (2010). *Introduction to Machine Learning*. The MIT Press, London UK.
- [2] Asliyan, Rifat. (2011). *Syllable Based Speech Recognition*. Computer and Information Science. Diambil dari: <http://www.intechopen.com/books/speech-technologies/syllable-based-speech-recognition>. (3 Desember 2014).
- [3] Bevan, Nigel. (1997). *Quality and Usability: A New Framework*. National Physical Laboratory. UK.
- [4] Ciampi, Antonio. Zhang, Fulin. (2002). *A New Approach to Training Backpropagation Artificial Neural Network: Empirical Evaluation on Tens Dataset on Clinical Studies*. McGill University. Canada.
- [5] Gorunescu, Florin. (2011). *Data Mining: Concepts, Models and Techniques*. Verlag Berlin Heidelberg, Springer. Jerman.
- [6] Guillet, Fabrice. Hamilton, Howard J. (2007). *Quality Measures in Data Mining*. Verlag Berlin Heidelberg, Springer. Jerman.
- [7] Han, J & Kamber, Micheline. (2007). *Data Mining Concepts, Models and Techniques*. Second Edition, Morgan Kaufmann Publisher. Elsevier.
- [8] Heaton, Jeff. (2010). *Programming Neural Networks With Encog 2 In Java*. Heaton Research. Inc, USA.
- [9] Hong, X., Harris, C., Brown, M., & Chen, S. (2002). *Backward Elimination Methods for Associative Memory Network Pruning*. Computers and Technology, (Reed 1993).
- [10] Kadhim, Jehan & Abdulrazzaq, Mohammad (2015). *Forecasting USD/IQD Future Values According to Minimum RMSE Rate*. Thi_Qar University. pg.271–285
- [11] Kohavi, R., & John, H. (1997). *Artificial Intelligence Wrappers for feature subset selection*. *elsevier*, 97(97), 273–324.
- [12] Kusumadewi, Sri & Hartati, Sri. (2010). *Neuro-Fuzzy Integrasi Sistem Fuzzy & Jaringan Syaraf*. Second Edition. Yogyakarta: Graha Ilmu.
- [13] Lakshmi, R, Raju Athira, Joy Teena Mary, S. Vijayalakshmi. (2012). *Breast Cancer Factor Preventable and Non-Preventable*. *Departement of Pharmacy Practice*. India.
- [14] Larose, D. (2005). *Discovering Knowledge in Data*. New Jersey, John Willey & Sons. Inc.
- [15] Liao, Warren. T. & Triantaphyllou. Evangelos. (2007). *Recent Advances in Data Mining of Enterprise Data: Algorithms and Applications*. Series: Computer and Operation Research. 6. 190.
- [16] Lim TS, Loh WY, Shih YS. (1999). *A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms*. Kluwer Academic Publishers: Boston.
- [17] Liu, Huan, Yu, Lei. (2005). *Toward Integrating Feature Selection Algorithms for Classification and Clustering*. Department of Computer Science and Engineering. Arizona State University.
- [18] Liu, Yuaning, Wang G., Chen, M., Dong, M., Zhu, X., Wang, S. (2011). *An Improved Particle Swarm Optimization for Feature Selection*. College of Computer Science and Technology. China.
- [19] Maimon, Oded & Rokach, Lior. (2010). *Data Mining and Knowledge Discovery Handbook*, Springer, New York.
- [20] Myatt, Glenn J. (2007). *Making sense of data : A Practical Guide to Exploratory data analysis and Data Mining*. John Wiley & Sons Inc, New Jersey.
- [21] Shukla, Anupam. Tiwari, Ritu. & Kala, Rahul. (2010). *Real Life Application of Soft Computing*. New York: Taylor and Francis Groups, LLC.
- [22] Siang, Jong Jek (2009). *Jaringan Syaraf Tiruan dan Pemrogramannya menggunakan MATLAB*. Penerbit Andi. Yogyakarta.
- [23] Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. Wiley.
- [24] Witten, I. Frank, E., & Hall. (2011). *Data Mining: Practical Machine Learning and tools*. Morgan Kaufmann Publisher, Burlington.