

ARSITEKTUR JARINGAN SYARAF TIRUAN UNTUK PEMODELAN PROSES EKSTRAKSI ATURAN DENGAN SEARCH TREE

Rina Dewi Indah Sari¹ dan Gunawan²

¹Teknik Informatika STMIK ASIA Malang

²Teknik Informatika STTS Surabaya

rideinsar30@gmail.com dan gunawan@stts.edu

ABSTRAK

Jaringan syaraf tiruan akan dipakai untuk ekstraksi rule dalam pengklasifikasian data. Tidak semua hasil pelatihan jaringan syaraf tiruan backpropagation dapat dilakukan proses ekstraksi aturan dengan search tree. Dibutuhkan pemodelan khusus untuk arsitektur jaringan syaraf tiruan sebelum dilakukan pelatihan. Setiap value atribut yang ada harus diwakili oleh sebuah node input pada jaringan. Hal ini mengakibatkan jumlah node input yang cukup signifikan.

Permasalahan yang diangkat adalah disposisi material (bahan baku) pada PT. Gatra Mapan Malang. Dataset terdiri dari 6 atribut non kelas dan 1 atribut kelas dengan total 20 distinct value atribut non kelas dan 2 value atribut kelas. Formulasi permasalahan ke dalam jaringan syaraf dengan menggunakan 20 input neuron yang masing-masing mewakili value atribut non kelas dan 2 output neuron yang mewakili value atribut kelas. Dalam uji coba cross validasi memberikan akurasi rata-rata pelatihan jaringan syaraf tiruan 98.378% dan akurasi rata-rata dari ekstraksi rule sebesar 91.687%.

Kata Kunci: *arsitektur JST, backpropagation, search tree, ekstraksi rule*

ABSTRACT

Artificial neural network will be used for rule extraction in classifying the data. Not all the weight backpropagation neural network training can be done with the rule extraction process search tree. It takes a special modeling for architecture neural network before training. Each existing attribute value must be represented by an input node on the network. This resulted in the number of input nodes significant.

Issues raised is the disposition of raw materials on PT. Gatra Mapan Malang. The dataset consists of 6 non class attributes, one attribute class with a total of 20 distinct values of non class attributes and 2 value of class attribute. Role in the problem formulation to neural network by using 20 input neurons, each represents the value of non class attributes and 2 output neurons representing the class attribute values. Cross validation data show average accuration from training of neural network is 98.378% and average accuration from rule extraction is 91.687%.

Keywords: Classification, artificial neural network, backpropagation, search tree, rule extraction

I. PENDAHULUAN

Para ahli tidak puas hanya dengan tingkat akurasi tinggi yang ditunjukkan oleh jaringan syaraf. Karena cara penalaran yang digunakan jaringan syaraf untuk mencari jawaban tidak dapat dilakukan dengan cepat, ini memberikan bukti perlunya untuk mengekstraksi pengetahuan yang terdapat pada data terlatih jaringan syaraf dan menunjukkannya secara simbolik untuk mendapatkan output-nya.

Salah satu pendekatan yang digunakan untuk ekstraksi aturan adalah metode berbasis pencarian ruang solusi. Metode ini terdiri dari 2 tahap, yang pertama adalah tahap pencarian aturan, dan tahap kedua adalah proses pengecekan validitas rule yang telah diperoleh dari tahap pertama. Pencarian aturan dimulai dari kemungkinan aturan dengan 1 premis sampai aturan dengan banyak premis. Jika aturan dengan 1 premis telah valid, maka tidak perlu dicari aturan dengan premis banyak.

Ekstraksi aturan berbasis pencarian ruang solusi dilakukan terhadap hasil pelatihan jaringan syaraf tiruan metode backpropagation. Namun demikian, tidak semua hasil pelatihan jaringan syaraf tiruan backpropagation dapat dilakukan proses ekstraksi aturan. Hal ini dikarenakan terdapat pemodelan khusus untuk arsitektur jaringan syaraf tiruan yang akan dilakukan proses ekstraksi aturan.

Pada kesempatan kali ini akan dibahas arsitektur jaringan syaraf tiruan untuk pemodelan proses ekstraksi aturan dengan *search tree*. Untuk studi kasus digunakan data disposisi material incoming pada PT. Gatra Mapan Malang yang bergerak dibidang produksi furniture.

II. KAJIAN TEORI

1. Backpropagation-Feed Forward Neural Network

Jaringan syaraf tiruan merupakan salah satu representasi buatan dari otak manusia yang selalu mencoba untuk mensimulasikan proses pembelajaran pada otak manusia tersebut. Backpropagation merupakan algoritma pembelajaran yang terawasi (*supervised learning*) dan biasanya digunakan oleh JST dengan banyak lapisan untuk mengubah bobot-bobot yang terhubung dengan neuron-neuron yang ada pada lapisan tersembunyinya (*hidden layer*).

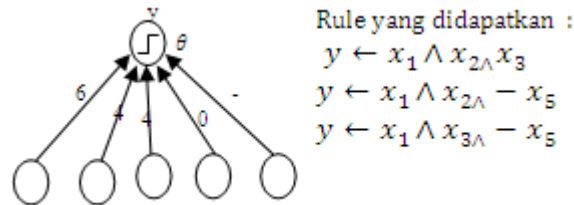
Algoritma backpropagation menggunakan error output untuk mengubah nilai bobot-bobotnya dalam arah mundur (*backward*). Untuk mendapatkan error ini, tahap perambatan maju (*forward propagation*) harus dikerjakan terlebih dahulu. Pada saat perambatan maju, neuron-neuron diaktifkan dengan menggunakan fungsi aktivasi sigmoid, yaitu: $f(x) = 1/(1 + e^{(-x)})$

2. Ekstraksi Rule dari Jaringan Syaraf Tiruan

Proses ekstraksi aturan ini dilakukan setelah tahap pembelajaran dengan jaringan syaraf tiruan selesai dan menghasilkan bobot-bobot pelatihan. Bobot-bobot hasil pelatihan dari jaringan syaraf tiruan tersebut yang nantinya akan digunakan untuk proses ekstraksi rule/aturan.

Gambar 1 menunjukkan hubungan arsitektur ekstraksi dan rule yang dihasilkan serta mengilustrasikan proses ekstraksi rule dari sebuah jaringan syaraf tiruan yang sangat sederhana. Jaringan syaraf tiruan tersebut hanya terdiri dari lapisan input (terdiri dari 5 node input) dan lapisan output yang hanya memiliki 1 node. Nilai boolean true

diwakili oleh nilai 1 dan nilai boolean false diwakili oleh nilai 0. Bentuk rule yang dihasilkan berbentuk *conjunctive rule* dan menggambarkan kondisi-kondisi input yang harus dipenuhi agar output bernilai 1. Sebagai contoh rule $y \leftarrow x_1 \wedge x_2 \wedge \neg x_5$ mengindikasikan bahwa agar y bernilai 1 (*true*) maka input x_1 , x_2 harus bernilai 1 (*true*) dan x_5 harus bernilai 0 (*false*).



Gambar 1. Hubungan antara arsitektur ekstraksi aturan dan rule yang didapatkan

Untuk jaringan syaraf tiruan yang memiliki hidden layer maka jaringan syaraf tiruan tersebut akan dipecah menjadi beberapa bagian. Bagian yang dihasilkan adalah jaringan syaraf tiruan yang terdiri dari input layer dan hidden layer dan jaringan syaraf tiruan yang terdiri dari output layer dan hidden layer.

3. Pendekatan Berbasis Pencarian Ruang Solusi

Pendekatan berbasis pencarian ruang solusi akan digunakan dalam proses ekstraksi aturan. Ada 2 (dua) proses utama yang terdapat dalam pendekatan ini yaitu:

a. Pencarian rule ke dalam ruang kandidat

Proses ini diawali dengan membentuk pohon pencarian aturan. Tiap node merepresentasikan antisenden dari sebuah aturan. Sisi antar node menunjukkan hubungan spesialisasi (dengan arah ke bawah). Node yang terdapat pada lapisan teratas merepresentasikan rule yang paling general dan node pada lapisan terbawah merepresentasikan rule yang paling spesifik yang hanya meng-cover satu kondisi saja. Tidak seperti proses pencarian kebanyakan yang berhenti ketika ditemukan goal node yang pertama, pencarian pada ekstraksi aturan akan terus dilakukan sampai semua (atau sebagian besar) rule-rule yang general telah ditemukan.

b. Test validitas aturan

Proses pengecekan aturan ini tergantung pada fungsi aktivasi yang digunakan oleh jaringan syaraf tiruan. Pengecekan ini dilakukan dengan menghitung nilai aktivasi yang dibangkitkan oleh output layer dan jika nilai aktivasi yang dihasilkan memenuhi suatu kriteria sesuai dengan fungsi aktivasinya maka aturan tersebut dianggap valid.

III. PEMBAHASAN

Data yang digunakan dalam penelitian ini adalah data hasil inspeksi material khusus bahan baku pada PT. Gatra Mapan Malang. Bahan baku (raw material) yang dibahas hanya untuk Sheet, PB, MDF, dan Plywood. Data yang digunakan adalah data hasil inspeksi material oleh bagian QA bahan baku mulai tahun 2011 sampai 2012. Total data hasil inspeksi oleh staf QA yang akan digunakan dalam penelitian ini adalah

1288 record data. Data hasil inspeksi tersebut kemudian dibentuk menjadi dataset. Dataset hasil inspeksi material terdiri dari 7 atribut yaitu 6 atribut non kelas dan 1 atribut kelas.

Tabel 1. Atribut hasil inspeksi material

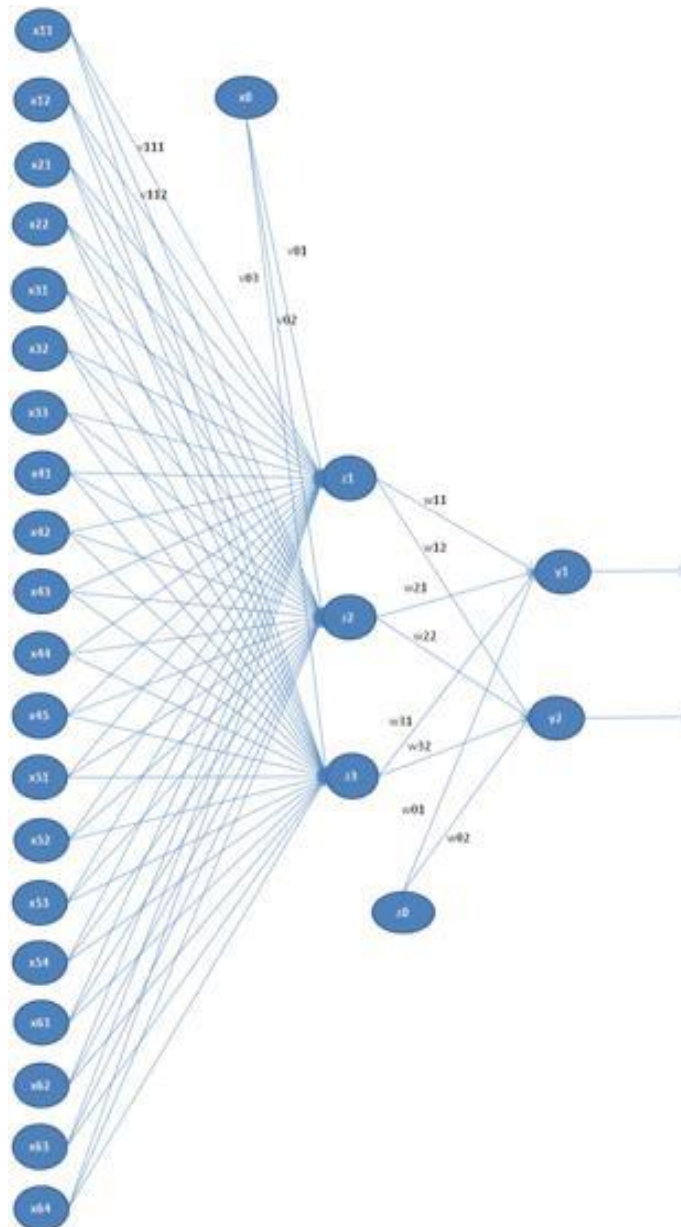
No	Kriteria	Value	Σ Value
1	Hasil laminasi atas	Normal, mruntus (kasar)	2
2	Hasil laminasi bawah	Normal, mruntus (kasar)	2
3	Ketebalan	sesuai, bervariasi, selisih 5mm	3
4	Kondisi	baik, basah, berjamur, cacat, kotor	5
5	Permukaan atas	normal, kasar, berlubang, bergelombang	4
6	Permukaan bawah	normal, kasar, berlubang, bergelombang	4
Total Value Atribut			20

Dataset inspeksi material akan diubah menjadi training set dengan 22 atribut yaitu 20 atribut non kelas ($x_{11}, x_{12}, x_{21}, x_{22}, x_{31}, x_{32}, x_{33}, x_{41}, x_{42}, x_{43}, x_{44}, x_{45}, x_{51}, x_{52}, x_{53}, x_{54}, x_{61}, x_{62}, x_{63}, x_{64}$) dan 2 atribut kelas (y_1, y_2). Masing-masing atribut memiliki 2 distinct value yaitu 0 dan 1. Value 0 menunjukkan bahwa value atribut tersebut tidak terpilih, sedangkan value 1 menunjukkan bahwa atribut tersebut terpilih.

Tabel 2. Trainingset Inspeksi Material

x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	y	y
1	1	2	2	3	3	3	4	4	4	4	4	5	5	5	5	6	6	6	6	1	2
1	2	1	2	1	2	3	1	2	3	4	5	1	2	3	4	1	2	3	4		
1	0	1	0	1	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	1
1	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	1	0	0	0	1
1	0	1	0	1	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0
1	0	1	0	1	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0
1	0	1	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
0	1	1	0	0	1	0	0	1	0	0	0	0	1	0	0	0	1	0	0	1	0
1	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	1	0
0	1	0	1	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	1
0	1	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0

Training set pada tabel di atas selanjutnya akan diformulasikan dalam arsitektur jaringan yang ditunjukkan pada gambar 4. Berdasarkan penotasian yang telah dilakukan, maka training set inspeksi material akan dimodelkan dengan arsitektur 21-3-2. Maksudnya adalah jumlah layer input 21 neuron, jumlah layer hidden 3 neuron dan jumlah layer output 2 neuron. Pada penelitian ini, jumlah neuron pada layer hidden dapat diubah-ubah sesuai setting parameter yang dilakukan sebelum proses pembelajaran. Jumlah layer hidden dibatasi mulai 1 neuron sampai 50 neuron. Pada contoh pembahasan dalam sub bab ini digunakan 3 neuron pada layer hidden.



Gambar 2. Arsitektur JST data Inspeksi Material

PEMBELAJARAN (TRAINING) DATA DENGAN JST BACKPROPAGATION

Untuk melakukan pembelajaran dibutuhkan parameter jaringan saraf tiruan. Parameter jaringan yang digunakan ada 4 (empat) yaitu:

Jumlah node hidden = 3

Learning rate = 0.2

Maksimum epoch = 300

Target error = 0.0001

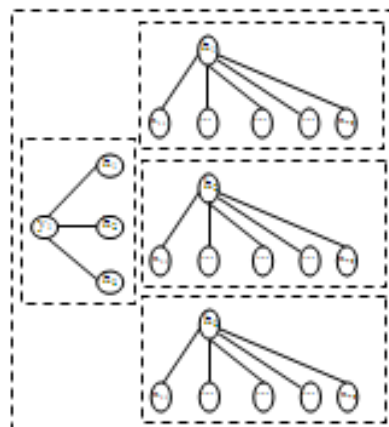
Setelah dilakukan proses pembelajaran dataset inspeksi material menggunakan arsitektur jaringan pada gambar 2, perulangan proses JST berhenti pada epoch ke-84 dengan SSE = 0.0000992. Hal ini menunjukkan bahwa target error telah terpenuhi sebelum mencapai maximum epoch yang ditetapkan.

Tabel 3. Bobot Hasil Pembelajaran Dataset Inspeksi Material

Bobot input dengan hidden			
	z_1	z_2	z_3
x_{11}	$v_{111} = 0.217$	$v_{112} = 0.332$	$v_{113} = 0.4194$
x_{12}	$v_{121} = -0.6848$	$v_{122} = -1.3166$	$v_{123} = 1.536$
x_{21}	$v_{211} = -1.1108$	$v_{212} = 0.9855$	$v_{213} = -0.3673$
x_{22}	$v_{221} = 0.0989$	$v_{222} = -2.0286$	$v_{223} = 2.0713$
x_{31}	$v_{311} = 0.0484$	$v_{312} = 1.4233$	$v_{313} = -0.7604$
x_{32}	$v_{321} = 0.9527$	$v_{322} = -1.1412$	$v_{323} = 1.4271$
x_{33}	$v_{331} = 1.0891$	$v_{332} = 0.1145$	$v_{333} = -0.5105$
x_{41}	$v_{411} = 2.8361$	$v_{412} = 0.6898$	$v_{413} = -3.061$
x_{42}	$v_{421} = -0.994$	$v_{422} = -1.2727$	$v_{423} = 2.1307$
x_{43}	$v_{431} = 0.2714$	$v_{432} = -0.4867$	$v_{433} = -0.412$
x_{44}	$v_{441} = -1.1155$	$v_{442} = -0.6594$	$v_{443} = 2.6838$
x_{45}	$v_{451} = 0.9431$	$v_{452} = 0.7758$	$v_{453} = -0.4735$
x_{51}	$v_{511} = 0.9847$	$v_{512} = 1.923$	$v_{513} = -1.2122$
x_{52}	$v_{521} = 1.2562$	$v_{522} = 0.7097$	$v_{523} = -0.7935$
x_{53}	$v_{531} = -1.2814$	$v_{532} = -3.1435$	$v_{533} = 2.8995$
x_{54}	$v_{541} = -1.3886$	$v_{542} = -1.7173$	$v_{543} = 1.5698$
x_{61}	$v_{611} = -0.9116$	$v_{612} = -0.3871$	$v_{613} = 0.1577$
x_{62}	$v_{621} = -0.2134$	$v_{622} = -1.575$	$v_{623} = 1.5779$
x_{63}	$v_{631} = 0.9597$	$v_{632} = -0.1973$	$v_{633} = -0.4434$
x_{64}	$v_{641} = -0.6791$	$v_{642} = -0.6744$	$v_{643} = 0.2932$
x_0	$v_{01} = -0.2453$	$v_{02} = -0.8291$	$v_{03} = 1.3801$
Bobot hidden dengan output			
	y_1	y_2	
z_1	$w_{11} = 2.6584$	$w_{12} = -2.7302$	
z_2	$w_{21} = 3.7367$	$w_{22} = -3.7734$	
z_3	$w_{31} = -5.7653$	$w_{32} = 5.7471$	
z_0	$w_{01} = -1.2116$	$w_{02} = 1.2997$	

PEMBENTUKAN POHON Pencarian (SEARCH TREE)

Arsitektur jaringan syaraf secara keseluruhan dipecah menjadi beberapa jaringan dengan arsitektur yang lebih kecil yaitu dengan 2 layer saja. Pada pemecahan awal, neuron output (y_1, y_2) akan menjadi inputan dan neuron hidden akan menjadi outputnya. Sedangkan pada pemecahan selanjutnya, layer hidden (h_1, h_2, h_3) menjadi inputan dan layer input menjadi outputnya. Hal ini menunjukkan bahwa dalam setiap kelas bisa menghasilkan beberapa rule dari h_1, h_2 dan h_3 . Demikian pula dalam setiap hidden layer dapat menghasilkan beberapa rule dari $x_{11}, x_{12}, \dots, x_{72}$.

**Gambar 3. Pemecahan Arsitektur Jaringan Syaraf Tiruan untuk Ekstraksi Rule**

Selanjutnya akan dibentuk pohon pencarian aturan. Masing-masing bagian hasil pemecahan arsitektur jaringan pada gambar 5 akan dibentuk menjadi pohon dengan membangkitkan seluruh kemungkinan solusi dalam ruang pencarian. Setiap node yang degenerate akan diurutkan terlebih dahulu sesuai nilai aktivasi yang diharapkan. Berdasarkan bobot hasil pembelajaran akan buat sebuah pohon pencarian.

PENCARIAN RULE DALAM RUANG KANDIDAT SOLUSI

Proses pencarian rule dilakukan mulai dari level 1 yaitu rule dengan 1 atribut dan dimulai dari paling kiri. Proses pencarian dilakukan dengan melakukan pengecekan apakah rule yang ter-generate pada setiap node valid. Apabila hasil pengecekan tidak valid, maka pencarian pada level tersebut berhenti dan dilanjutkan pencarian pada level selanjutnya. Hal ini dikarenakan node tersebut sudah diurutkan sehingga apabila node disebelah kirinya tidak valid maka dapat dipastikan bahwa node disebelah kanannya juga tidak valid. Hal ini akan mempercepat proses pencarian rule.

OPTIMASI RULE

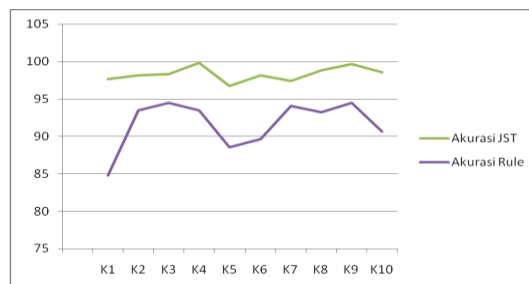
Sesuai dengan arsitektur jaringannya, maka seluruh rule hasil optimasi (18 rule) selanjutnya akan digabungkan menggunakan operator OR sehingga menjadi 2 buah rule sesuai atribut kelasnya. Sehingga menghasilkan dua buah rule sebagai berikut:

1. IF (permukaan_atas=1 AND ketebalan=1 AND hasil_laminasi_bawah=1 AND kondisi=1) OR (kondisi=1 AND permukaan_atas=1 AND ketebalan=1 AND permukaan_bawah=1) OR (kondisi=1 AND permukaan_atas=2) OR (kondisi=1 AND ketebalan=3 AND hasil_laminasi_atas=1) OR (kondisi=1 AND ketebalan=3 AND permukaan_atas=1) THEN disposisi=1
2. IF (hasil_laminasi_bawah=2 AND ketebalan=2) OR (permukaan_atas=4 AND hasil_laminasi_bawah=1 AND permukaan_bawah=1 AND hasil_laminasi_atas=2) OR (hasil_laminasi_bawah=2 AND permukaan_bawah=2) OR (permukaan_atas=3) OR (permukaan_atas=4 AND hasil_laminasi_bawah=1 AND permukaan_bawah=1 AND ketebalan=1) OR (permukaan_atas=4 AND kondisi=4) OR (permukaan_atas=3 AND ketebalan=2) OR (permukaan_atas=3 AND hasil_laminasi_atas=2) OR (permukaan_atas=3 AND hasil_lmns_atas=1 AND permukaan_bawah=1) OR (permukaan_atas=3 AND kondisi=4) OR (hasil_laminasi_bawah=2 AND hasil_laminasi_atas=2) OR (permukaan_atas=3 AND permukaan_bawah=2) OR (permukaan_atas=3 AND hasil_laminasi_atas=1 AND hasil_laminasi_bawah=1 AND kondisi=3) THEN disposisi=2

UJI COBA

Pengujian ini bertujuan untuk mengukur akurasi bobot hasil pelatihan dan rule yang dihasilkan terhadap dataset.

No	Data Testing	Jumlah Epoch	SSE	Akurasi JST	Akurasi Rule	Waktu (menit)
1	K1	15	0.00099	97.67	84.78	159
2	K2	15	0.00094	98.23	93.48	135
3	K3	15	0.00093	98.34	94.48	133
4	K4	15	0.00091	99.87	93.48	144
5	K5	15	0.00092	96.78	88.54	152
6	K6	14	0.00097	98.23	89.67	166
7	K7	14	0.00099	97.44	94.10	172
8	K8	14	0.00092	98.90	93.23	151
9	K9	13	0.00094	99.67	94.45	153



Gambar 4. Hasil Uji Coba

Testing (validasi) dilakukan pada setiap fold data yang dimiliki, sisa fold data yang tidak digunakan untuk validasi akan digunakan sebagai data pelatihan. Dari hasil uji coba terlihat bahwa setiap fold data menghasilkan akurasi yang cukup besar, di dapatkan akurasi rata-rata dari pelatihan jaringan syaraf tiruan sebesar 98.378% dan akurasi rata-rata dari ekstraksi aturan sebesar 91.687%.

IV. PENUTUP

Berdasarkan pembahasan yang telah dilakukan bahwa pemodelan dataset untuk pelatihan JST dan ekstraksi aturan harus menggunakan representasi biner. Pemodelan data tidak dapat menggunakan representasi bipolar karena setiap atribut hanya diwakilkan oleh 2 keadaan yaitu terpilih (1) dan tidak terpilih (0). Setiap value input atribut dari dataset harus diwakilkan oleh sebuah node/ neuron input pada pemodelan arsitektur jaringan syaraf tiruan backpropagation yang akan digunakan untuk proses ekstraksi aturan dengan search tree.

V. DAFTAR PUSTAKA

- [1] Bertalya, *Konsep Data Mining Klasifikasi: Pohon Keputusan*, Universitas Gunadarma. 2009
- [2] Chrisna, Temmy, Bijaksana, Arif, Muntina, Eddy, *Ekstraksi Rule Dari Jaringan Syaraf Tiruan Menggunakan Metode Search Based Pada Data Mining*, STT Telkom Bandung. 2005
- [3] Dancey, Darren, *Tree Based Methods for Rule Extraction from Artificial Neural Networks*, Manchester Metropolitan University, United Kingdom. 2008
- [4] Duch, Włodzisław, Kordos, Mirosław, *Search-based Training for Logical Rule Extraction by Multilayer Perceptron*, School of Computer Engineering Institute of Computer Science Nanyang Technological University The Silesian University of Technology, Singapore. 2005
- [5] Kusumadewi, Sri. *Artificial Intelligent (Teknik dan Aplikasinya)*, Graha Ilmu, Yogyakarta. 2003
- [6] Mark W, Craven, *Using Neural Network For Data Mining*, School of Computer Science Carnigie Mellon Univercity, Pittsburgh. 1995