

PREDIKSI PENYAKIT HATI DENGAN MENGGUNAKAN MODEL ALGORITMA NEURAL NETWORK

Wati Erawati

Program Studi Manajemen Informatika, AMIK BSI Jakarta
Jl. Kamal Raya No.18, Ringroad, Cengkareng, Jakarta barat
wati.erawati@yahoo.com

ABSTRACT

Liver is one of vital organs for human body. Hepatosit is the main important part in a liver which is a unique epitel cell configuration. Liver disease should be predicted based on clinically tested because sometimes a doctor usually making a decision by using his/her intuition rather than to collect hidden data in a database. This problem causing a refraction, missed diagnostic, and over medical payment that influence service quality of a patient. Therefore, medically automatic diagnose system will be useful to carry on those problems. In this research, Neural Network algorithm method is used to get liver disease prediction, Neural Network algorithm will be improved by using Adaboost method which is implemented into patient who suffer from liver disease. The result of this experiment method are divided into 80%, 70%, and 60%, the accuracy points are 70.99%, 69.60%, 68.57%.

Key Word: *Liver disease, Neural Network, Predictio*

PENDAHULUAN

Dunia kesehatan mengumpulkan sejumlah besar data kesehatan, namun beberapa data kesehatan masih tersembunyi. Informasi untuk membuat keputusan yang efektif. Dalam bidang kesehatan keakuratan dalam prediksi sebuah penyakit. Dan dunia kesehatan memerlukan keputusan yang efektif dalam mengambil keputusan dan keakuratan dalam prediksi suatu penyakit sangat penting. Hati merupakan salah satu organ yang cukup besar dan penting pada tubuh kita. Bagian yang penting pada hati ini terdiri dari hepatosit, yang merupakan sel epitel dengan konfigurasi yang unik. Pada dasarnya hati ini merupakan kelenjar eksorin, oleh karena mensekresi cairan

empedu yang dialirkan kedalam duodenum. Selain itu juga merupakan kelenjar endokrin dan penyaring darah (Lumongga, 2008). Data mining merupakan langkah penting dalam penemuan pengetahuan dari serangkaian dataset dalam ukuran yang besar. Dalam beberapa tahun terakhir, data mining telah menemukan peranan penting dalam setiap bidang termasuk perawatan kesehatan (Anbarasi, et al, 2010).

Data mining adalah pencarian untuk hubungan dan pola global yang ada di database yang besar tapi tersembunyi diantara sejumlah besar data. Sedangkan menurut (Santoso, 2007) data mining adalah kegiatan pengumpulan, pemakaian dan history untuk menemukan keteraturan, pola atau hubungan dalam set data

berukuran besar. Pada penelitian ini penulis akan menggunakan teknik algoritma *Neural Network* dengan melakukan optimalisasi perulangan (*iterations*) pada model algoritma neural network dengan optimal, sehingga hasil prediksi lebih akurat pada penyakit hati.

BAHAN DAN METODE

Berdasarkan latar belakang di atas, permasalahan yang dapat dirumuskan dan akan dibahas adalah seberapa akurat model algoritma Neural Network memecahkan masalah dalam prediksi penyakit hati. Mengacu pada paparan di atas, tentunya penulis berharap peranan algoritma ini memberikan kontribusi yang besar sebagai penentuan akurasi yang lebih akurat.

Berdasarkan tujuan penelitian di atas, penulis menggunakan metode algoritma dengan *Neural Network* untuk menentukan akurasi dalam prediksi penyakit hati.

HASIL DAN PEMBAHASAN

1. Data Mining

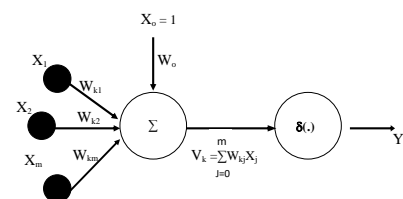
Data mining sering juga disebut *knowledge discovery in database* (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining ini bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan. Sehingga istilah *pattern recognition* sekarang jarang digunakan karena ia termasuk bagian dari data mining (Santoso, 2007).

Data mining meliputi langkah-langkah menentukan variabel atau fitur yang penting untuk dipakai dalam klasifikasi dan regresi. Data mining memegang peran penting dalam bidang industri, kesehatan,

keuangan, cuaca, ilmu dan teknologi. Data mining berkenaan dengan pengolahan data dalam skala besar.

2. Algoritma Neural Network

Neural Network (NN) atau jaringan sayaraf tiruan (JST) adalah suatu usaha untuk meniru fungsi otak manusia. Otak manusia diyakini terdiri dari jutaan unit pengolahan kecil, yang disebut neuron, yang bekerja secara paralel. Neuron saling terhubung satu sama lain melalui koneksi neuron, setiap individu neuron mengambil input dari satu set neuron. Keluaran dikumpulkan oleh neuron lain untuk diproses lebih lanjut. Para otak manusia adalah jaringan kompleks neuron dimana koneksi tetap melanggar dan membentuk. Banyak model mirip dengan otak manusia telah diusulkan (Shukla et al, 2010).



Sumber : Santoso (2007)

Gambar 1. Model Neuron

3. Algoritma Back-Propagation

Salah satu metode untuk mentraining multilayer *neural network* adalah algoritma back-propagasi yang menggunakan *learning rule gradient descent*. Algoritma ini sangat bermanfaat, cukup handal dan mudah dipahami. Selain itu banyak algoritma yang lain mendasarkan prosesnya pada *back-propagation*.

Algoritma mempunyai pengaturan hubungan yang sangat sederhana yaitu: jika keluaran memberikan hasil yang salah, maka penimbang (*weight*) dikoreksi supaya galatnya dapat diperkecil dan respon

jaringan selanjutnya diharapkan akan mendekati nilai yang benar. Algoritma ini juga berkemampuan untuk memperbaiki penimbang pada lapisan tersembunyi (*hidden layer*) (Purnomo & Kurniawan, 2006).

Langkah pembelajaran dalam algoritma backpropagation adalah sebagai berikut:

- Inisialisasi bobot jaringan secara acak (biasanya antara -0.1 sampai 1.0)
- Untuk setiap data pada data training, hitung input untuk simpul berdasarkan nilai input dan bobot jaringan saat itu, menggunakan rumus:

$$Input_j = \sum_{i=1}^n O_i W_{ij} + \theta_j$$

Keterangan:

O_i = Output simpul i dari layer sebelumnya

W_{ij} = bobot relasi dari simpul i pada layer sebelumnya ke simpul j

θ_j = bias (sebagai pembatas)

- Berdasarkan input dari langkah dua, selanjutnya membangkitkan output untuk simpul menggunakan fungsi aktivasi sigmoid :

$$Output = \frac{1}{1 + e^{-input}}$$

- Hitung Nilai *Error* antara nilai yang diprediksi dengan nilai yang sesungguhnya menggunakan rumus :

$$Error_j = Output_j (1 - Output_j) (Target_j - Output_j)$$

Keterangan:

$Output_j$ = Output aktual dari simpul j

$Target_j$ = Nilai target yang sudah diketahui pada data training

- Setelah nilai *Error* dihitung, selanjutnya dibalik ke layer sebelumnya (*backpropagated*). Untuk menghitung nilai *Error* pada *hidden layer*, menggunakan rumus:

$$Error_j = Output_j (1 - Output_j) \sum_{k=1}^n Error_k W_{jk}$$

Keterangan:

$Output_j$ = Output aktual dari simpul j

$Error_k$ = error simpul k

W_{jk} = Bobot relasi dari simpul j ke simpul k pada layer berikutnya

- Nilai *Error* yang dihasilkan dari langkah sebelumnya digunakan untuk memperbarui bobot relasi menggunakan rumus :

$$W_{ij} = W_{ij} + l \cdot Error_j \cdot Output_j$$

Keterangan:

w_{ij} = bobot relasi dari unit i pada layer sebelumnya ke unit j

l = *learning rate* (konstanta, nilainya antara 0 sampai dengan 1)

$Error_j$ = *Error* pada output layer simpul j

$Output_j$ = Output dari simpul i

4. Penyakit Hati

Hati merupakan salah satu organ yang cukup besar dan penting pada tubuh kita. Bagian yang penting pada hati ini terdiri dari hepatosit, yang

merupakan sel epitel dengan konfigurasi yang unik. Pada dasarnya hati ini merupakan kelenjar eksorin, oleh karena mensekresi cairan empedu yang dialirkan kedalam duodenum. Selain itu juga merupakan kelenjar endokrin dan penyaring darah (Lumongga, 2008).

Organ ini mempunyai fungsi yang banyak antara lain sebagai :

1. Pembentukan dan sekresi empedu
2. Tempat penyimpanan glycogen, yang merupakan buffer bagi glukosa darah
3. Sintesa urea
4. Metabolisme kolestrol dan lemak
5. Sintesa dan sekresi endokrin untuk plasma protein, termasuk faktor pembekuan darah
6. Detoksifikasi berbagai macam obat dan racun
7. Membersihkan bakteri dari darah
8. Prosesing beberapa hormon steroid dan reservoar vitamin D
9. Katabolisme hemoglobindari sel darah merah yang tidak terpakai lagi.

Secara umum hepatitis didefinisikan sebagai suatu penyakit yang ditandai dengan suatu peradangan yang terjadi pada organ tubuh seperti hati. Hepatitis diakibatkan berbagai faktor dimana tiap faktor mempunyai karakter khas, maka timbullah berbagai macam hepatitis yang berbeda satu sama lain Gejala atau Tanda-tanda penyakit hepatitis :

Gejala hepatitis A :

1. Penyakit kuning
2. Kuku, mata, kulit penderita mengalami warna kekuningan dan tampak jelas sekali

3. Kelelahan (*fatigue*)
4. Nyeri lambung
5. Kehilangan nafsu makan
6. Mual-mual
7. Diare
8. Demam (*fever*)

Gejala Hepatitis B :

1. Mual-mual
2. Muntah-muntah
3. Hilang nafsu makan
4. Sakit kepala
5. Penyakit kuning

Gejala Hepatitis C :

1. Penurunan nafsu makan
2. Lambung tidak nyaman
3. Mual-mual
4. Muntah-muntah
5. Penyakit kuning

Dilaporkan bahwa 80% penderita hepatitis C kronis, tidak menunjukkan gejala apapun, tapi dapat diketahui menderita hep-C melalui test laboratorium apakah dia mengidap HCV (hepatitis C virus).

Dari mereka yang terkena HCV 40% dapat sembuh total, 60% akan menjadi carrier (tidak menjadi sakit tapi dapat menularkan HCV ke orang lain). Sementara 20% lagi berlanjut menjadi sirosis hati, dan 20% akan menderita kanker hati. Penyakit HCV dilarang minum minuman beralkohol.

Penyebaran Virus Hepatitis :

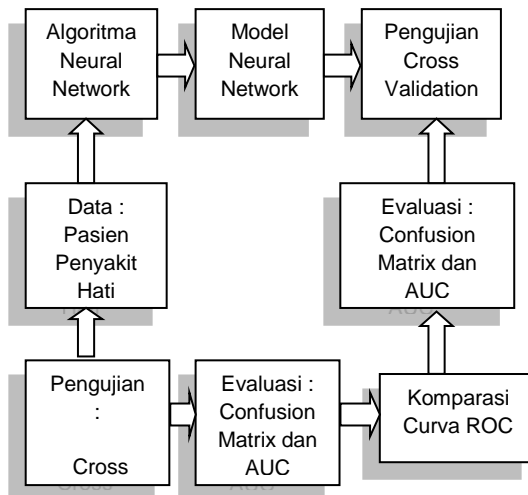
1. HAV
HAV disebarkan melalui kotoran dan tinja penderita, orang-orang yang tidak punya perilaku hidup bersih dan sehat seperti tidak mencuci tangan dengan sabun setelah keluar dari toilet umum, tidak cuci tangan saat akan makan, dll.
2. HBV
HBV ditularkan melalui darah ketika terjadi kontak terutama bagian tangan yang terluka dengan darah yang terinfeksi,

- juga melalui cairan tubuh (sperma).
3. HCV
HCV penyebarannya mirip dengan hepatitis B, yakni melalui darah. Khusus pada hepatitis C, penyebaran terbesar terjadi melalui transfusi darah, terutama pada darah yang belum dilakukan screening. Hingga saat ini hepatitis C belum ada vaksinnya.
 4. Hepatitis Alkoholik
Disebabkan oleh alkohol dan berbeda dengan hepatitis karena virus, meskipun demikian ada juga banyak aspek yang sama. Respon dari tubuh karena minum alkohol secara berlebihan adalah :
 - a. Penambahan suplay darah ke organ yang terkena
 - b. Organ membengkak dan memerah
 - c. Peningkatan agen kekebalan tubuh seperti sel darah putih dan senjata kimianya
 - d. Rasa nyeri.

Penelitian tentang prediksi penyakit hati cukup banyak. Berikut ini beberapa penelitian terkait yang mengangkat topik tentang penyakit hati :

 - a. Komparasi dari beberapa algoritma antara lain Naïve bayes classifier, C4.5, Back Propagation Neural Network, dan Support Vector Machine mengenai database pasien penyakit hati (Ramana et al., 2011). Penelitian ini selain mengkomparasi metode data mining juga mengkombinasi data pasien penyakit hati yang diperoleh.
 - b. Penelitian dengan teknik klasifikasi data mining diantaranya klasifikasi data pasien penyakit hepatitis dengan menggunakan algoritma backpropagation and naïve bayes (Karlik, 2011).
 - c. Penelitian berikutnya dilakukan untuk prediksi penyakit jantung menggunakan metode algoritma C4.5 dan ditingkatkan kinerja dari algoritma C4.5 menggunakan metode adaboost yang diimplementasikan pada data pasien penyakit jantung (Rohman, 2009). Hasil penelitian menunjukkan metode adaboost mampu meningkatkan metode Algoritma C4.5 denganobyek penyakit jantung.
 - d. Penelitian berikutnya dilakukan untuk prediksi penyakit jantung menggunakan metode algoritma *Neural Network* dan ditingkatkan kinerja dari algoritma *Neural Network* menggunakan metode adaboost yang diimplementasikan pada data pasien penyakit jantung. Hasil penelitian menunjukkan metode adaboost mampu meningkatkan metode Algoritma *Neural Network* denganobyek penyakit jantung.
 5. Rancangan Penelitian
Dalam penelitian ini menggunakan data pasien yang melakukan pemeriksaan penyakit hati yang didapat dari UCI (Universitas California Invene) Machine Learning Repository (Janosi & Steinbrunn, 2011). Data pemeriksaan penyakit hati ini akan diolah menggunakan algoritma *Neural Network* dan dengan menggunakan metode optimasi adaboost sehingga diperoleh metode yang paling akurat dan dapat digunakan sebagai aturan dalam prediksi penyakit hati.

Dataset penelitian prediksi penyakit hati untuk training dibagi kedalam tiga tahapan, yaitu dataset dengan 80% berjumlah 424, dataset dengan 70% berjumlah 371, dan dataset dengan 60% berjumlah 318.



Sumber : Santoso (2007)

Gambar 2. Rancangan Penelitian

Tahapan yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Pengumpulan Data

Dalam penelitian ini, pengumpulan data adalah data pasien yang memeriksakan penyakit hati dengan hasil yang didapat sebanyak 582 orang yang diperiksa dan sebanyak 415 pasien terdeteksi sakit, sehingga 167 pasien terdeteksi sehat. Dengan atribut dari setiap penyakit hati yang diperiksa adalah umur, jenis kelamin, total bilirubin, direct bilirubin, Alkaline, Alamine, Aspartate, Total Proteins, Albumin, Ration.

2. Pengolahan Awal Data

Setelah melakukan pengumpulan data dari dataset pengujian, langkah selanjutnya adalah

melakukan pengolahan data. Sehingga dapat menghasilkan perhitungan yang baik.

3. Metode Yang Diusulkan

Melakukan pengolahan data dengan menggunakan Algoritma Neural Network.

4. Eksperimen dan Pengujian Metode

Pengolahan data dengan metode yang telah ditentukan menggunakan RapidMiner 5.

5. Evaluasi dan Validasi Hasil

Pada tahap evaluasi, disebut tahap klasifikasi karena pada tahap ini akan ditentukan pengujian untuk akurasi. Tahap pengujiannya adalah melihat hasil akurasi pada proses klasifikasi algoritma *Neural Network* dan klasifikasi Algoritma *Neural Network* berbasis metode adaboost serta evaluasi dengan *ROC Curve*.

6. Pengolahan Data Awal

Setelah melakukan pengumpulan data dari dataset pengujian, langkah selanjutnya adalah melakukan pengolahan data. Sehingga dapat menghasilkan perhitungan yang baik. Untuk mendapatkan data yang berkualitas, beberapa teknik yang dilakukan sebagai berikut (Vercellis, 2009) :

1. Data Validation

Untuk mengidentifikasi dan menghapus data ganjil, data yang tidak konsisten dan data yang tidak lengkap (*missing value*).

2. *Data Integration and Transformation*

Untuk meningkatkan akurasi dan efisiensi algoritma. Data yang digunakan dalam penulisan ini bernilai kategorikal. Data

ditransformasikan ke dalam software rapidminer. Tabel atribut dan tabel kategorikal atribut terlihat di tabel 1.

Sebelum dilakukan pelatihan pada dataset, data primer harus dinormalisasikan dahulu dengan tujuan agar proses pelatihan berjalan dengan cepat dan mampu digunakan untuk melakukan pelatihan. Persamaannya sebagai berikut :

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

Keterangan : x^1 = Nilai Transform
 x = Nilai Asli
 a = Nilai Minimal
 b = Nilai Maximal

Tabel 1. Data Input Output

Keterangan	Variabel	Inisial
Input	Age	X1
	Gender	X2
	TtlBilirubin	X3
	DirectBilirubin	X4
	Alkaline	X5
	Alamine	X6
	Aspartate	X7
	TtlProtein	X8
	Albumin	X9
	Ratio	X10
Output	Hasil	Y

Sumber : Hasil Uji Coba Rapid Miner

Perhitungan Neural Network

Pembuatan model *Neural Network* dilakukan pada dataset yang terdiri dari 10 atribut yang merupakan atribut dari penyakit hati, dan kelas yang mewakili sakit atau sehat. Atribut penyakit hati merupakan faktor yang dibuat secara acak, dan kelas merupakan hasil penyakit hati (sehat atau sakit).

Dari dataset yang ada, penelitian ini mengambil salah satu eksperimen penyakit hati yang terdiri dari 10 simpul yaitu 1,2,3,4,5,6,7,8,9,10 yang merupakan atribut dari dataset penyakit hati. Hidden Layer sebanyak 2 simpul yaitu X dan Y dan memiliki output layer terdiri dari 2 simpul yaitu O yang mempresentasikan Output atau Class.

Langkah – langkah dalam perhitungan training *Neural Network* (Algoritma *Backpropagation*) :

1. Inialisasi bobot awal secara acak

Nilai input awal tergambar pada tabel 4.1 dimana tiap simpul di input layer diambil dari dataset uji penyakit hati, sedangkan nilai bobot awal untuk input layer dan hidden layer pada tabel 4.2 dimana diberi random -100 s/d 100. Begitu pula di tabel 4.3 adalah simpul bias yang terhubung dengan hidden layer dan output layer.

Tabel 2. Nilai Input Awal

I ₁	I ₂	I ₃	I ₄	I ₅	I ₆	I ₇	I ₈	I ₉	I ₁₀	Output/Class
6	0	0	0	1	1	1	6	3	0	
5	.	.	.	8	6	8	.	.	.	1
	1	7	1	7			8	3	9	

Sumber : Hasil Uji Coba Rapid Miner

Data tabel di atas dari I₁ sampai I₁₀ merupakan data input yang diperoleh dari data pasien penyakit hati.

Tabel 3. Nilai Bobot Awal

W _{1x}	W _{2x}	W _{3x}	W _{4x}	W _{5x}	W _{6x}	W _{7x}	W _{8x}	W _{9x}	W _{10x}
5.	3.	2.	1.	11	11	-	0.	0.	-
2	6	3	3	.1	.1	1	8	3	3.
6	9	9	9	03	31	1	4	2	73
8	2	4	8			8	9	1	6

W _{1y}	W _{2y}	W _{3y}	W _{4y}	W _{5y}	W _{6y}	W _{7y}	W _{8y}	W _{9y}	W _{10y}
-	-	0.	6.	1.	4.	1.	15	-	-
4.	0.	82	4	5	3	3	.2	3.	10
09	65	3	0	2	7	8	23	5	.6

4	9	7	2	7	7	2	59
						2	3

Sumber : Hasil Uji Coba Rapid Miner

Data tabel 3 di atas dari W_{1x} sampai W_{10y} merupakan data hidden layer yang diperoleh dari pengolahan dengan Neural Network yang terhubung dengan nilai input output di tabel 2

Tabel 4. Tabel Nilai Bias awal

B _x	B _y	B ₀₁	B ₀₂
16.992	4.818	0.020	-0.020

Sumber : Hasil Uji Coba Rapid Miner

Data tabel 3. B_x, B_y, B₀₁, dan B₀₂ merupakan nilai bias dari input, output dan hidden layer yang diolah dengan Neural Network.

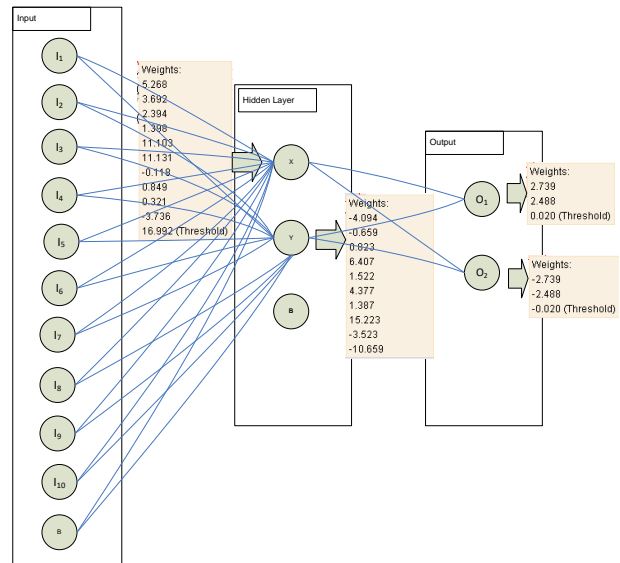
2. Menghitung Input, Output dan Error

$$Input_x = (I_1 * W_{1x}) + (I_2 * W_{2x}) + (I_3 * W_{3x}) + (I_4 * W_{4x}) + (I_5 * W_{5x}) + (I_6 * W_{6x}) + (I_7 * W_{7x}) + (I_8 * W_{8x}) + (I_9 * W_{9x}) + (I_{10} * W_{10x}) + B_x$$

$$= (65 * 5.268) + (0.1 * 3.692) + (0.7 * 2.394) + (0.1 * 1.398) + (187 * 11.103) + (16 * 11.131) + (18 * 0.118) + (6.8 * 0.849) + (3.3 * 0.321) + (0.9 * -3.736) + 16.992$$

$$= (342.42) + (0.3692) + (1.6758) + (0.1398) + (2076.261) + (178.096) + (2.124) + (5.7732) + (1.0593) + (-3.3624) + 16.992 = 2.621.5479$$

$$Output_x = 1 / (1 + e^{2.621.5479}) = 14.5909880$$



Sumber : Hasil Uji Coba Rapid Miner

Gambar 3. Arsitektur Neural Network

KESIMPULAN

Dari penelitian yang dilakukan, penentuan nilai parameter dengan model yang dibentuk oleh Neural Network menghasilkan akurasi yang lebih baik. Hasil ini sangat penting dalam penentuan parameter mana yang paling baik, sehingga menghasilkan akurasi yang tinggi. Hasil akurasi dapat dilihat dengan model algoritma Neural Network menghasilkan akurasi sebesar 70,99%.

Adapun dari model yang terbentuk, nantinya dapat dikembangkan atau diimplementasikan ke dalam sebuah aplikasi. Sehingga dapat membantu dan memudahkan bagi para praktisi kesehatan dalam mendiagnosa penyakit hati, dan hasil diagnosa pun bisa lebih akurat dan terpercaya.

Didalam penulisan ini pun penulis memberikan saran, adapun saran dari penulis antara lain adanya

penelitian lebih lanjut dengan dataset yang lebih banyak dan lebih baik, selain itu algoritma ini dapat digunakan pada studi kasus lain, tidak hanya dalam bidang kesehatan tetapi juga pada bidang –bidang ilmu lainnya.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada tim jurnal techno yang telah meluangkan waktu untuk memberikan dan memeriksa template jurnal ini.

DAFTAR PUSTAKA

- Anbarasi, M., Anupriya, E., & Iyengar, N. (2010). Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm. *International Journal of Engineering Science and Technology Vol. 2(10), 2010, 5370-53,5370-5376.*
- Janosi, A., & Steinbrunn, W. (2011, November 2013). *UCI Machine Learning Repository*. Retrieved from UCI Machine Repository.
- Karlik, Bekir. 2011. *Hepatitis Disease Diagnosis Using Backpropagation and the Naïve Bayes Classifier*. Konya, Turkey.
- Lumongga, Fitriani. 2008. Struktur Liver. Medan
- Purnomo, M. H., & Kurniawan, A. (2006). *Supervised Neural Network dan Aplikasinya*. Yogyakarta: Graha Ilmu.
- Rohman, Abdul. 2009. Penerapan Algoritma C4.5 Berbasis Adaboost Untuk Prediksi Penyakit Jantung. Medan
- Santoso, B. 2007. *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Shukla, A., Tiwari, R., & Kala, R. 2010. *Real Life Applications of Soft Computing*. United States of America on: Taylor and Francis Group, LLC.
- Wu, X., & Kumar, V. 2009. *The Top Ten Algorithms in Data Mining*. Boca Raton, London, New York: Taylor & Francis Group, LLC.
- Wu, X., & Kumar, V. 2009. *The Top Ten Algorithms in Data Mining*. Boca Raton, London, New York: Taylor & Francis Group, LLC.

