# KOMPARASI ALGORITMA *MULTI LAYER PERCEPTRON* DAN *RADIAL BASIS FUNCTION* UNTUK DIAGNOSA PENYAKIT JANTUNG

**Ahmad Setiadi**

Akademi Manajemen Informatika dan Komputer Bina Sarana Informatika Karawang
Program Studi Manajemen Informatika
Jl. Banten No. 1 Karangpawitan, Karawang Telp : (0267)8454893
ahmad.ams@bsi.ac.id

***ABSTRACT***

*Neural network as a data mining model has many algorithms with different accuracy level. This research use UCI machine learning repository's data to compare the accuracy level of Multilayer Perceptron (MLP) and Radial Basis Function (RBF) algorithm in predicting the heart disease. Confusion matrix and the ROC (Receiver Operating Characteristic) curve method is used to measure the performance of the both algorithms. Based on the test results of the implementation, proved that the MLP algorithm has the higher value of accuracy than the RBF algorithm. Using the Confusion Matrix, the MLP algorithm has higher value of accuracy with 87.3% than the RBF algorithm with 81.1%. Using the ROC curve, the MLP algorithm also has higher AUC (Area Under the Curve) value with 0.949 than the RBF algorithm with 0.911. Using confusion matrix, the accuracy value of both algorthms are included as good classification, because the AUC value is in the range of 0.80 until 0.90. Using ROC Curve, the accuracy values are included as excellent classification, because the AUC value is in the range of 0.90 until 1.00.*

***Keywords : Heart Disease, Neural Network Model, Multilayer Perceptron, Radial Basis Function, Confusion Matrix, ROC Curve***

## PENDAHULUAN

Penyakit jantung adalah penyakit *cardiovascular* yang menjadi masalah kesehatan utama di seluruh negara. Penyakit jantung saat ini menjadi salah satu penyakit penyebab kematian tertinggi. Kematian akibat penyakit jantung di seluruh Amerika Serikat mencapai 959.227 penderita, yakni 41,4 % dari seluruh kematian atau setiap hari terdapat 2.600 penduduk meninggal akibat penyakit jantung (Hananta & Muhammad, 2011). Di Indonesia, kematian yang disebabkan penyakit jantung mencapai 9,7% dengan 5,1% merupakan penderita jantung iskemik (Departemen Kesehatan RI, 2009).

Jantung adalah salah satu organ vital tubuh yang berfungsi untuk memompa darah bersih ke seluruh tubuh dan darah kotor ke paru-paru. Jika terjadi gangguan pada jantung maka fungsi pemompaan darah akan terganggu bahkan bisa berakibat pada kematian. Penyakit Jantung adalah penyempitan atau penyumbatan *(arteriosclerosis)* pembuluh arteri koroner yang disebabkan oleh penumpukan dari zat-zat lemak (kolesterol, *trigliserida*) yang makin lama makin banyak dan menumpuk di bawah lapisan terdalam *(endotelium)* dari dinding pembuluh nadi. Akibat dari adanya peyumbatan maka dengan sendirinya suplai energi kimiawi ke otot jantung berkurang, sehingga terjadi gangguan keseimbangan antara suplai dan kebutuhan darah (Hananta & Muhammad, 2011).

Dengan tersumbatnya arteri koroner, maka hal ini akan mengurangi atau menghentikan aliran darah men*supply* oksigen ke otot-otot jantung, sehingga mengganggu kerja jantung sebagai pemompa darah. Bila sampai otot-otot jantung kekurangan *supply* darah maka jantung akan menjadi lemah dan tidak dapat menyediakan darah ke seluruh bagian tubuh. Faktor gejala yang terdiagnosa sebagai penyakit jantung diantaranya adalah jenis sakit dada *(chest pain)*¸ tekanan darah tinggi *(tresbps)*, kolesterol *(chol)*, nilai tes EKG *(resting electrodiagraphic (restecg)*), denyut jantung *(thalach)* dan kadar gula (*fasting blood sugar* (FBS) (Mahmood & Kuppa, 2010).

Penyakit jantung meliputi *aortic regurgition*, *cardiogenic shock*, *congenital heart disease*, *coronary artery disease* (CAD), *hypertrophic cardiomyopathy*, *ischemic cardiomyopathy*, *peripartum*

*cardiomyopa-thy*, *tricuspid regurgitation*, dan lain-lain. (Shukla, Tiwari, & Kala, 2010)

Diagnosa penyakit jantung perlu diprediksi karena keputusan klinis seringkali dibuat oleh dokter berdasarkan intuisi dan bukan pengalaman pada pengetahuan yang didapat atas data yang tersembunyi dalam database (Palaniappan & Awang, 2008).

Teknik analisa konvensional secara manual yang selama ini digunakan tidak lagi efektif digunakan untuk mendiagnosa karena teknik ini menyebabkan bias, kesalahan dan biaya medis yang berlebihan yang mempengaruhi kualitas layanan yang diberikan kepada pasien.

Seiring dengan perkembangan sistem berbasis pengetahuan medis, tuntutan akan adanya penggunaan sistem pengetahuan berbasis komputer sebagai teknik analisa dalam mendiagnosa penyakit menjadi semakin penting. oleh karenanya, saat inilah waktu yang tepat untuk mengembangkan sistem pengetahuan berbasis komputer yang modern, efektif dan efisien dalam mendiagnosa penyakit (Mehdi & Yaghobi, 2009).

Untuk mengurangi waktu yang dibutuhkan untuk mendiagnosa dan meningkatkan tingkat akurasi diagnosa, dibutuhkan adanya pengembangan suatu sistem diagnosa medis yang baik dan dapat dihandalkan. Diagnosa medis secara alamiah sangat sulit dan membutuhkan proses berfikir yang rumit. Oleh karena itu, metode *soft computing* seperti *neural network* menunjukkan potensi yang menjanjikan untuk dikembangkan guna menghasilkan diagnosa medis yang tepat untuk penyakit jantung (Hanan, Manza, & Ramteke, 2010).

Penggunaan *data mining* dengan model *neural network* sebagai salah satu pilihan untuk diagnosa penyakit jantung dapat menjadi alternatif pilihan yang tepat. Dalam model *neural network* tersebut terdapat beberapa algoritma, diantaranya *Multilayer Perceptron* (MLP) dan *Radial Basis Function* (RBF). Namun sampai saat ini belum diketahui algoritma yang paling akurat dalam penentuan diagnosa untuk prediksi penyakit jantung. Untuk itu maka dalam penelitian ini akan dilakukan komparasi *data mining* berbasis *neural network* untuk mengetahui algoritma yang memiiki akurasi lebih tinggi dalam mendeteksi penyakit jantung.

## BAHAN DAN METODE

Berdasaran tipenya, pengumpulan data dibagi dua, yaitu data primer dan data sekunder. Data primer adalah data yang dikumpulkan pertama kali dan untuk melihat apa yang sesungguhnya terjadi, sedangkan data sekunder adalah data yang sebelumnya pernah dibuat oleh seseorang baik diterbitkan atau tidak (Kothari, 2004).

Pada penelitian ini, sumber data diperoleh secara sekunder, yaitu melalui pengumpulan data yang pernah dibuat sebelumnya dan dipublikasi melalui *website*, yaitu data diambil melalui alamat web: http://archive.ics.uci.edu/ml/.

Jenis penelitian yang dilakukan adalah jenis penelitian eksperimen komparatif, yaitu metode penelitian yang menguji kebenaran sebuah hipotesis dengan statistik dan menghubungkannya dengan masalah penelitian. Jenis penelitian ini membandingkan dua objek yang berbeda, misalnya membandingkan dua algoritma yang berbeda dengan melihat hasil statistik masing-masing mana yang lebih baik (Kothari, 2004).

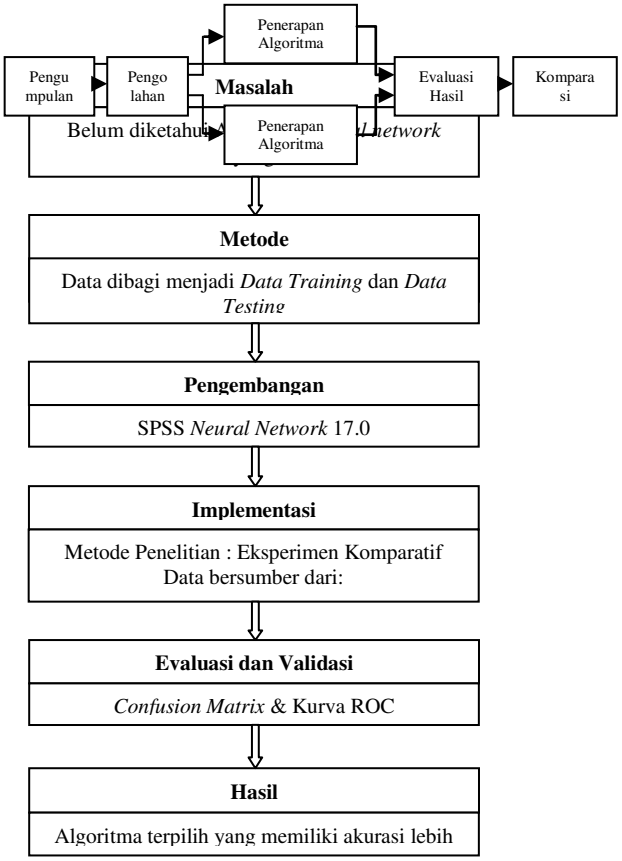
Dalam penelitian ini, penulis menyusun sebuah kerangka pemikiran penelitian yang akan digunakan sebagai pedoman atau acuan, sehingga penelitian ini dapat dilakukan dengan konsisten. Kerangka pemikiran tersebut adalah:

1. **Masalah**
   Permasalahan yang ingin dipecahkan pada penelitian ini adalah belum diketahuinya algoritma *neural network* yang memiliki akurasi lebih tinggi.
2. **Metode**
   Untuk dapat diketahui mana algoritma yang memiliki tingkat akurasi yang lebih tinggi, data akan dibagi menjadi *Data Training*, yaitu data yang digunakan untuk Pembentukan model dan *Data Testing,* yaitu data yang digunakan untuk menguji model yang terbentuk. Data tersebutk akan diolah menggunakan algoritma MLP dan *RBF*.
   Hasil pengolahan data kedua algoritma tersebut nantinya akan dikomparasi.
3. **Pengembangan**
   Perangkat lunak SPSS *Neural Nework* 17.0 digunakan untuk megolah data.
4. **Implementasi**
   Pada penelitian objek ini, penelitian yang dijadikan dasar penelitian adalah data pasien yang melakukan pemeriksaan jantung yang didapat dari UCI

(University California, Irvene) *Machine Learning Repository.*

5. **Evaluasi dan Validasi**
   Hasil pengolahan data menggunakan kedua algoritma yang digunakan, yaitu MLP dan RBF akan dievaluasi dengan menggunakan metode *Confusion Matrix* dan Kurva ROC.
6. **Hasil**
   Hasil yang diharapkan dari penelitian ini adalah diketahuinya algoritma yang memiliki akurasi yang lebih.

Untuk mendukung kerangka pemikiran di atas, penelitian ini dilakukan dengan menjalankan beberapa langkah proses penelitian seperti terlihat pada gambar berikut :




Gambar 1. Kerangka Pemikiran Penelitian

Gambar 2. Tahap Penelitian

## 7. Pengumpulan Data

Data yang digunakan dalam penelitian ini bersumber dari alamat web: http://archive.ics.uci.edu/ml/. Data merupakan hasil pemeriksaan terhadap 270 orang yang melakukan pemeriksaan jantung. Variabel atau atribut yang ada sebanyak 14 variabel, yang terdiri dari 13 variabel *predictor* dan 1 variabel tujuan, seperti terlihat pada Tabel 1:

Tabel 1. Variabel Data

| | Name | Type | Width | Decimals | Label | Values | Missing | Columns | Align | Measure |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | age | Numeric | 11 | 0 | | None | None | 11 | Right | Scale |
| 2 | sex | Numeric | 11 | 0 | | None | None | 11 | Right | Nominal |
| 3 | CPT | Numeric | 11 | 0 | | None | None | 11 | Right | Nominal |
| 4 | RBP | Numeric | 11 | 0 | | None | None | 11 | Right | Scale |
| 5 | SC | Numeric | 11 | 0 | | None | None | 11 | Right | Scale |
| 6 | FBS | Numeric | 11 | 0 | | None | None | 11 | Right | Nominal |
| 7 | RER | Numeric | 11 | 0 | | None | None | 11 | Right | Nominal |
| 8 | MHRA | Numeric | 11 | 0 | | None | None | 11 | Right | Scale |
| 9 | EIA | Numeric | 11 | 0 | | None | None | 11 | Right | Nominal |
| 10 | Oldpeak | Numeric | 11 | 1 | | None | None | 11 | Right | Scale |
| 11 | STSegment | Numeric | 11 | 0 | ST Segment | None | None | 11 | Right | Nominal |
| 12 | NumberofMV | Numeric | 11 | 0 | Number of MV | None | None | 11 | Right | Scale |
| 13 | Thal | Numeric | 11 | 0 | | None | None | 11 | Right | Scale |
| 14 | Output | Numeric | 11 | 0 | | None | None | 11 | Right | Nominal |

Sumber: Hasil pengolahan data (2014)

Dari 270 orang yang melakukan pemeriksaan jantung tersebut, memiliki variabel dengan *data value* yang berbeda-beda, seperti terlihat pada Tabel 2:

Tabel 2. Data *Value* dari Variabel

| 1 | age | sex | CPT | RBP | SC | FBS | RER | MHRA | EIA | Oldpeak | ST Segment | Number of MV | Thal | Outpu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 70 | 1 | 4 | 130 | 322 | 0 | 2 | 109 | 0 | 2.4 | 2 | 3 | 3 | 2 |
| 3 | 67 | 0 | 3 | 115 | 564 | 0 | 2 | 160 | 0 | 1.6 | 2 | 0 | 7 | 1 |
| 4 | 57 | 1 | 2 | 124 | 261 | 0 | 0 | 141 | 0 | 0.3 | 1 | 0 | 7 | 2 |
| 5 | 64 | 1 | 4 | 128 | 263 | 0 | 0 | 105 | 1 | 0.2 | 2 | 1 | 7 | 1 |
| 6 | 74 | 0 | 2 | 120 | 269 | 0 | 2 | 121 | 1 | 0.2 | 1 | 1 | 3 | 1 |
| 7 | 65 | 1 | 4 | 120 | 177 | 0 | 0 | 140 | 0 | 0.4 | 1 | 0 | 7 | 1 |
| 8 | 56 | 1 | 3 | 130 | 256 | 1 | 2 | 142 | 1 | 0.6 | 2 | 1 | 6 | 2 |
| 9 | 59 | 1 | 4 | 110 | 239 | 0 | 2 | 142 | 1 | 1.2 | 2 | 1 | 7 | 2 |
| 10 | 60 | 1 | 4 | 140 | 293 | 0 | 2 | 170 | 0 | 1.2 | 2 | 2 | 7 | 2 |
| 11 | 63 | 0 | 4 | 150 | 407 | 0 | 2 | 154 | 0 | 4.0 | 2 | 3 | 7 | 2 |
| 12 | 59 | 1 | 4 | 135 | 234 | 0 | 0 | 161 | 0 | 0.5 | 2 | 0 | 7 | 1 |
| 13 | 53 | 1 | 4 | 142 | 226 | 0 | 2 | 111 | 1 | 0.0 | 1 | 0 | 7 | 1 |
| 14 | 44 | 1 | 3 | 140 | 235 | 0 | 2 | 180 | 0 | 0.0 | 1 | 0 | 3 | 1 |

Sumber: Hasil pengolahan data (2014)

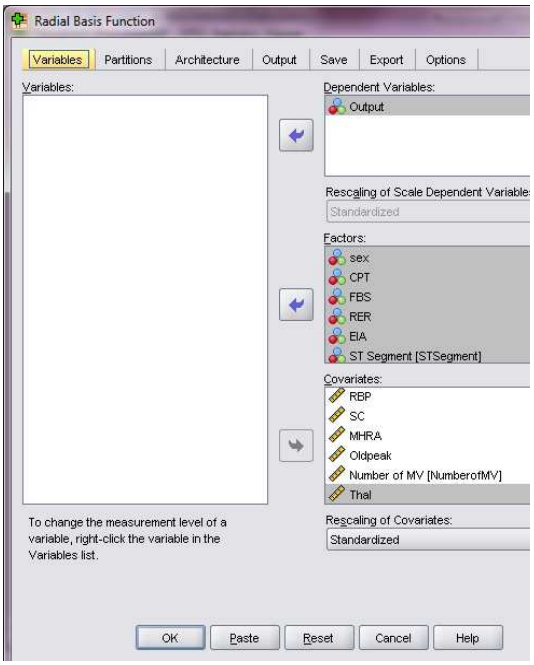
## 8. Pengolahan Data Awal

Untuk mendapatkan hasil pengolahan data yang objektif, data terlebih dahulu akan diacak *(random)*, dimana ditentukan inisialisasi pembangkit aktif (*active generator initialization*), yaitu nilai awal (*starting point*) berupa nilai tetap (*fixed value*) : 9191972.

Sumber: Hasil pengolahan data (2014)

Gambar 3. Random Data

Dari variable yang ada, akan ditentukan mana yang merupakan variable Tujuan dan mana yang merupakan variable *predictor*. Untuk mendapatkan data yang valid, variable *predictor* dipilah kembali menjadi *factor* dan *covariates*, tergantung pada jenis ukuran data, seperti terlihat pada gambar 4:

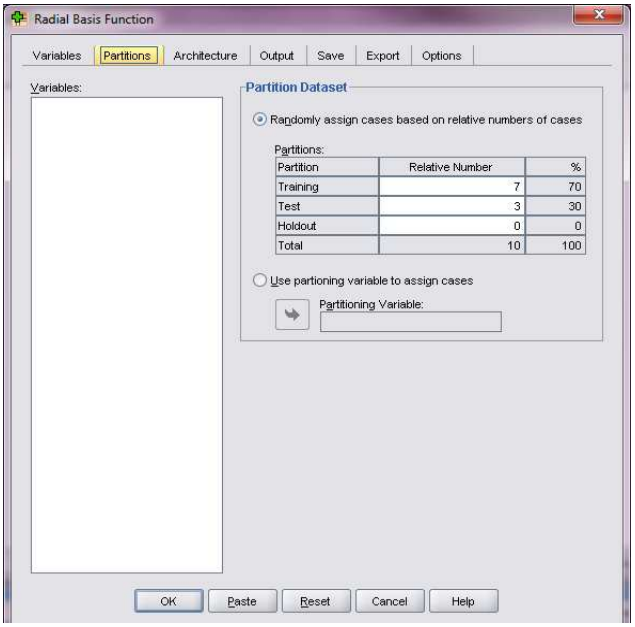



Sumber: Hasil pengolahan data (2014)

Gambar 4. Pengaturan Jenis Variabel

Langkah selanjutnya adalah melakukan partisi data. Langkah ini dilakukan dengan tujuan untuk memisahkan data menjadi dua bagian, yaitu data *Training* dan data *Testing*.

Data *training* adalah sejumlah data yang digunakan untuk merangkai jaringan *neural network*, dimana data tersebut akan digunakan untuk memperoleh model algoritma. Sedangkan data *testing* adalah sejumlah data yang digunakan untuk menguji model yang telah terbentuk. Dalam penelitian ini, data *training* diambil sekitar 70% dari total data dan data *testing* akan menggunakan sekitar 30% data lainnya.




Sumber: Hasil pengolahan data (2014)

Gambar 5. Partisi Data

## HASIL DAN PEMBAHASAN

### Implementasi Algoritma *Radial Basis Function* (RBF)

Data yang telah terbagi menjadi *data training* dan *data testing* kemudian diterapkan ke dalam algoritma RBF, dengan hasil partisi data seperti terlihat pada table 3:

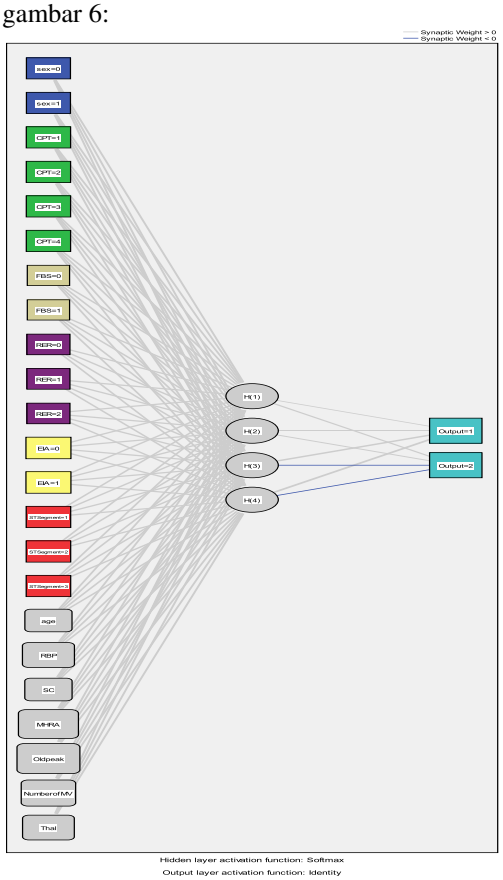
**Tabel 3. Keterangan Proses Data RBF**

| | | N | *Percent* |
|---|---|---|---|
| *Sample* | *Training* | 196 | 72.6% |
| | *Testing* | 74 | 27.4% |
| *Valid* | | 270 | 100.0% |
| *Excluded* | | 0 | |
| Total | | 270 | |

Sumber: Hasil pengolahan data (2014)

Pada table diatas terlihat, dari total 270 data yang diolah, semua data dalam kondisi valid sehingga tidak ada data yang tertolak *(excluded)*. Dari total 270 data tersebut, 196 data (72,6%) digunakan sebagai data *Training* dan 74 data (27,4%) digunakan sebagai data *Testing*.

*Neural Network* yang terbentuk dari pengolahan data tersebut terlihat pada gambar 6:




Sumber: Hasil pengolahan data (2014)

Gambar 6. *Neural Network* RBF yang terbentuk

Pada gambar di atas terlihat bahwa, *neural network* yang terbentuk terdiri dari 23 lapisan masukan *(input layer),* 4 lapisan tersembunyi *(hidden layer)* dan 2 lapisan keluaran *(output layer).*

**Implementasi Algoritma *Multilayer Perceptron* (MLP)**

Data yang telah terbagi menjadi *data training* dan *data testing* kemudian diterapkan ke dalam algoritma MLP, dengan hasil partisi data seperti terlihat pada table 3:

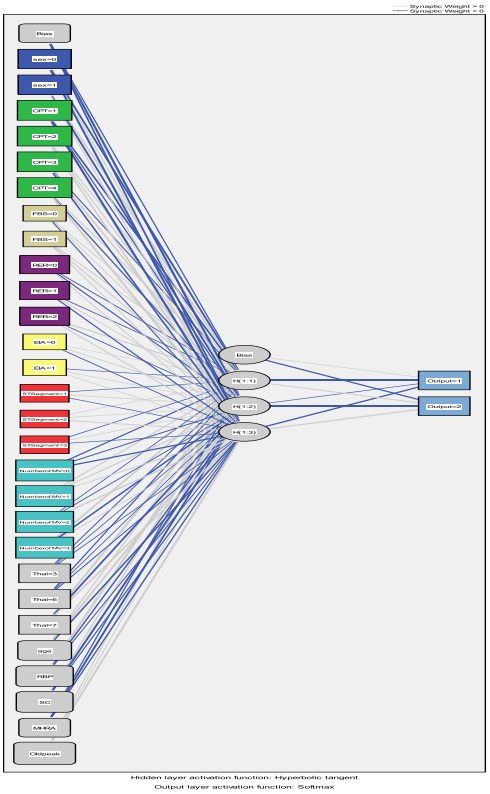
**Tabel 4. Keterangan Proses Data MLP**

| | | N | *Percent* |
|---|---|---|---|
| *Sample* | *Training* | 199 | 73.7% |
| | *Testing* | 71 | 26.3% |
| *Valid* | | 270 | 100.0% |
| *Excluded* | | 0 | |
| Total | | 270 | |

Sumber: Hasil pengolahan data (2014)

Pada table diatas terlihat, dari total 270 data yang diolah, semua data dalam kondisi valid sehingga tidak ada data yang tertolak *(excluded)*. Dari total 270 data tersebut, 199 data (73,7%) digunakan sebagai data *Training* dan 71 data (26,3%) digunakan sebagai data *Testing*.

*Neural Network* yang terbentuk dari pengolahan data tersebut terlihat pada gambar 7:




Sumber: Hasil pengolahan data (2014)

Gambar 7. *Neural Network* MLP yang terbentuk

Pada gambar di atas terlihat bahwa, *neural network* yang terbentuk terdiri dari 29 lapisan masukan *(input layer),* 4 lapisan tersembunyi *(hidden layer)* dan 2 lapisan keluaran *(output layer).*

**Evaluasi Hasil Penerapan**

Data yang telah terbagi menjadi *data training* dan *data testing* dan kemudian telah diterapkan ke dalam algoritma MLP dan RBF akan menghasilkan nilai ketepatan *(correct)* prediksi hasil pengujian model algoritma. Metode yang digunakan untuk mengevaluasi adalah metode *Confusion Matrix. Confusion matrix* digunakan karena akan memberikan keputusan yang diperoleh dalam *training* dan *testing* (Bramer, 2007) dan *Confusion matrix* juga akan memberikan penilaian *performance* klasifikasi berdasarkan objek dengan benar atau salah (Gorunescu, 2011), seperti terlihat pada tabel 5:

Tabel 5. *Confusion Matrix*

| | | *Predicted* **RBF** | | | *Predicted* **MLP** | | |
|---|---|---|---|---|---|---|---|
| *Sample* | *Observed* | 1 | 2 | *Percent Correct* | 1 | 2 | *Percent Correct* |
| *Training* | 1 | 90 | 19 | 82.6% | 102 | 7 | 93.6% |
| | 2 | 13 | 74 | 85.1% | 13 | 77 | 85.6% |
| | *Overall Percent* | 52,6% | 47,4% | 83,7% | 42.2% | 89.9% | 89.9% |
| *Testing* | 1 | 34 | 7 | 82.9% | 39 | 2 | 95.1% |
| | 2 | 7 | 26 | 78.8% | 7 | 23 | 76.7% |
| | *Overall Percent* | 55.4% | 44.6% | **81.1%** | 64,.% | 35.2% | **87,3%** |

Sumber: Hasil pengolahan data (2014)

Berdasarkan table di atas, terlihat bahwa hasil pengujian menggunakan metode *confusion matrix* membuktikan bahwa algoritma MLP memiliki akurasi yang lebih tinggi **(87,35%)** dibandingkan dengan algoritma RBF **(81,1%).**

Hal tersebut dapat terlihat secara rinci pada nilai variable prediksi setiap variable, seperti terlihat pada Tabel 6:

Tabel 6. Hasil Prediksi Data

| | dpeak | STSegment | NumberofMV | Thal | Output | MLP_PredictedValue | MLP_PseudoProbability_1 | MLP_PseudoProbability_2 | RBF_PredictedValue | RBF_PseudoProbability_1 | RBF_PseudoProbability_2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.4 | 2 | 3 | 3 | 2 | 2 | 0.063 | 0.937 | 2 | 0.116 | 0.884 |
| 2 | 1.6 | 2 | 0 | 7 | 1 | 1 | 0.732 | 0.268 | 1 | 0.995 | 0.005 |
| 3 | 0.3 | 1 | 0 | 7 | 2 | 1 | 0.723 | 0.277 | 1 | 0.641 | 0.359 |
| 4 | 0.2 | 2 | 1 | 7 | 1 | 2 | 0.077 | 0.923 | 2 | 0.072 | 0.928 |
| 5 | 0.2 | 1 | 1 | 3 | 1 | 1 | 0.871 | 0.129 | 1 | 0.715 | 0.285 |
| 6 | 0.4 | 1 | 0 | 7 | 1 | 1 | 0.819 | 0.181 | 2 | 0.407 | 0.593 |
| 7 | 0.6 | 2 | 1 | 6 | 2 | 2 | 0.185 | 0.815 | 2 | 0.431 | 0.569 |
| 8 | 1.2 | 2 | 1 | 7 | 2 | 2 | 0.064 | 0.936 | 2 | 0.089 | 0.911 |
| 9 | 1.2 | 2 | 2 | 7 | 2 | 2 | 0.043 | 0.957 | 2 | 0.325 | 0.675 |
| 10 | 4.0 | 2 | 3 | 7 | 2 | 2 | 0.043 | 0.957 | 2 | 0.140 | 0.860 |
| 11 | 0.5 | 2 | 0 | 7 | 1 | 2 | 0.433 | 0.567 | 2 | 0.452 | 0.548 |
| 12 | 0.0 | 1 | 0 | 7 | 1 | 2 | 0.148 | 0.852 | 2 | 0.041 | 0.959 |
| 13 | 0.0 | 1 | 0 | 3 | 1 | 1 | 0.916 | 0.084 | 1 | 0.935 | 0.065 |
| 14 | 2.6 | 2 | 2 | 3 | 2 | 2 | 0.313 | 0.687 | 2 | 0.358 | 0.642 |
| 15 | 0.0 | 1 | 1 | 3 | 1 | 1 | 0.755 | 0.245 | 1 | 0.856 | 0.144 |
| 16 | 1.6 | 2 | 0 | 3 | 1 | 1 | 0.982 | 0.018 | 2 | 0.430 | 0.570 |
| 17 | 1.8 | 2 | 2 | 7 | 2 | 2 | 0.042 | 0.958 | 2 | 0.037 | 0.963 |

Sumber: Hasil pengolahan data (2014)

Pada Tabel 6 diatas terlihat, pada baris 1, 2, 5, 7,8,9, 10, 13, 15 dan 17 hasil prediksi algoritma MLP dan RBF tepat. Pada baris 6 dan 16, hasil prediksi MLP tepat sedangkan hasil prediksi RBF tidak tepat. Sedangkan pada baris 3, 4, 11, dan 12 hasil prediksi MLP dan RBF sama-sama tidak tepat.

Setelah dievaluasi menggunakan metode *confusion matrix*, penerapan algoritma MLP dan RBF juga dievaluasi menggunakan kurva ROC *(Receiver Operating Characteristic)*. Berdasarkan evaluasi yang dilakukan, untuk algoritma MLP dihasilkan nilai AUC (*Area Under the Curve*) **0,949** dan untuk algoritma RBF dihasilkan nilai AUC **0,911**.

**Analisa Komparasi Hasil**

Untuk mengetahui algoritma mana yang memiliki akurasi yang lebih tinggi dalam memprediksi penyakit jantung, digunakan metode *Confusion Matrix* dan Kurva ROC dengan hasil komparasi yang terlihat pada tabel 3:

Tabel 7 : Hasil Komparasi

| | **RBF** | **MLP** |
|---|---|---|
| ***Confusion Matrix*** | **0,811** | **0,873** |
| Kurva ROC (AUC) | **0,911** | **0,949** |

Sumber: Hasil pengolahan data (2014)

Tabel hasil komparasi di atas memperlihatkan bahwa model algoritma MLP memiliki tingkat keakuratan yang lebih tinggi dibandingkan RBF, baik dengan menggunakan metode *confusion matrix* maupun kurva ROC.

Pada *data mining*, nilai AUC dapat dibagi menjadi beberapa kelompok , yaitu (Gorunescu, 2011):

a. 0.90 sampai 1.00 : klasifikasi sangat baik (*excellent classification*)
b. 0.80 sampai 0.90 : klasifikasi baik (*good classification*)
c. 0.70 sampai 0.80 : klasifikasi cukup (*fair clasifcation*)
d. 0.60 sampai 0.70 : klasifikasi buruk (*poor classification*)
e. 0.50 sampai 0.60 : klasifikasi salah (*failure classification*)

Berdasarkan pengelompokkan di atas maka dapat disimpulkan bahwa dengan menggunakan metode *confusion matrix*, algoritma MLP dan RBF termasuk dalam klasifikasi baik (*good classification*) karena memiliki nilai AUC antara 0.80 sampai 0.90, sedangkan dengan menggunakan metode kurva ROC (AUC) algoritma MLP dan RBF termasuk dalam klasifikasi sangat baik (*excellent classification*) karena memiliki nilai AUC antara 0.90 sampai 1.00.

## KESIMPULAN

Tujuan dari penelitian ini adalah untuk mengetahui model *Neural Network* yang memiliki nilai prediksi penyakit Jantung yang lebih akurat. Hasil pengujian algoritma MLP dan RBF menghasilkan kesimpulan bahwa baik menggunakan metode *Confusion Matrix* maupun kurva ROC, algoritma MLP memiliki tingkat akurasi yang lebih tinggi dibandingkan RBF. Menggunakan metode C*onfusion Matrix*, MLP memiliki tingkat akurasi 0,873 sedangkan RBF memiliki tingkat akurasi 0,811. Menggunakan metode Kurva ROC, algoritma MLP menghasilkan nilai AUC 0,949 sedangkan RBF 0,911.

Penelitian ini juga menyimpulkan bahwa dengan menggunakan metode *Confusion Matrix* menghasilkan nilai pengujian yang termasuk dalam klasifikasi Baik (*Good Classification*). Sedangkan dengan menggunakan metode pengujian Kurva ROC hasil pengukuran terhadap algoritma MLP dan RBF menghasilkan nilai AUC yang termasuk dalam klasifikasi Sangat Baik (*Excellent Classification*). Dengan kedua klasifikasi tersebut, maka algoritma MLP dan RBF dapat digunakan untuk prediksi penyakit jantung.

## DAFTAR PUSTAKA


Bramer, M. 2007. *Principle of Data Mining*. London: Springers-Verlag.

Departemen Kesehatan Republik Indonesia. 2009. Profil Kesehatan Indonesia 2008. Jakarta

Gorunescu, F. 2011. "*Data Mining Concepts, Models and Techniques"*. Berlin Heidelberg: Springer Verlag.

Hananta, I. Y., & Muhammad, H. F. 2011. Dietisien Deteksi Dini & Pencegahan 7 Penyakit Penyebab Mati Muda. Yogyakarta: Media Pressindo.

Hannan, Shaikh Abdul, R. R. Manza, R. J. Ramteke. 2010. *Generalized Regression Neural Network and Radial Basis Function for Heart Disease diagnosis*. Maharashtra, India. International Journal of Computer Application (0975-8887). Volume 7- No.13.

Kothari, C.R. 2004. *Research Methology Methods and Techniques.* India: New Age International Limited.

Mahmood, A.M, & Kuppa, M. R. 2010. *Early Detection of Clinical Parameters in Heart Disease by Improving Decision Tree Algorithm.* 2011 Second Vaagdevi International Conference in Informations Technology for real World Problem,24-28.

Mehdi, Neshat, & Yaghobi, Mehdi. 2009. *Designing a Fuzzy Expert System of Diagnosing the Hepatitis B Intensity Rate and Comparing it with Adaptive Neural Network Fuzzy System*. Proceeding of the world congress on engineering and computer science 2009,Vol II, WCECS 2009, ISBN:978-988-18210-2-7. October 20-22, 2009. pp 1-6

Palaniappan, S. & Awang, R. 2008. *Intelligent Heart Disease Prediction System Using Data Mining Technique.* IJCSNS International Journal of Computer Science and Network Security. Vol. 8, August 2008, 343-350.

Shukla, A., Tiwari, R., & Kala R. 2010. *Real Life Applications of Soft Computing*. United States of America: CRC Press Taylor and Francis Group.

Sudoyo AW, Setiyohadi B, Alwi I, Simadibrata M, Setiati S. 2009. Buku Ajar Ilmu Penyakit Dalam Jilid II edisi V. Jakarta: Interna Publishing.


## BIODATA PENULIS


**Ahmad Setiadi, M.Kom**, lahir di Tangerang pada tanggal 1 Mei 1975. Tamat pendidikan Strata satu dari STMIK (Sekarang Universitas) Budi Luhur pada tahun 1988 dengan menyandang gelar Sarjana Komputer (S.Kom) dan tamat pendidikan Strata Dua dari STMIK Nusa Mandiri Jakarta dengan menyandang gelar Magister ilmu Komputer (M.Kom) pada tahun 2012. Sejak tahun 1999 menjadi dosen tetap pada Akademi Bina Sarana Informatika.