

## IMPLEMENTASI METODE *BACKPROPAGATION NEURAL NETWORK* UNTUK MENDIAGNOSIS PENYAKIT KANKER PAYUDARA

Muhamad Bahrul Ulum<sup>1</sup>, Ravie Kurnia Laday<sup>2</sup>

<sup>1</sup> Program Studi Sistem Informasi, Institut Sains dan Teknologi Al-Kamal, Jakarta, Indonesia  
(Tel : +62 857 7447 5605; E-mail: aabahrul@gmail.com)

<sup>2</sup> Program Studi Teknik Informatika, Institut Sains dan Teknologi Al-Kamal, Jakarta, Indonesia  
(Tel : +62 838 9817 4319; E-mail: ravie.4n2sc@gmail.com)

### Abstract

Breast cancer is one of lethal diseases feared by women. Early symptoms that are hard to recognize are the main factors which caused early treatment to patients receive few attention. One of the ways to solve this is by diagnosing someone to know whether or not the individual harbor malignant cancer. The research in this paper will explain and implement backpropagation neural network method to diagnose breast cancer. The implementation process is by conducting tests using breast cancer data set. In initial stage, backpropagation neural network method conducted training to training data set and then the model obtained from the training's network would be tested on test data. The experiment used 569 observation data divided into two parts; which were 70% training data set and 30% test data set for each class. Based on the result of the experiment, backpropagation neural network method can be used to diagnose breast cancer. The accuracy value obtained from this method was 100%. Choosing training data sets that will be used for learning will affect the accuracy level of the system.

*Keyword : Neural network, breast cancer, classification.*

### 1. Pendahuluan

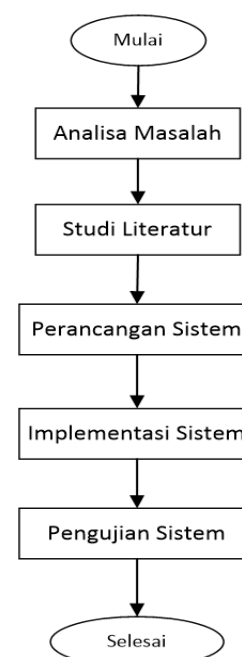
Kanker payudara (*breast cancer*) merupakan salah satu penyakit mematikan yang ditakuti bagi kaum wanita. Gejala awal yang kurang dikenali, merupakan faktor utama yang menyebabkan penanganan dini penderita tidak mendapat perhatian lebih. Sehingga proses pengobatan dilakukan apabila tingkat penyakit ini telah memasuki level yang tinggi. Salah satu cara adalah dengan melakukan diagnosis terhadap seseorang, untuk mengetahui apakah orang tersebut terkena kanker ganas atau tidak. Teknik untuk mendiagnosa penyakit pada manusia dapat menggunakan teknik klasifikasi, salah satunya adalah jaringan syaraf tiruan (JST). Pada proses diagnosis penyakit, pasien akan dikelompokkan menderita penyakit tersebut atau tidak. Untuk itu, diperlukan teknik klasifikasi yang mampu melakukan pembelajaran (*learning*) sehingga model jaringan yang dikembangkan dapat menghasilkan pola dari permasalahan tersebut.

Dalam dunia informatika, khususnya kecerdasan sistem dapat menggunakan teknik klasifikasi untuk mendiagnosa suatu penyakit. Ada beberapa teknik klasifikasi berbasis pembelajaran (*learning*) yang dapat digunakan salah satunya adalah dengan menggunakan jaringan syaraf tiruan *backpropagation*. Banyak peneliti yang menggunakan metode *learning* dengan pendekatan metode lainnya seperti Bayes dan SVM, antara lain algoritma perangkian *game online* [2], penentuan markov model [3]. Penelitian lain yang telah dilakukan yang berkaitan antara lain adalah mereka menggunakan

*linear programming* untuk mendiagnosis dan meramal penyakit kanker payudara [1].

### 2. Metodologi

Pada penelitian ini, metode penelitian yang digunakan memiliki alur sebagai berikut.



Gambar 1. Alur penelitian

**A. Analisa masalah**

Penelitian ini bertujuan untuk mengembangkan sistem yang dapat digunakan untuk mendiagnosa penyakit kanker payudara pada diri seseorang. Pengembangan tersebut menggunakan teknik klasifikasi yang berbasis teknik pembelajaran berdasarkan data fakta yang diberikan, yaitu jaringan syaraf tiruan *backpropagation*. Proses pelatihan dan pengujian dilakukan pada komputer *notebook* dengan spesifikasi processor Pentium® Dual Core CPU T4500 @T230 GHz, memori 2 GB. Sedangkan perangkat lunak yang digunakan menggunakan MATLAB 7.13.0.564 (R2011b).

**B. Studi literatur**

Pada penelitian ini, data yang digunakan adalah data diagnosis kanker payudara yang dibuat oleh Street *et al.*,(1993). Data set tersebut dapat didownload pada “*UCI Machine Learning Repository*” (Wisconsin, 2012).

**C. Perancangan Sistem**

Proses perancangan sistem dengan melakukan pelatihan terhadap data latih. Sehingga data akan dibagi menjadi dua kelompok, yaitu data latih dan data uji. Selanjutnya data latih tersebut akan dilatih untuk memperoleh model jaringan yang tepat.

**D. Implementasi Sistem**

Pada proses ini adalah mengimplementasikannya dengan menggunakan MATLAB 7.13.0.564 (R2011b).

**E. Pengujian Sistem**

Pada proses pengujian ini dilakukan menggunakan data latih.

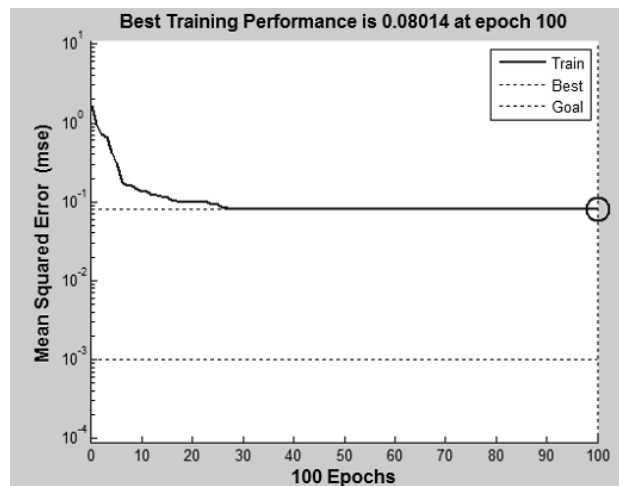
**3. Pembahasan**

Berikut adalah langkah-langkah untuk melakukan klasifikasi.

1. Data pada masing-masing kelas dibagi menjadi 2 bagian, yakni data latih dan data uji. Untuk data latih menggunakan 70% data observasi pada masing-masing kelas (148 *malignant* dan 250 *benign*), sedangkan sisanya sebanyak 30% digunakan sebagai data uji (64 *malignant* dan 107 *benign*).
2. Sebelum data diproses, terlebih dahulu dilakukan proses normalisasi dan praprosesing, agar nilai masing-masing data dapat dihitung dengan nilai yang lebih kecil tanpa kehilangan karakteristiknya.
3. Setelah semua data input dan target dinormalisasi, langkah selanjutnya adalah membangun jaringan syaraf tiruan *backpropagation*.
4. Langkah selanjutnya adalah menentukan parameter-parameter yang mempengaruhi *learning rate* pada proses pelatihan.
5. Setelah input, target, jaringan, dan parameter ditentukan, selanjutnya adalah melakukan proses pelatihan.

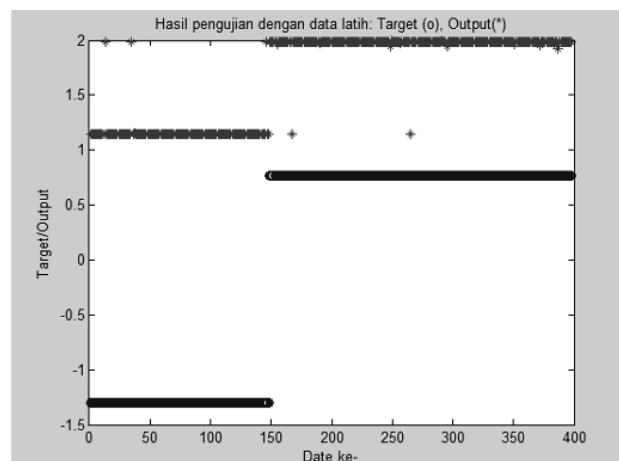
6. Proses pelatihan akan berlangsung hingga mendapatkan nilai *error* minimum atau nilai *epoch* maksimum. Selanjutnya melakukan pengujian (simulasi) terhadap *JST backpropagation* hasil pelatihan. Fungsi *poststd* digunakan untuk men denormalisasi target hasil simulasi.
7. Apabila proses pengujian telah dilakukan, selanjutnya dapat menampilkan hasil data simulasi dalam bentuk grafik.

Berdasarkan hasil pengujian tersebut dapat dianalisis kelas yang dihasilkan (target) apakah sesuai dengan kelas asal (fakta). Selanjutnya hasil analisis tersebut dapat digunakan untuk mengevaluasi model klasifikasi serta total akurasi. Gambar 2 di bawah ini merupakan grafik hasil performa dari jaringan yang sudah dilatih.



**Gambar 2.** Grafik pelatihan

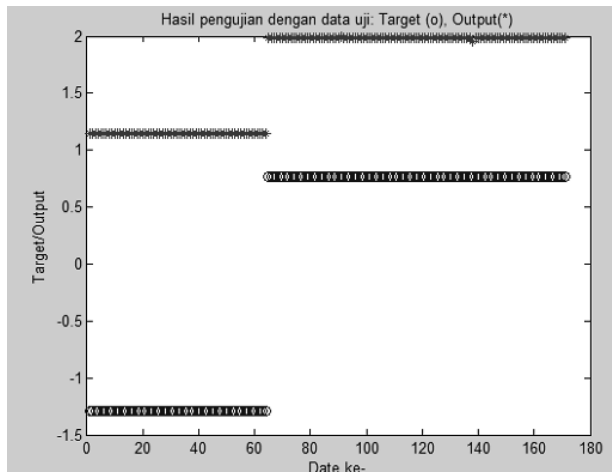
Sedangkan untuk hasil pengujian yang dilakukan dengan menggunakan data latih dapat dilihat pada Gambar 3 di bawah ini.



**Gambar 3.** Hasil pengujian dengan data latih

Berdasarkan gambar di atas, terdapat data pada kelas *benign* yang salah klasifikasi kedalam kelas *malignant* sebanyak 2 data, demikian halnya dengan data *malignant* yang salah klasifikasi kedalam

kelas *benign*. Untuk pengujian terhadap data uji dapat dilihat pada Gambar 4 di bawah ini.



**Gambar 4.** Hasil pengujian dengan data uji

Evaluasi kinerja model klasifikasi Backpropagation Neural Network dilakukan dengan menghitung persentase ketepatan suatu data yang masuk ke kelas tertentu. Evaluasi tersebut direpresentasikan dalam bentuk *confusion matrix*. Untuk menghitung tingkat akurasi dari seluruh percobaan, dengan menggunakan persamaan berikut:

$$\text{Akurasi} = \frac{\text{Jumlah klasifikasi benar}}{\text{Total data klasifikasi}} \times 100\%$$

Hasil perhitungan tingkat akurasi secara keseluruhan menggunakan data uji direpresentasikan pada Tabel 1 di bawah ini.

**Tabel 1.** Tingkat Akurasi

Kelas asal	Kelas Target		Total Data	Akurasi
	Malignant	Benign		
Malignant	64	0	64	100%
Benign	0	107	107	100%
Total Akurasi				100%

#### 4. Kesimpulan

Berdasarkan hasil percobaan, dengan menggunakan 569 data observasi yang dibagi menjadi dua bagian, yaitu data latih sebanyak 70% untuk masing-masing kelas dan data uji sebanyak 30% pada masing-masing kelas. Metode jaringan syaraf tiruan *backpropagation* dapat digunakan untuk mendiagnosis penyakit kanker payudara. Nilai akurasi yang dihasilkan dengan menggunakan metode ini adalah sebesar 100%. Pemilihan data latih yang akan digunakan sebagai pembelajaran akan mempengaruhi tingkat akurasi dari sistem tersebut.

#### Daftar Pustaka

- [1] Mangasarian OL, Street WN, Wolberg WH. Breast cancer diagnosis and prognosis via linear programming. *Operations Research*. 43(4):570-577. 1995.
- [2] Ruby CW, Chih JL. A Bayesian Approximation Method for Online Ranking. *Journal of Machine Learning Research*. 12:267-300. 2011.
- [3] Stéphane R, Joelle P, Brahim Cd, Pierre K. A Bayesian Approach for Learning and Planning in Partially Observable Markov Decision Processes. *Journal of Machine Learning Research*. 12:1729-1770. 2011.
- [4] UCI Data Set. 2015. *Breast Cancer Wisconsin (Diagnostic) Data Set*. <http://archive.ics.uci.edu/ml/datasets/> [diakses maret 2015]
- [5] Wisconsin. 2015. Breast Cancer Wisconsin (Diagnostic) Data Set. <http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29>. (Diakses maret 2015).