

Diagnosis Kanker Payudara Menggunakan Jaringan Saraf Tiruan Dengan Metode Backpropagation

Diana Astria Gultom¹, Wesley Yando Tantra²

¹Fakultas Ilmu Komputer, Universitas Pelita Harapan

E-mail: diana.gultom@gmail.com

²Fakultas Ilmu Komputer, Universitas Pelita Harapan

E-mail: [wy70008@student.uph.edu](mailto:w70008@student.uph.edu)

Abstract – Based on the World Health Organization (WHO), breast cancer ranks eighth which causes the largest mortality rate in the world. Based on data from the Ministry of Health of the Republic of Indonesia, breast cancer is ranked second where the first position is cervical cancer. According to data from WHO, in 2015 the mortality rate due to breast cancer in the world reached 571,000 or 6.48% of the total mortality in the world. While in Indonesia, the figure is 20,025 or 1.41% of the total number of deaths in Indonesia. Increased rates of breast cancer can be caused by several risk factors such as genetic and family history, previous tumor or breast cancer history, history of early menstruation, history of late menopause, obesity, reproductive history, hormonal, poor diet, alcohol consumption, due to radiation ultraviolet light, and environmental factors. In Indonesia, more than 80% of cases were found difficult to make treatment efforts because the cases are at an advanced stage. Early diagnosis of breast cancer can be done with the Data Mining process with the Artificial Neural Network method and the Backpropagation algorithm. In conducting this research researchers used data available at the UCI - Machine Learning Repository: Breast Cancer Repository. The results showed that the level of accuracy in diagnosing breast cancer using artificial neural networks with the Backpropagation method in the testing process reached 94.634% and when the training process was 99.372%.

Keywords: breast cancer, backpropagation, diagnosis, genetic

Abstrak – Berdasarkan World Health Organization (WHO), kanker payudara menempati urutan ke-delapan yang menyebabkan angka mortalitas terbesar di dunia. Berdasarkan data Kementerian Kesehatan Republik Indonesia, kanker payudara berada pada urutan kedua dimana posisi pertama merupakan kanker leher rahim. Menurut data dari WHO, pada tahun 2015 angka mortalitas akibat dari kanker payudara di dunia mencapai 571.000 atau sebesar 6,48% dari keseluruhan mortalitas di dunia. Sedangkan di Indonesia berada pada angka 20.025 atau senilai 1,41% dari keseluruhan jumlah kematian yang ada di Indonesia. Meningkatnya angka dari kanker payudara ini dapat disebabkan dari beberapa faktor resiko seperti genetik dan riwayat keluarga, riwayat tumor atau kanker payudara sebelumnya, riwayat menstruasi dini, riwayat menopause lambat, obesitas, riwayat reproduksi, hormonal, pola makan yang buruk, konsumsi alkohol, akibat radiasi sinar ultraviolet, dan faktor lingkungan. Di Indonesia, lebih dari

80% kasus ditemukan sulit melakukan upaya pengobatan karena kasus berada pada stadium yang lanjut. Diagnosis dini kanker payudara dapat dilakukan dengan proses Data Mining dengan metode Jaringan Saraf Tiruan dan algoritma Backpropagation. Dalam melakukan penelitian ini peneliti menggunakan data yang ada pada UCI - Machine Learning Repository: Breast Cancer Repository. Hasil penelitian menunjukkan bahwa tingkat keakuratan mendiagnosis kanker payudara menggunakan jaringan saraf tiruan dengan metode Backpropagation pada proses testing mencapai 94,634% dan saat proses training sebesar 99,372%.

Kata Kunci: kanker payudara, backpropagation, diagnosis, genetik

PENDAHULUAN

Kanker merupakan salah satu penyakit yang paling berbahaya karena proses penyembuhan yang terbilang sulit dan efek yang ditimbulkan setelahnya. Banyak yang masih mengira bahwa kanker merupakan hal yang sama dengan tumor, pada kenyataannya tumor yang muncul tidak dapat disebut sebagai kanker. Sel kanker dapat muncul jika terjadi kesalahan pada siklus sel dalam tubuh seperti pembelahan diri sel yang terjadi secara tidak sempurna.

Wanita perlu lebih memperhatikan kesehatan payudara terutama ketika munculnya benjolan yang dapat tergolong kanker atau bukan. Benjolan pada payudara yang sering terjadi disebut kista payudara atau fibroadenoma. Fibroadenoma merupakan salah satu jenis tumor jinak dalam payudara yang berbentuk benjolan kecil dan tidak terasa sakit.

Kanker payudara adalah penyakit tumor yang paling berbahaya dapat menyerang pada bagian payudara dimana jaringan penunjang payudara, kelenjar dan saluran kelenjar. Terjadinya kanker payudara disebabkan dengan adanya gen yang rusak dimana gen tersebut berfungsi untuk mengatur pemisahan dan pertumbuhan. Kanker payudara dapat terjadi pada kaum wanita dan kaum pria, akan tetapi kemungkinan

terbesar dapat terjadi pada kaum wanita dibandingkan pria.

Berdasarkan World Health Organization (WHO), kanker payudara menempati urutan ke-delapan yang menyebabkan angka mortalitas terbesar di dunia. Berdasarkan data Kementerian Kesehatan Republik Indonesia, kanker payudara berada pada urutan kedua dimana posisi pertama merupakan kanker leher rahim. Menurut data dari WHO, pada tahun 2015 angka mortalitas akibat dari kanker payudara di dunia mencapai 571.000 atau sebesar 6,48% dari keseluruhan mortalitas di dunia. Sedangkan di Indonesia berada pada angka 20.025 atau senilai 1,41% dari keseluruhan jumlah kematian yang ada di Indonesia. Meningkatnya angka dari kanker payudara ini dapat disebabkan dari beberapa faktor resiko seperti genetik dan riwayat keluarga, riwayat tumor atau kanker payudara sebelumnya, riwayat menstruasi dini (dibawah usia 12 tahun), riwayat menopause lambat (diatas usia 55 tahun), obesitas, riwayat reproduksi (tidak mempunyai anak dan tidak pernah menyusui), hormonal, pola makan yang buruk, konsumsi alkohol, akibat radiasi sinar ultraviolet, dan faktor lingkungan. Di Indonesia, lebih dari 80% kasus ditemukan sulit melakukan upaya pengobatan karena kasus berada pada stadium yang lanjut. Kanker payudara pada umumnya dapat di deteksi secara

awal dengan cara yaitu SADARI dimana pengertiannya adalah Periksa Payudara Sendiri, SADANIS yang artinya Periksa Payudara Klinis dan mammografi skrining (Kementrian Kesehatan Republik Indonesia, 2017).

Diagnosis dini kanker payudara dapat dilakukan dengan proses Data Mining dengan metode Jaringan Saraf Tiruan dan algoritma Backpropagation. Algoritma backpropagation biasa digunakan untuk meramal kemungkinan yang dapat terjadi. Jaringan Saraf Tiruan Backpropagation mampu memberikan suatu gambaran seperti obyek secara menyeluruh hanya dengan sebagian dari obyek lainnya. Dalam melakukan penelitian ini peneliti menggunakan data yang ada pada UCI - Machine Learning Repository: Breast Cancer Repository.

Kanker Payudara

Kanker ini merupakan pertumbuhan sel-sel yang ada dalam payudara bertumbuh secara tidak normal. Kanker payudara pada umumnya berupa munculnya sel-sel abnormal yang menyebar luas secara. Umumnya adanya kanker payudara dapat ditandai dengan beberapa ciri-ciri seperti adanya gumpalan kecil yang terasa sakit pada bagian payudara, terdapat cekungan pada permukaan payudara, terjadi pembengkakan, terkadang terasa gatal di bagian puting, dan ada merah-merah disekitar payudara. Saat penyebaran sel kanker, dapat timbul seperti gejala nyeri pada tulang, berat badan menurun secara drastis, dan penyakit kuning

Pada umumnya, sel kanker payudara dapat tumbuh di berbagai area yang berbeda pada payudara. Kebanyakan sel kanker payudara mulai tumbuh di saluran yang membawa ASI ke bagian puting dinamakan ductal cancer. Ada juga yang mulai tumbuh pada kelenjar susu yang

dinamakan lobular cancer, dan ada beberapa jenis kanker payudara lainnya. Ada juga yang disebut sebagai sarkoma dan limfoma dimana sel kanker dimulai pada jaringan lainnya yang terdapat pada payudara. Kedua jenis kanker ini tidak dianggap sebagai kanker payudara. Tidak semua jenis kanker payudara dapat menyebabkan benjolan pada payudara. Banyak kanker payudara yang ditemukan ketika dilakukan skrining mamografi dimana dapat mendeteksi kanker payudara stadium awal sebelum menimbulkan gejala-gejala yang dapat dirasakan.

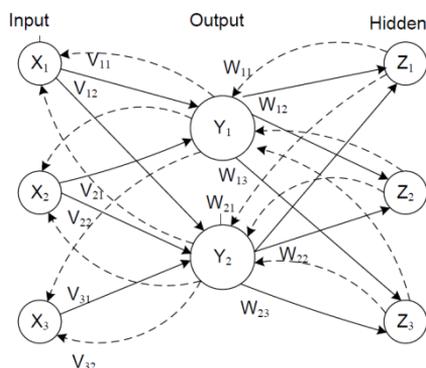
Backpropagation

Salah satu metode yang ada pada jaringan saraf tiruan adalah backpropagation. Metode backpropagation merupakan suatu metode yang baik dalam menangani pengenalan pola yang rumit. Backpropagation mengajarkan jaringan agar memiliki keseimbangan antara kemampuan jaringan untuk memahami atau mengenali hubungan yang digunakan serta kemampuan jaringan memberikan jawaban yang benar terhadap pola input yang serupa dengan pola yang dipakai selama pelatihan (Taufikurrahman, 2014). Backpropagation adalah algoritma pengkajian yang dapat ditinjau dan umumnya digunakan untuk pengklasifikasian secara linear dengan banyak layer untuk mengonversikan bobot yang terhubung dengan setiap neuron pada hidden layer (Kusumadewi, 2004).

Setiap unit yang terdapat di input layer akan terhubung dengan unit yang berbeda, dan setiap unit yang ada pada hidden layer akan terhubung dengan setiap unit yang terdapat pada layer keluaran.

Jaringan saraf tiruan metode backpropagation terdiri dari Jaringan Lapisan Banyak (Puspitaningrum, 2006), yaitu :

1. Lapisan masukan (input layer). Pada input layer terdiri dari unit input, mulai dari unit input pertama hingga unit input n.
2. Lapisan tersembunyi (hidden layer). Pada hidden layer terdiri dari unit hidden, mulai dari unit hidden pertama hingga unit hidden p.
3. Lapisan keluaran (output layer). Pada output layer terdiri dari unit output, mulai dari unit output pertama hingga unit output m.



Gambar 1. Arsitektur jaringan Backpropagation

Berdasarkan gambar diatas, dapat dilihat cara kerja jaringan backpropagation yaitu awalnya jaringan diinisialisasi dengan bobot yang telah ditentukan dengan bilang acak. Setelah itu, sampel akan dimasukkan ke dalam jaringan. Sampel terdiri atas pasangan vektor masukan dan vektor target keluaran. Untuk mengetahui apakah hasil jaringan keluaran telah sesuai dengan yang diharapkan, maka hasil dari jaringan atau kemungkinan yang dihasilkan akan dibandingkan dengan target keluaran.

Karena adanya perbedaan antara prediksi keluaran dengan target keluaran, maka akan dihasilkan error

yang kemudian akan dihitung untuk memperbarui bobot-bobot koneksi yang terkait dan kemudian mempropagasikan error kembali. Proses akan terus berlanjut hingga mencapai hasil yang diharapkan atau sampai kondisi berhenti dipenuhi. Umumnya kondisi berhenti yang sering digunakan adalah jumlah iterasi. Jika jumlah iterasi yang dilakukan jaringan telah melebihi jumlah iterasi yang ditentukan, atau nilai error yang didapat lebih kecil dari batas toleransi, maka iterasi akan terhenti.

METODE PENELITIAN

Langkah-langkah dalam pembuatan model diagnosis kanker payudara dengan menggunakan metode jaringan saraf tiruan backpropagation adalah melakukan pengumpulan data, implementasi metode jaringan saraf tiruan backpropagation untuk melakukan pengklasifikasian data, training dan testing keakuratan, serta analisis hasil prediksi backpropagation.

Penelitian diawali dengan menentukan latar belakang dan tujuan penelitian kemudian merumuskan ruang lingkup. Dilakukan studi literatur untuk memperluas wawasan mengenai cara kerja metode backpropagation dan tahapan-tahapan yang diperlukan untuk mendiagnosis kanker payudara dengan menggunakan metode backpropagation. Kemudian dilanjutkan dengan melakukan pengumpulan data dan dilakukan implementasi metode backpropagation melalui pembagian data, training dan testing backpropagation. Selanjutnya menganalisis hasil prediksi backpropagation dan tahapan terakhir adalah menarik kesimpulan dan memberi saran.

Metode analisis *backpropagation* untuk mendiagnosis penyakit kanker payudara akan dibantu dengan menggunakan aplikasi MATLAB.

1. Pembagian data *training* dan *testing*
 Proses *training* dan *testing* perlu dilakukan untuk proses diagnosis dengan menggunakan *backpropagation*. Proses *training* digunakan untuk menguraikan proses *backpropagation* sedangkan proses *testing* digunakan untuk evaluasi kemampuan *backpropagation* sebagai alat diagnosis. Pembagian data *training* dan *testing* dibagi dengan komposisi sebagai berikut :
 - a. Data *training* diambil sebanyak 70% dari total data
 - b. Data *testing* diambil sebanyak 30% dari total data

2. *Training backpropagation*
 Akan dilakukan *training backpropagation* dengan menggunakan aplikasi MATLAB. Algoritma dalam *training backpropagation* terdapat tiga fase (Kusumadewi, 2004) :

- Langkah 1
 Lakukan inialisasi semua bobot dengan bilangan acak kecil
- Langkah 2
 Jika kondisi penghentian belum terpenuhi, maka lakukan langkah 3 hingga 9
- Langkah 3
 Untuk setiap langkah *training data*, lakukan langkah 4 hingga 9

Fase I : *feed forward*

- Langkah 4
 Tiap unit masukan menerima sinyal dan akan meneruskannya ke unit tersembunyi
- Langkah 5
 Hitung semua keluaran di unit tersembunyi z_j ($j = 1, 2, \dots, p$)

$$z_{net_j} = v_{j0} + \sum_{i=1}^n x_i v_{ji}$$

$$z_j = f(z_{net_j}) = \frac{1}{1 + e^{-z_{net_j}}}$$

- Langkah 6

Hitung semua keluaran jaringan di unit y_k ($k = 1, 2, \dots, m$)

$$y_{net_k} = w_{k0} + \sum_{j=1}^p z_j w_{kj}$$

$$y_k = f(y_{net_k}) = \frac{1}{1 + e^{-y_{net_k}}}$$

Fase II : *backpropagation*

- Langkah 7
 Hitung faktor δ unit keluaran berdasarkan kesalahan di setiap unit keluaran y_k ($k = 1, 2, \dots, m$)
 $\delta_k = (t_k - y_k) f'(y_{net_k}) = (t_k - y_k) y_k (1 - y_k)$
 t_k = target keluaran
 δ_k = merupakan unit kesalahan yang akan dipakai dalam perubahan bobot lapisan dibawahnya.

Hitung perubahan bobot w_{kj} dengan laju pemahaman α

$$\Delta w_{kj} = \alpha \delta_k z_j, \quad (k = 1, 2, \dots, m; j = 1, 2, \dots, p)$$

- Langkah 8
 Hitung faktor δ unit tersembunyi berdasarkan kesalahan di setiap unit tersembunyi z_j ($j = 1, 2, \dots, p$)

$$\delta_{net_j} = \sum_{k=1}^m \delta_k w_{kj}$$

Faktor δ unit tersembunyi

$$\delta_j = \delta_{net_j} f'(z_{net_j}) =$$

$$\delta_{net_j} z_j (1 - z_j)$$

Hitung perubahan bobot v_{ji}

$$\Delta v_{ji} = \alpha \delta_j x_i, \quad (j = 1, 2, \dots, p; i = 0, 1, \dots, n)$$

Fase III : *weight update*

- Langkah 9
 Hitung semua perubahan bobot yang terjadi. Perubahan bobot garis yang menuju ke unit keluaran
 $w_{kj}(\text{baru}) = w_{kj}(\text{lama}) + \Delta w_{kj}$, ($k = 1, 2, \dots, m; j = 0, 1, 2, \dots, p$)

Perubahan bobot garis yang menuju ke unit tersembunyi
 $v_{ji}(\text{baru}) = v_{ji}(\text{lama}) + \Delta v_{ji}$, ($j = 1, 2, \dots, p; i = 0, 1, 2, \dots, n$)

Keterangan :

α : laju pembelajaran

n : jumlah data pembelajaran

δ_{net_j} : jumlah *input* pada *hidden layer* dari unit pada *output layer* y_k

y_{net_k} : sinyal masukan untuk unit *output* y_k

y_k : unit *output* k

x_i : unit *input* i

v_{jo} : bias pada unit *hidden*

w_{ko} : bias pada unit *output*

δ_k : informasi *error* pada unit *output* y_k yang dilakukan *backward pass* kembali ke unit *hidden*

δ_j : informasi *error* pada unit *hidden* z_j

Δv_{ji} : koreksi bobot antara *hidden layer* z_j dengan *input layer* x_i

Δw_{kj} : koreksi bobot antara *output layer* y_k dengan *hidden layer* z_j

v_{ji} : bobot antara *hidden layer* z_j dengan *input layer* x_i yang telah disesuaikan

z_j : unit *hidden* j

w_{kj} : bobot antara *output layer* y_k dengan *input layer* z_j yang telah disesuaikan

z_{net_j} : sinyal *input* untuk unit *hidden* z_j

3. Testing backpropagation

Dari proses *training* diperoleh data *input* yang telah di normalisasi dan menentukan konfigurasi jaringan, kemudian tahap selanjutnya adalah melakukan diagnosis dengan metode *backpropagation*.

HASIL DAN PEMBAHASAN

Pada tahap dari penelitian ini adalah mengukur performa *Backpropagation*

yang digunakan untuk melakukan prediksi pada penelitian ini.

1. Jumlah Hidden Neuron

Pada penelitian ini juga dicari berapakah jumlah *hidden neuron* yang optimal untuk mendapatkan hasil yang diharapkan. Jumlah *hidden neuron* akan berpengaruh pada hasil akurasi prediksi saat data *training* ataupun data *testing*. Dengan jumlah *neuron* antara 0 hingga 10, diperoleh nilai akurasi terbaik untuk data *training* adalah 99.372 pada jumlah *hidden neuron* ke-10

Tabel 1. Akurasi Prediksi *Testing* Pada Jumlah *Hidden Neuron*

Jumlah <i>Hidden Neuron</i>	Akurasi Prediksi <i>Testing</i>
0	65.063
1	97.699
2	97.699
3	98.536
4	98.536
5	98.745
6	98.745
7	98.745
8	98.745
9	98.536
10	99.372

Tabel 1 menunjukkan secara lengkap hasil percobaan pada akurasi prediksi *Backpropagation* terhadap jumlah *hidden neuron*. Jumlah *hidden neuron* yang diambil untuk mendapatkan hasil prediksi yang optimal pada penelitian ini hanya saat proses *training* karena keakuratan *Backpropagation* dalam melakukan prediksi dilihat saat *training*.

2. Keakuratan Prediksi dengan Confusion Matrix

Proses *training* dan proses *testing* akan dievaluasi dengan menggunakan *Confusion Matrix*.

Pada penelitian ini, keakuratan *Backpropagation* dalam mendiagnosis penyakit kanker payudara dapat diketahui berdasarkan penanda pada seluruh data pada proses *training* dan proses *testing*. Dari hasil proses *Backpropagation*, didapat tabel *Confusion Matrix* pada setiap kondisi masing-masing proses *training* dan proses *testing* seperti pada tabel 2.

Tabel 2. *Confusion Matrix* Hasil Proses *Backpropagation*

Proses	TP	TN	FP	FN
<i>Training Backpropagation</i>	166	309	1	2
<i>Testing Backpropagation</i>	63	131	9	2

Berdasarkan hasil yang terdapat pada tabel *Confusion Matrix* di atas, selanjutnya akan dilakukan perhitungan *classification accuracy*, *sensitivity* dan *specificity* untuk proses *training Backpropagation* dan *testing Backpropagation*.

Tabel 3. Kinerja dari Setiap Proses

Proses	<i>Classification Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
<i>Training Backpropagation</i>	0.9937	0.988	0.996
<i>Testing Backpropagation</i>	0.9463	0.969	0.935

Pada tabel 3 di atas terlihat bahwa keabsahan dari setiap proses. Dari tabel di atas terlihat bahwa nilai akurasi untuk proses *training Backpropagation* adalah 0.99372 dan nilai akurasi untuk proses *testing Backpropagation* adalah 0.94634.

KESIMPULAN

Berdasarkan hasil penelitian diagnosis kanker payudara menggunakan Jaringan Saraf Tiruan dengan metode *Backpropagation*, maka dapat diambil kesimpulan yaitu tingkat keakuratan mendiagnosis kanker payudara menggunakan jaringan saraf tiruan dengan metode *Backpropagation* pada proses testing mencapai 94,634% dan saat proses training sebesar 99,372%.

DAFTAR PUSTAKA

- [1] American Cancer Society, Diambil kembali dari American Cancer Society Web site: <https://www.cancer.org/content/cancer/en/cancer/acs-medical-content-and-news-staff.html>, 2017.
- [2] Bahrul Ulum, M., & Kurnia Laday, R, Implementasi Metode *Backpropagation Neural Network* Untuk Mendiagnosis Penyakit Kanker Payudara. Seminar Nasional APTIKOM (SEMNASSTIKOM), 2016.
- [3] Hermawan, A, Jaringan Saraf Tiruan. Yogyakarta: Andi, 2006.
- [4] Hoffer, J. A., Ramesh, V., & Topi, H, Modern Database Management. Prentice Hall, Inc, 2012.
- [5] HSB, R. R, Mendeteksi Penyakit Gigi Menggunakan Jaringan Saraf Tiruan Dengan Menggunakan Metode *Backpropagation* dan Metode *Hopefield*, 2014.
- [6] Jiawei Han, M. K, Data Mining : Concepts and Techniques - 3rd edition. Morgan Kaufmann Publishers, 2012.
- [7] Kementerian Kesehatan Republik Indonesia, Kementerian Kesehatan. Diambil kembali dari Kementerian Kesehatan Republik Indonesia, 2017.
- [8] Kusriani, & Taufiq Luthfi, E, Algoritma Data Mining. Yogyakarta: ANDI, 2009.

- [9] Kusumadewi, S, Membangun Jaringan Syaraf Tiruan Menggunakan Matlab & Axcellink. Yogyakarta: Graham Ilmu, 2004.
- [10] Larose, D. T, Discovering Knowledge in Data : An Introduction To Data Mining. Jhon Wiley & Sons Inc, 2005.
- [11] MacLennan, J., Tang, Z., & Crivat, B, Data Mining With Microsoft SQL Server 2008. USA: Wiley Publishing, Inc, 2009.
- [12] Nursela, D. A, Penerapan Algoritma C4.5 Untuk Klasifikasi Tingkat Keganasan Kanker Payudara, 2014.
- [13] Puspitaningrum, D, Pengantar Jaringan Saraf Tiruan. Yogyakarta: Andi, 2006.
- [14] Sathya, R., & Abraham, A, Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification, 2013.
- [15] Setiawan, A., & Sidik Sasongko, P, Sistem Deteksi Dini Penyakit Tuberculosis Menggunakan Jaringan Saraf Tiruan Backpropagation Dengan Algoritma Nguyen Widrow Dan Optimasi Adaptive Learning Rate, 2016.
- [16] Siang, J. J, Jaringan Syaraf Tiruan dan Pemrogramannya Menggunakan MATLAB. Yogyakarta: Andi, 2005.
- [17] Suwarno, & Abdillah, A, Penerapan Algoritma Bayesian Regularization Backpropagation Untuk Memprediksi Penyakit Diabetes, 2016.
- [18] Taufikurrahman, M, Peramalan Nilai Tukar Mata Uang Rupiah Terhadap Dolar Amerika Serikat Tahun 2014 Dengan Menggunakan Jaringan Syaraf Tiruan Model Backpropagation, 2014.
- [19] Vita Via, Y., Nugroho, B., & Syafrizal, A, Sistem Pendukung Keputusan Klasifikasi Tingkat Keganasan Kanker Payudara Dengan Metode Naive Bayes Classifier, 2015.