Klasifikasi Data Microarray dengan Metode *Artificial Neural Network* dan *Genetic Algorithm* untuk Kasus Deteksi Kanker

Ilham Yunirakhman[1], Prof. Dr. Adiwijaya, S.Si., M.Si.[2], Widi Astuti, S.T, M.Kom.[3]

[1,2,3]Fakultas Informatika, Universitas Telkom, Bandung
[1]ilhamrakhman@students.telkomuniversity.ac.id, [2]adiwijaya@telkomuniversity.ac.id,
[3]widiwdu@telkomuniversity.ac.id

Abstrak

Kanker adalah salah satu penyakit yang paling mematikan di dunia. Pada tahun 2012, terdapat 32,6 juta orang yang positif mengidap kanker dan 8,2 juta kematian yang disebabkan oleh kanker. Terdapat banyak cara yang bisa dilakukan untuk mendeteksi kanker sejak dini, salah satu caranya adalah dengan melakukan klasifikasi fitur pada data DNA microarray. Salah satu metode yang digunakan untuk mendeteksi kanker adalah metode *Artificial Neural Network* (ANN) – *Backpropagation* dengan bantuan *Genetics Algorithm* (GA). ANN digunakan sebagai metode klasifikasi untuk memprediksi kanker, sedangkan GA digunakan sebagai metode untuk mereduksi dimensi dari fitur DNA Microarray yang memiliki dimensi yang sangat besar. Pada penelitian ini dilakukan perbandingan antara metode ANN dan metode ANN-GA *hybrid*. Metode ANN-GA terbukti lebih efektif dari ANN karena dapat menghasilkan nilai akurasi 93.08% dan mereduksi dimensi hingga 51% dengan waktu *running time* lebih cepat hingga 42.2%.

Kata Kunci : *Artificial Neural Network* (ANN), ANN-GA *hybrid*, DNA Microarray, *Genetics Algorithm* (GA)

Abstract

Cancer is one of deadliest disease in the world. In 2012, there were 32,6 million people diagnosed positive for cancer and 8,2 million of death caused of cancer. There are many ways that we can do to detect cancer since early stage, one of it is using *Artificial Neural Network* (ANN) – *Backpropagation* methods with *Genetics Algorithm* (GA). ANN is used as classifier methods for predicting cancer and GA is used as dimension reduction method for DNA microarray features which has wide dimension. In this research ANN methods will be compared with ANN-GA hybrid. The result from ANN-GA methods has proven more effective than ANN because it can give 93.08% accuration and reduce dimension up to 51% with faster running time up to 42.2%.

Key words : *Artificial Neural Network* (ANN), ANN-GA *hybrid*, DNA Microarray, *Genetics Algorithm* (GA)

1 Pendahuluan

Latar Belakang

Kanker adalah salah satu penyakit yang paling mematikan di dunia. Pada tahun 2012, terdapat 14.1 juta kasus baru kanker dari sekitar 32.6 juta orang yang positif mengidap kanker dan terdapat 8.2 juta kematian yang disebabkan oleh kanker[2]. Menurut Kementrian Kesehatan Republik Indonesia ada beberapa faktor yang dapat menyebabkan risiko Penyakit kanker. Diantaranya adalah : (1) Faktor genetik, (2) Faktor karsinogen, yang diantaranya merupakan zat kimia, radiasi, virus, hormon, dan iritasi kronis, (3) Faktor perilaku atau gaya hidup, diantaranya yaitu merokok, pola makan yang tidak sehat, konsumsi alkohol, dan kurangnya aktifitas fisik [1].
Ada banyak cara yang bisa dilakukan untuk deteksi kanker sejak dini. Salah satu caranya adalah melalui faktor genetik dengan merepresentasikan data kanker pada DNA microarray dimana ribuan untaian molekul dan dinamai dengan fitur. Data dengan jumlah fitur yang sangat banyak memberikan beban komputasi yang sangat besar dan *running time* time lama, sehingga metode reduksi dimensi dibutuhkan untuk mengurangi beban komputasi dan waktu *running time*.

Beberapa penelitian mengenai klasifikasi data microarray dengan metode ANN telah dilakukan, salah satunya adalah klasifikasi kanker payudara menggunakan metode *Artificial Neural Network* dengan menggunakan struktur *Backpropagation* dan *Genetics Algorithm* untuk optimalisasi parameter pada data kanker payudara dan menghasilkan rata-rata nilai akurasi sebesar 97%[3]. Melihat hasil akurasi dan kemampuan metode tersebut pada kasus kanker payudara maka metode tersebut dijadikan sebagai acuan untuk penelitian ini dengan merubah beberapa struktur pada penelitian tersebut [3] GA digunakan sebagai metode untuk mencari parameter *learning rate* dan *hidden neuron.* Pada penelitian ini, GA digunakan sebagai metode untuk mereduksi dimensi dengan tetap mempertahankan fitur-fitur yang penting, sedangkan *Artificial Neural Nework* (ANN) – *Backpropagation* digunakan sebagai metode classifier.

Topik dan Batasannya

Permasalahan yang dijadikan di dalam penelitian ini adalah :

1. Bagaimana hasil performansi dari algoritma ANN dan GA untuk diagnosa pada penyakit kanker.
2. Berapa nilai hyperparameter terbaik agar klasifikasi kanker *colon* menjadi optimal.

Batasan masalah dalam penelitian ini adalah :

1. Data kanker *colon* yang diambil dari Kent Ridge Biomedical Dataset Repository.
2. Data dibagi menjadi 80% *training*, 20% *testing*.

Tujuan

Dari permasalahan yang ada, berikut tujuan yang hendak dicapai dalam penelitian tugas akhir ini, yaitu:

1. Mengetahui dan mencari struktur terbaik dalam *Artificial Neural Network* (ANN) dan *Genetic Algorithm* (GA) dalam mereduksi dimensi pada microarray data.
2. Mengetahui hasil performansi *Artificial Neural Network* (ANN) dan *Genetic Algorithm* (GA) pada data microarray.

Organisasi Tulisan

Pada bagian 2 berisi mengenai literatur dari penelitian, pada bagian 3 berisi langkah-langkah pengerjaan dalam penelitian, pada bagian 4 berisi mengenai pengerjaan penelitian dan pada bagian 5 berisi kesimpulan dari penelitian.

2 Studi Terkait

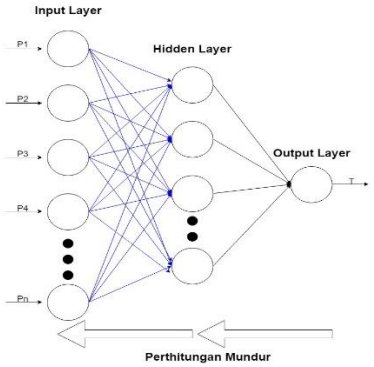
2.1 Studi Literatur

Penelitian untuk klasifikasi pada data miccroarray beberapa tahun terakhir telah banyak dilakukan, seperti penelitian yang dilakukan oleh Adam Mizza [3] yang melakukan klasifikasi dengan menggunakan ANN sebagai *classifier* dan GA untuk mengoptimalisasi *learning rate* dan *hidden neuron* yang menghasilkan nilai akurasi sebesar 97% pada data kanker payudara. Adiwijaya [2] melakukan penelitian yang membandingkan antara *Support Vector Machine* dan *Levenberg-Marquardt Backpropagation* (LMBPP) dengan menggunakan *Principal Component Analysis* (PCA) dimana SVM menghasilkan nilai akurasi sebesar 94.98% dan LMBP menghasilkan nilai akurasi sebesar 96.07%. Penelitian lain dilakukan oleh Aydadenta [4] yang membandingkan beberapa metode seperti *Support Vector Machine* (SVM), *Artificial Neural Network* (ANN), *Naive Bayes*, *k-Nearest Neighbor* (kNN), dan C4 .5 untuk klasifikasi data kanker dan menghasilkan akurasi *Support Vector Machine* (SVM) sebesar 66.67%, *Artificial Neural Network* (ANN) sebesar 66.67%, *Naive Bayes* sebesar 41.67%, *k-Nearest Neighbor* (kNN) sebesar 58.33%, dan C4 .5 sebesar 58.33%.

Menurut yip [6] terdapat 6 jenis klasifikasi yang dapat digunakan, salah satunya adalah *Artificial Neural Network* (ANN). ANN memiliki kelebihan dapat menghasilkan akurasi yang cukup tinggi untuk data yang berdimensi besar atau data yg mengandung noise, namun memiliki kelemahan berupa beban komputasi yang cukup tinggi dan parameter yang berbeda untuk setiap kasus.

2.2 *Artificial Neural Network* (ANN)

ANN adalah sistem komputasi dimana arsitektur dan operasi diilhami dari pengetahuan tentang sel syaraf biologis dalam otak. Istilah ANN digunakan karena jaringan syaraf ini diimplementasikan dengan menggunakan program komputer yang mampu menyelesaikan sejumlah proses perhitungan selama proses pembelajaran, cara kerja jaringan syaraf tiruan meniru cara kerja otak manusia [8].



Gambar 1. Struktur ANN [10]

Arsitektur dasar pada ANN terdiri dari tiga tahap yaitu input, hidden layer, dan output layer. Dalam arsitektur jaringan maju, sinyal mengalir dari input hingga output secara ketat ke arah maju [6]. Pada ANN terdapat algoritma Backpropagation dimana terdapat perhitungan maju (feedforward) dan mundur (backward). Diketahui bersama bahwa prediksi yang dihasilkan dari model Neural Network akan lebih akurat jika parameter seperti jumlah unit lapisan tersembunyi (hidden layer) dan learning rate dapat dioptimalkan secara benar dan tepat [12]. Berikut adalah rumus perhitungan pada ANN :

Pada *hidden layer* :

$$V1 = P \times W1 + B1 \quad (1)$$

$$A1 = \frac{1}{1 + e^{-V1}} \quad (2)$$

Pada *output layer* :

$$V2 = A1 \times W2 + B2 \quad (3)$$

$$A2 = \frac{1}{1 + e^{-V2}} \quad (4)$$

$$E = T - A2 \quad (5)$$

$$MSE = \frac{\sum E^2}{N} \quad (6)$$

Perhitungan mundur :

$$D2 = A2 \times (1 - A2) \times E \quad (7)$$

$$D1 = A1 \times (1 - A1) \times (W2 \times D2) \quad (8)$$

$$dW1 = lr \times D1 \times P \quad (9)$$

$$dW2 = lr \times D2 \times P \quad (10)$$

$$dB1 = lr \times D1 \quad (11)$$

$$dB2 = lr \times D2 \quad (12)$$

$$W1 = W1 + dW1 \quad (13)$$

$$W2 = W2 + dW2 \quad (14)$$

$$B1 = B1 + dB1 \quad (15)$$

$$B2 = B2 + dB2 \quad (16)$$

Keterangan :

V1 : nilai pada hidden layer; V2 : nilai pada output layer; P : matriks pada input layer; B1 : bias pada bobot; B2 : bias pada output; A1 : nilai pada hidden layer (perhitungan maju); A2 : nilai pada output layer (perhitungan maju); W1 : bobot pada *hidden layer*; W2 : bobot pada output layer; T : matriks

target ; MSE : mean square error; E : error; N : jumlah data; *lr* : learning rate; D1 : nilai pada hidden layer (perhitungan mundur); D2 : nilai pada output layer (perhitungan mundur); dW1 : delta W1; dW2 : delta W2; dB1 : delta B1; dan dB2 : delta B2.

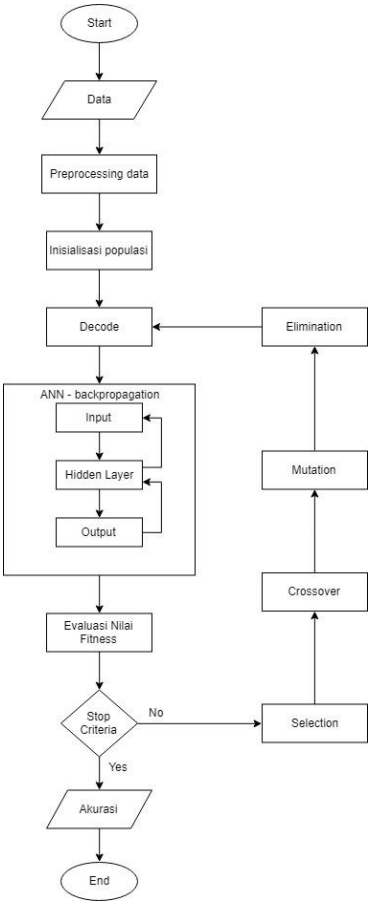
2.3 *Genetics Algorithm* (GA)

Menurut Suyanto, 2005 [7] algoritma genetika merupakan metode adaptif yang biasa digunakan untuk memecahkan suatu pencarian nilai dalam sebuah masalah optimasi. Algoritma ini sendiri pertama kali dipublikasikan pada tahun 1975 oleh Jhon Holland di Amerika Serikat [11]. Algoritma genetika berdasar pada dua buah teori yaitu "evolusi" dan "genetika" dimana setiap individu dalam populasi akan terkena seleksi yang didasarkan dari nilai fitness-nya agar diperoleh keturunan yang lebih baik.

2.4 ANN-GA *hybrid*

Metode *Artificial Neural Network* dan *Genetics Algorithm hybrid* atau yang sering disingkat ANN-GA Hybrid terbentuk karena algoritma ANN memiliki beberapa kelemahan yang cukup berpengaruh seperti waktu training yang cukup lama dan hasil yang dihasilkan merupakan optimum lokal bukan optimum global [9]. Maka dari itu untuk memperbaiki beberapa kelemahan yang dimiliki oleh ANN digunakan GA yang sudah terbukti efektif dalam mereduksi dimensi dengan memanfaatkan fitur *crossover* dan mutasi sehingga memungkinkan fitur yang sebelumnya tidak terpilih menjadi bisa terpilih, hal ini membuat GA mampu memberikan solusi yang optimum. Pada penelitian ini, fitur yang terpilih oleh GA akan diklasifikasikan menggunakan ANN dengan merubah beberapa parameter seperti jumlah *hidden neuron* (5,10,15,20,25) dan jumlah *epoch* (25,50,75,100). Nilai akurasi yang didapatkan dari tiap individu pada proses ANN akan digunakan sebagai nilai *fitness* pada GA.

3 Sistem yang Dibangun



Gambar 2: flow chart Penelitian

3.1 Data yang digunakan pada penelitian ini menggunakan DNA *Microarray Data* untuk kasus kanker *colon* yang diambil dari Kent Ridge Data Set Repository dan tersedia di http://leo.ugr.es/elvira/DBCRepository/. Pada sistem yang akan dibangun, Dataset dibagi menjadi 2, yaitu data *training* dan data *testing*. 80% data training dan 20% data testing. Berikut adalah data yang digunakan pada penelitian ini.

Tabel 1. Dataset

Data	Kelas	Jumlah Sampel	Jumlah *Fitur*
Colon	2	62 (41 negative, 22 positive)	2000

3.2 *Preprocessing Data*

Dikarenakan *Range* pada masing masing atribut pada data terlalu jauh maka dilakukan normalisasi data dengan rumus sebagai berikut :

$$a = \min(data) \quad (17)$$

$$b = \max(data) \quad (18)$$

$$ra = 0.9 \quad (19)$$

$$rb = 0.1 \quad (20)$$

$$Dn = \frac{(ra - rb) \times (data - a)}{(b - a)} + rb \quad (21)$$

Keterangan:
a = data paling rendah; b = data paling tinggi; ra = data paling tinggi nilainya; rb = data paling rendah nilainya; Dn = data hasil normalisasi.

3.3 Inisialisasi populasi / *generate chromosomes*

Setelah normalisasi data lalu munculkan nilai random 0 atau 1 untuk tiap *fitur* lalu proses ini diulangi sebanyak 10 kali hingga menciptakan 10 individu dan kemudian disebut dengan populasi 1. Contoh :

0	1	1	0	1	1	0	1	……	1
0	0	0	1	0	1	1	1	……	0
:									
1	1	0	1	0	0	0	1	……	1

3.4 *Decode* / membaca chromosomes

Pada proses ini individu yang telah dibentuk pada proses sebelumnya diterjemahkan pada data untuk digunakan sebagai variable input pada ANN. Dengan contoh :

0	1	1	0	1	……	1

menjadi

0	5215.04	4288.616	0	3147.402	……..	89.4537

3.5 ANN – *Backpropagat*ion

Proses ini digunakan untuk mendapatkan akurasi yang selanjutnya dijadikan sebagai nilai fitness. Setelah variable input didapatkan ANN-*Backpropagation* dilakukan dengan tahapan:

a. Tahap *Input*. Nilai *input* merupakan nilai yang telah dipilih pada tahap *decode.*
b. Tahap *hidden Layer*. Pada tahap ini masing-masing nilai input dihitung dengan menggunakan rumus perhitungan pada *hidden layer.*
c. Tahap *output*. Pada tahap ini nilai pada *hidden layer* dihitung dengan menggunakan rumus *output*.
d. Tahap *backward*. Pada tahap ini nilai pada *output* dihitung dengan menggunakan rumus perhitungan mundur.
e. Tahap *backward*. Pada tahap ini nilai pada *hidden layer* dihitung dengan menggunakan rumus perhitungan mundur.
f. Setelah itu dilakukan kembali tahap d dan e hingga nilai kriteria terpenuhi.
g. Setelah itu nilai bobot yang didapatkan dari *training* digunakan untuk menguji data *testing* sehingga mendapatkan hasil akurasi.

3.6 *Fitness Evaluation*

Pada proses ini nilai akurasi yang didapatkan dari tiap *chromosomes* pada proses ANN digunakan sebagai nilai *fitness* pada GA. Dengan rumus akurasi sebagai berikut :

$$akurasi = \frac{jumlah\ pola\ benar}{jumlah\ data\ testing} \quad (22)$$

3.7 *Stop criteria*

stop criteria adalah dimana jika sudah mencapai nilai iterasi maka proses GA akan berhenti, jika belum terpenuhi maka proses GA akan berlanjut

3.8 *Selection*

Pada proses ini *chromosomes* pada populasi awal akan dipilih menggunakan teknik *roulette wheel* dan disebut dengan populasi 2.

3.9 *Crossover*

Pada tahap ini setiap individu dalam populasi 2 akan dibangkitkan suatu nilai random, jika nilai pc (peluang *crossover*) pada individu memiliki nilai dibawah yang ditentukan maka terjadi crossover, dengan pc = 0,5.

3.10 *Mutation*

Pada tahap ini setiap fitur pada individu di populasi 2 dibangkitkan suatu nilai random, apabila nilai random < pm (peluang mutasi) maka nilai gen akan diganti dengan inversnya. Hal ini dilakukan untuk mengembalikan nilai yang hilang pada tahap sebelumnya, dengan pm = 0,1.

3.11 *Elimination*

Dalam proses ini 2 individu dengan nilai *fitness* terendah dari populasi 1 (sebelum *selection*) akan ditukar dengan *individu* dengan nilai *fitness* terbaik dari populasi 2 (setelah *mutation*)

3.12 Tahap selanjutnya adalah kembali ke tahap 2 (*decode*) dan terus berulang sampai kriteria terpenuhi.

4 Evaluasi

4.1 Skenario Pengujian

Pada penelitian ini penulis menggunakan 3 skenario pengujian yang dilakukan. Skenario pertama adalah menguji pengaruh jumlah *hidden neuron* dan *epoch* untuk nilai akurasi pada ANN. Skenario kedua adalah menguji pengaruh *hidden neuron* dan *epoch* untuk nilai akurasi pada ANN-GA *hybrid*. Skenario ketiga adalah membandingkan hasil performansi antara ANN dan ANN-GA *hybrid*.

Pengujian ini menggunakan parameter :

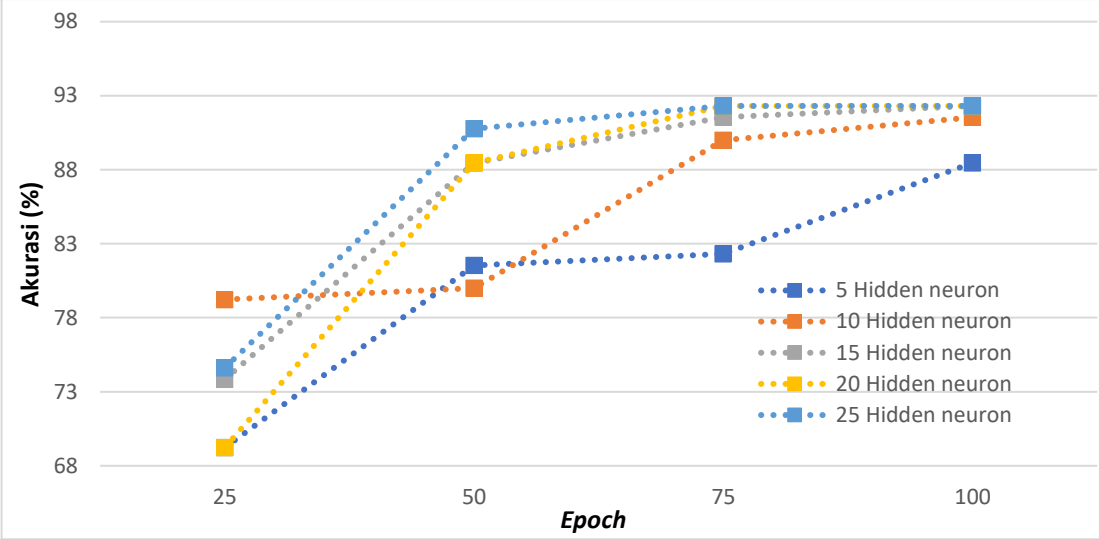
Pm (peluang mutasi) = 0.1	Jumlah generasi pada GA = 10 generasi
Pc (peluang crossover) = 0.5	Epoch = 25, 50, 75, 100.
Lr (learning rate) = 0.1	Hidden Neuron = 5, 10, 15, 20, 25.

Pengujian ini dilakukan pada komputer dengan spesifikasi hardware yang digunakan: processor : intel core i7-4700HQ ; RAM : 10GB DDR3 dan spesifikasi software yang digunakan: OS : Windows 10 pro; aplikasi : MATLAB – R2015a.

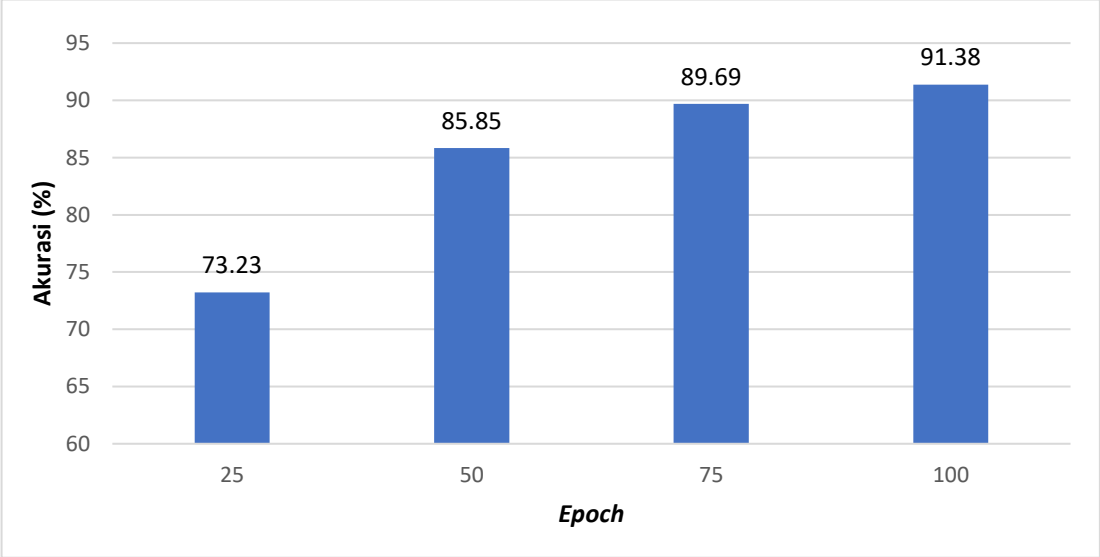
4.2 Pengujian

4.2.1 Pengaruh hidden neuron dan epoch untuk nilai akurasi pada ANN

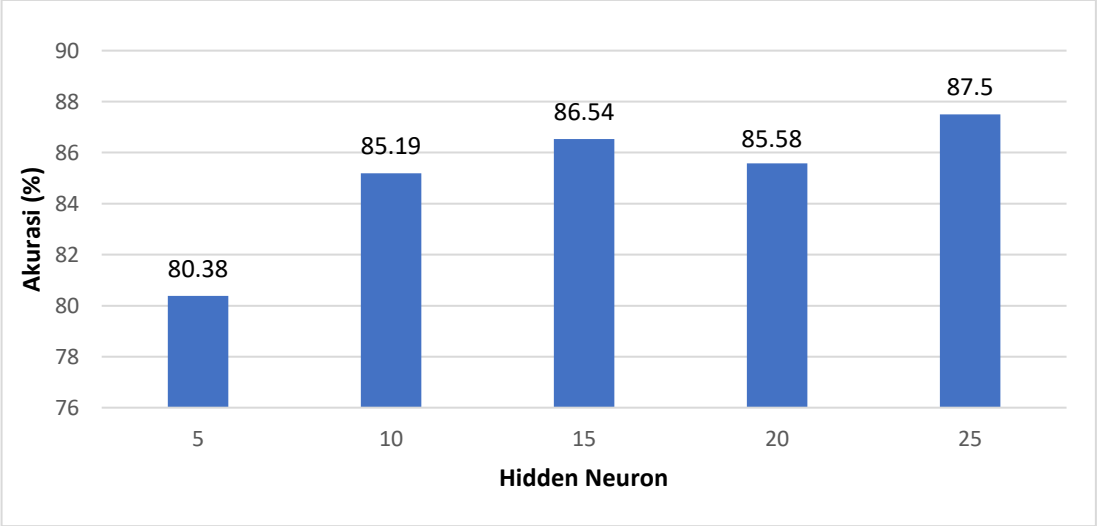
Pengujian ini dilakukan untuk mengetahui pengaruh dari *hidden neuron* dan *epoch* pada struktur ANN terhadap hasil akurasinya dengan menggunakan 5 *hidden neuron* yaitu 5, 10, 15, 20 dan 25. Sedangkan jumlah *epoch* yang digunakan adalah 25, 50, 75 dan 100. Dikarenakan nilai akurasi pada ANN selalu berubah-ubah pada setiap kali pengujian, maka agar lebih akurat, pengujian ini dilakukan sebanyak 10 kali dan diambil nilai rata-ratanya. Berikut adalah tabel hasil pengujian ANN :



Gambar 3. Grafik pengujian ANN



Gambar 4. Grafik rata-rata akurasi setiap *hidden neuron* pada *epoch*



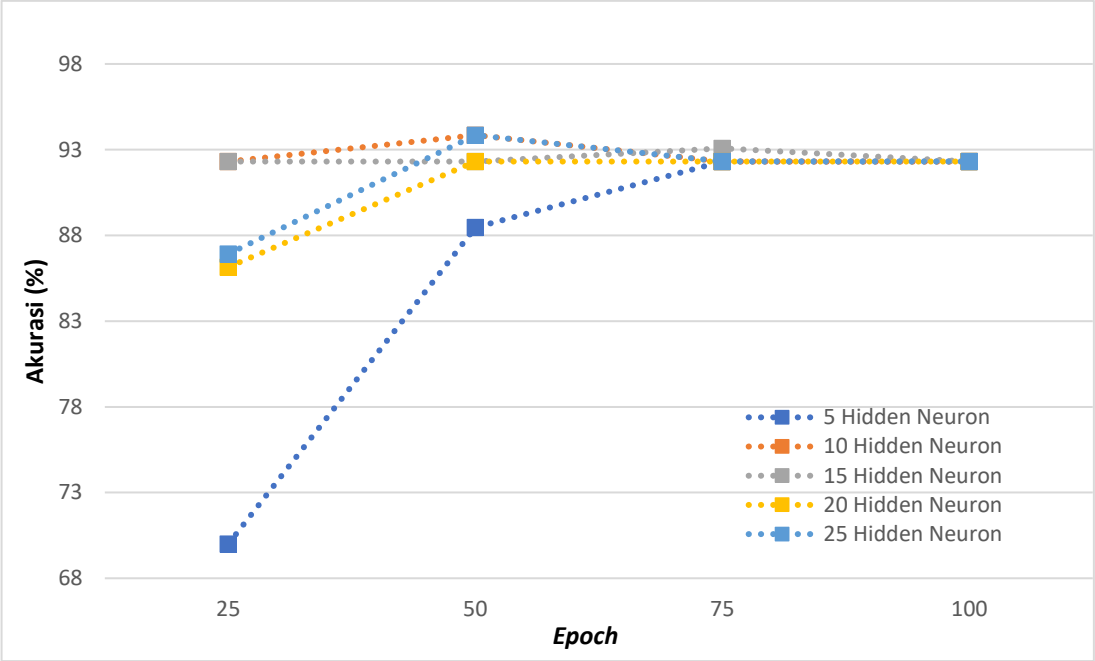
Gambar 5. Grafik rata-rata akurasi setiap *epoch* pada *hidden neuron*

Berdasarkan grafik *epoch* (Gambar 4) diatas menunjukan bahwa nilai akurasi berbanding lurus dengan banyaknya *epoch*, dimana semakin banyak jumlah *epoch* maka semakin tinggi nilai akurasi

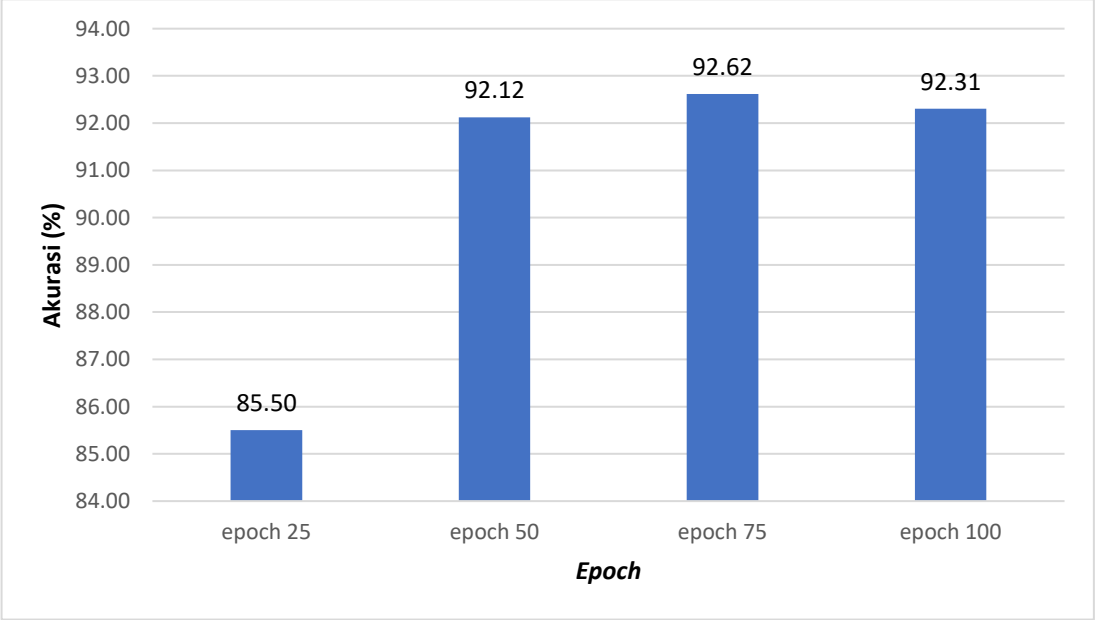
yang di dapat. Hal tersebut terjadi karena *epoch* merupakan proses pengulangan dalam metode learning ANN, dimana semakin banyak proses pengulangan terjadi maka semakin banyak pula proses learning pada ANN sehingga menghasilkan nilai akurasi yang semakin baik. Hasil pengujian *hidden neuron* menunjukan nilai akurasi tertinggi terdapat pada 25 *hidden neuron* yaitu sebesar 87.5%.

4.2.2 Pengaruh hidden neuron dan epoch untuk nilai akurasi pada ANN-GA

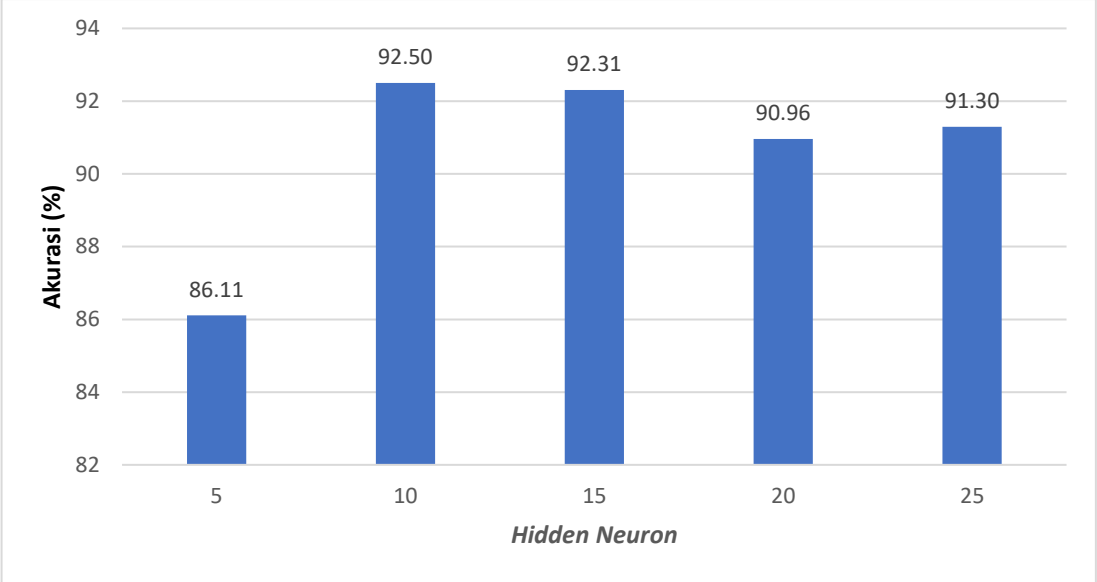
Pengujian ini dilakukan untuk mengetahui pengaruh dari *hidden neuron* dan *epoch* pada struktur ANN-GA *hybrid* terhadap hasil akurasinya dengan menggunakan 5 *hidden neuron* yaitu 5, 10, 15, 20 dan 25. Sedangkan jumlah *epoch* yang digunakan adalah 25, 50, 75 dan 100. Berikut adalah hasil pengujian ANN-GA :



Gambar 6. Grafik pengujian ANN- GA



Gambar 7. Grafik rata-rata akurasi setiap hidden neuron pada epoch



Gambar 8. Grafik rata-rata akurasi setiap epoch pada hidden neuron

Berdasarkan grafik pengujian *epoch* (Gambar 7) diatas menunjukan bahwa pada *epoch* 50 - 100 bernilai konvergen dengan nilai akurasi sebesar 92%. Hal tersebut dapat terjadi karena ANN sudah dioptimasi menggunakan GA sehingga pada *epoch* ke-50 sudah mencapai nilai konvergen. Hal ini dibuktikan oleh bilgehan, 2010 [13] dimana semakin banyak *epoch* maka nilai akurasi akan semakin tinggi dan menjadi konvergen. Hasil pengujian *hidden neuron* (Gambar 8) menunjukan nilai akurasi tertinggi terdapat pada 10 *hidden neuron* yaitu sebesar 92,5%.

4.2.3 Perbandingan nilai akurasi antara ANN dengan ANN-GA

Pengujian ini dilakukan untuk melihat apakah metode ANN-GA *hybrid* lebih optimal dibandingkan dengan metode ANN-*backpropagation*. Berikut merupakan tabel hasil hasil perbandingan antara ANN dengan ANN-GA :

Tabel 2. Perbandingan nilai akurasi pada ANN dan ANN-GA pada 25 *hidden neuron* dan 100 *epoch*

	ANN	ANN-GA
Nilai Akurasi	92.31 %	92.31 %
Fitur yang digunakan	2000 fitur	969 fitur
Waktu yang diperlukan	11,197 detik	7,649 detik

Tabel 3. Perbandingan nilai akurasi pada ANN dan ANN-GA pada 10 *hidden neuron* dan 75 *epoch*

	ANN	ANN-GA
Nilai Akurasi	90 %	93.08%
Fitur yang digunakan	2000 fitur	971 fitur
Waktu yang diperlukan	2,871 detik	1,660 detik

Tabel 4. Perbandingan nilai rata-rata akurasi pada ANN dan ANN-GA

	ANN	ANN-GA
Nilai Akurasi	85,03 %	90,63 %
Fitur yang digunakan	2000 fitur	984 fitur

Berdasarkan tabel perbandingan (Tabel 2, 3 dan 4) diatas menunjukkan bahwa ANN-GA lebih efektif dibandingkan ANN karena ANN-GA karena GA dapat mengoptimalisasi ANN dengan mereduksi fitur dan tetap mempertahankan fitur-fitur yang penting sehingga nilai akurasi yang dihasikan ANN-GA lebih tinggi dengan fitur yang lebih sedikit dibandingkan dengan ANN.

5 Kesimpulan

Berdasarkan hasil pengujian dan analisis yang telah dilakukan, maka dapat disimpulkan bahwa :

1. Metote ANN-GA *hybrid* terbukti lebih efektif daripada metode ANN-*Backpropagation* dalam mengklasifikasi data microarray karena dapat mereduksi fitur hingga 50,8% dengan nilai akurasi

sebesar 90,63% dan waktu *running time* lebih cepat hingga 42.2% dibandingkan ANN, sedangkan nilai akurasi metode ANN-*Backpropagation* sebesar 85,03% tanpa ada fitur yang direduksi.

2. Atribut terbaik yang didapatkan oleh ANN adalah 25 *hidden neuron* dan 100 *epoch* dengan nilai akurasi sebesar 92.3% tanpa ada fitur yang direduksi. Atribut terbaik yang didapatkan oleh ANN-GA adalah 10 *hidden neuron* dan 75 *epoch* dengan nilai akurasi sebesar 93,08% dengan fitur yang direduksi mencapai 51,55%.

Daftar Pustaka

[1] Pusat Data Dan Informasi Kementrian Kesehatan RI (InfoDATIN). 2015. 4 Februari-Hari Kanker Sedunia. Indonesia: Pusdatin.

[2] Adiwijaya, U. N. Wisesty, E. Lisnawati, A. Aditsania, D. S. Kusumo, (2018). Dimensionality Reduction using Principal Component Analysis for Cancer Detection based on Microarray Data Classification, Journal of Computer Science 14(11)

[3] Adam Z M, Bilqis A, Munif Abdul. 2012. Implementasi Algoritma Genetika pada Struktur Backpropagation Neural Network untuk Klasifikasi Kanker Payudara. Jurnal Teknik ITS Vol. 1, 2301-9271.

[4] Husna Aydadenta, Adiwijaya, 2018, A Clustering Approach for Feature Selection in Microarray Data Classification using Random Forest, Journal of Information Processing System 14(5)

[5] Abraham A. 2005. Artificial Neural Networks. Handbook of Measuring System Design.

[6] Yip Wai-Ki, Amin Samir B, Li Cheng. 2011. Chapter 10: A Survey of Classification Techniques for Microarray Data Analysis. In Springer Handbooks of Computational Statistics.

[7] Suyanto. 2008. Soft Computing, Membangun Mesin Ber-IQ Tinggi. Bandung : Penerbit Informatika.

[8] Siang JJ. 2009. Jaringan Syaraf Tiruan dan Pemrogramannya Menggunakan Matlab. Yogyakarta : Penerbit Andi.

[9] Inthachot M, Boonjing V, Intakosum S. 2016. Artificial Neural Network and Genetic Algorithm Hybrid Intelligence for Predicting Thai Stock Price Index Trend. Computational Intelligence and Neuroscience. Hindawi Publishing corporation. 3045254.

[10] Prathama, A. Y., Akhmad Aminullah, dan Ashar Saputra. 2017. Pendekatan ANN (*Artificial Neural Network*) untuk penentuan Prosentase Bobot Pekerjaan dan Estimasi Nilai Pekerjaan Struktur pada Rumah Sakit Pratama. Jurnal Teknosains. Vol 7 No. 1, 14-25.

[11] Holland, John H. 1975.Adaptation in Neural and Artificial Systems. University of Michigan Press. United State.

[12] Suyanto. 2002. Intelijensia Buatan. Bandung : Sekolah Tinggi Teknologi Telkom.

[13] Bilgehan, M. Turgut, P. 2010. Artificial Neural Network Approach to Predict Compressive Strength of Concrete through Ultrasonic Pulse Velocity. Research in Nondestructive Evaluation, 21: 1, 1 — 17.

[14] Nurfalah, A. Adiwijaya, and Suryani, A.A., (2016). Cancer Detection Based On Microarray Data Classification Using PCA And Modified Back Propagation. Far East Journal of Electronics and Communications, 16(2), p.269.

[15] 4Pratiwi, M. S., Aditsania, A., Adiwijaya. 2018. Cancer Detection Based on Microarray Data Classification using Genetic Bee Colony (GBC) and Conjugate Gradient Backpropagation with Modified Polak Ribiere (MBP-CGP). In 2018 International Conference on Computer, Control, Informatics and its Applications (IC3INA) (pp. 163-168). IEEE.

[16] Astuti, W., & Adiwijaya, A. (2019). Principal Component Analysis Sebagai Ekstraksi Fitur Data Microarray Untuk Deteksi Kanker Berbasis Linear Discriminant Analysis. JURNAL MEDIA INFORMATIKA BUDIDARMA, 3(2), 72-77.

[17] Adiwijaya, A. (2018). Deteksi Kanker Berdasarkan Klasifikasi Microarray Data. JURNAL MEDIA INFORMATIKA BUDIDARMA, 2(4), 181-186.

[18] Ma'ruf, F. A., Adiwijaya & Wisesty, U. N. (2019, March). Analysis of the influence of Minimum Redundancy Maximum Relevance as dimensionality reduction method on cancer classification based on microarray data using Support Vector Machine classifier. In Journal of Physics: Conference Series (Vol. 1192, No. 1, p. 012011). IOP Publishing.

[19] Purbolaksono, M. D., Widiastuti, K. C., Adiwijaya, Mubarok, M. S., & Ma'ruf, F. A. (2018, March). Implementation of mutual information and bayes theorem for classification microarray data. In Journal of Physics: Conference Series (Vol. 971, No. 1, p. 012011). IOP Publishing.

[20] Manik, A., Adiwijaya, A., & Utama, D. Q. (2019). Classification of Electrocardiogram Signals using Principal Component Analysis and Levenberg Marquardt Backpropagation for Detection Ventricular Tachyarrhythmia. Journal of Data Science and Its Applications, 2(1), 78-87.