

Deteksi Kanker Berdasarkan Klasifikasi *Microarray* Data Menggunakan *Wavelet Transform* dan *Backpropagation* Termodifikasi dengan *Conjugate Gradient Fletcher Reeves*

Syahrizal Rizkiana Rusamsi¹, Adiwijaya², Aniq Atiqi Rohmawati³
^{1,2,3} Program Studi S1 Ilmu Komputasi, Fakultas Informatika, Universitas Telkom

Jl. Telekomunikasi No. 1, Terusan Buah Batu Bandung 40257 Indonesia

¹ rusamsisyahrizal@gmail.com, ² adiwijaya@telkomuniversity.ac.id, ³ aniqatiqi@telkomuniversity.ac.id

Abstrak - Kanker adalah penyakit mematikan di dunia, penyebabnya adalah sel sel yang membelah diri secara tidak terkendali. Dalam perkembangannya sel-sel kanker ini dapat menyebar kebagian tubuh lainnya dengan cara memasuki aliran darah atau sistem limfatik sehingga dapat menyebabkan kematian. Hal ini terjadi akibat adanya kecacauan yang terjadi pada gen. Pemantauan ekspresi gen adalah salah satu yang paling mendasar dalam genetika, yaitu untuk mengukur mRNA bukan protein, karena urutan mRNA berhibridisasi dengan komplementer RNA dan DNA. *Microarray* merupakan teknologi yang dapat menyimpan ribuan gen yang diambil dalam beberapa sel manusia sekaligus. Data *microarray* cenderung memiliki dimensi yang sangat besar, sehingga untuk memodelkan data tersebut diperlukan reduksi dimensi dengan menggunakan *Wavelet Haar* untuk meningkatkan akurasi klasifikasi digunakan *Modified Back Propagation* (MBP) Fletcher Reeves. MBP merupakan modifikasi dari *Back Propagation Standar* (BP) dengan algoritma *Conjugate Gradient Fletcher Reeves* untuk mempercepat proses pelatihan. Dari hasil pengujian, skema *Wavelet Haar* dan *Modified Backpropagation Fletcher Reeves* dengan teknik *line search golden section search* memiliki akurasi yang terbaik dari skenario klasifikasi yaitu pada jenis data Kanker Leukimia dengan akurasi 88,9%.

Kata Kunci : Kanker, *Microarray*, *Wavelet Haar*, *Modified Back Propagation Fletcher Reeves*.

Abstract - Cancer is a deadly disease in the world, the cause is a cell cell that divides itself uncontrollably. In its development these cancer cells can spread to other body parts by entering the bloodstream or lymphatic system so that it can cause death. This happens as a result of what happens to genes. Gene expression monitoring is one of the most fundamental in genetics, ie to measure non-protein mRNAs, because the order of mRNAs hybridizes with complementary RNA and DNA. *Microarray* is a technology that can store thousands of genes that are taken in several human cells at once. *Microarray* data tend to have very large dimensions, so to model the data dimension reduction is required by using *Haar Wavelet* to improve the accuracy of *Fletcher Reeves Modified Back Propagation* (MBP) conversation. MBP is a modification of *Back Propagation Standard* (BP) with *Conjugate Gradient Fletcher Reeves* algorithm to speed up the training process. From the test results, *Wavelet Haar* and *Backpropagation Fletcher Reeves* with *line search golden section search* has the highest equation of all schemes in *Leukemia Cancer* data type with 88.9% accuracy.

Keyword : Cancer, *Microarray*, *Wavelet Haar*, *Modified Back Propagation Fletcher Reeves*.

1. PENDAHULUAN

Teknologi DNA *microarray* memungkinkan tingkat ribuan ekspresi gen bisa dipantau secara bersamaan melalui satu percobaan, hal ini berguna untuk membantu proses pengolahan data yang sangat besar. Analisis ekspresi gen ini sangat penting untuk berbagai bidang penelitian seperti bioinformatik, biostatistik, kedokteran dan penelitian dibidang kesehatan lainnya [1]. Reduksi dimensi merupakan langkah yang penting pada teknologi *microarray* sebelum dilakukannya klasifikasi. Metode reduksi dimensi adalah menghapus fitur yang tidak relevan dan berlebihan untuk meningkatkan akurasi. Klasifikasi adalah proses penentuan apakah pasien terdeteksi memiliki kanker atau tidak, tujuan klasifikasi adalah untuk mengidentifikasi gen informatif yang dapat digunakan memprediksi kelas dalam pengujian sampel baru. Perangkat lunak diperlukan untuk klasifikasi yang efektif dan analisis data dalam skala besar.

Dalam penelitian ini penulis menggunakan metode *Wavelet Haar* yang berfungsi untuk reduksi dimensi dan metode klasifikasi yang digunakan adalah *Conjugate Gradien BackPropagation* (MBP) *fletcher reeves*. MBP adalah modifikasi dari standar *back propagation* dengan mengimplementasikan algoritma *conjugate gradient* pada *back propagation training*.

2. KAJIAN PUSTAKA

2.1 Studi Literatur

Bai [2] dalam tesis masternya telah membangun *framework* berbasis PCA dan propagasi balik untuk mengklasifikasikan data *microarray*. Dia menggunakan beberapa dataset *microarray* seperti data kanker ovarium, data kanker usus besar, dan data leukemia, menghasilkan akurasi 96% ovarium, 95% akurasi untuk usus besar dan ketepatan 97% untuk leukemia. Dalam hal waktu pelatihan, Bai mengakui bahwa propagasi balik membutuhkan waktu lama untuk pelatihan, sekitar 20-23 detik untuk setiap data.

Huynh dkk[13] membangun kerangka kerja berbasis PCA dan propagasi mundur untuk mengklasifikasikan data *microarray*. PCA dilakukan dengan menerapkan dekomposisi nilai singular (SVD) dalam penelitian ini dilakukan dengan membandingkan keakuratan sistem yang diusulkan dengan metode pembelajaran ekstrim (ELM) dan sistem yang diusulkan mampu mengungguli ELM dengan akurasi 83,6% untuk data kanker usus besar.

2.2 Kanker

Kanker adalah penyakit akibat pertumbuhan tidak normal sel-sel jaringan tubuh yang berubah menjadi sel kanker. Dalam perkembangannya sel-sel kanker ini dapat menyebar kebagian tubuh lainnya sehingga dapat menyebabkan kematian. Statistik WHO menunjukkan bahwa pada tahun 2014 lebih dari 14 juta kasus baru teridentifikasi dan sekitar 8,2 juta kematian terkait kanker. Jumlah kasus baru diperkirakan akan meningkat 70% selama beberapa dekade [1].

2.3 Microarray Data

Microarray adalah teknologi yang mampu menyimpan ribuan ekspresi gen yang diambil dari beberapa jaringan manusia sekaligus. Dengan menganalisis data *microarray*, dapat diketahui apakah jaringan tersebut terkena kanker atau tidak. Studi ini memberikan performa kinerja yg cepat dan akurat untuk mendeteksi kanker berdasarkan klasifikasi data *microarray* [1].

Klasifikasi data *microarray* adalah *supervised learning* yang memprediksi kategori diagnostik dari sebuah sampel fenotipe ekspresi array. Dibutuhkan labeling sampel data *gene expression* agar menghasilkan model *classifier* yang mengklasifikasikan sampel data baru dengan penyakit yang telah ditetapkan.

2.4 Wavelet Haar

Wavelet adalah gelombang – gelombang kecil dimana gelombang – gelombang besar adalah gelombang sinusoida [3]. Wavelet digunakan untuk mentransformasikan sebuah sinyal yang sedang diteliti menjadi sebuah representasi baru yang bisa merepresentasikan informasi sinyal tersebut dengan lebih baik [4]. Ada berbagai jenis – jenis wavelet. Contoh wavelet paling sederhana yang adalah wavelet Haar, yang dikemukakan oleh Alfréd Haar di tahun 1909.

Suatu fungsi $\psi(t)$ yang bernilai real disebut wavelet jika memenuhi :

$$1. \int_{-\infty}^{\infty} \psi_{j,k}(t) dt = 0, \text{ Integral } \psi(t) \text{ adalah nol} \quad (1)$$

$$2. \int_{-\infty}^{\infty} \psi_{j,k}^2(t) dt = 1, \text{ Integral } \psi^2(t) \text{ adalah satu} \quad (2)$$

Setiap wavelet memiliki sebuah persamaan penting yang disebut *mother wavelet* atau *analysis wavelet*. Fungsi wavelet dihasilkan dari dilatasi dan translasi *mother wavelet* [7] yang dinyatakan dengan persamaan:

$$\psi_{a,b}(t) = a^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right); a > 0, -\infty < b < \infty \quad (3)$$

Selain fungsi *mother wavelet* $\psi(t)$, dalam analisis *wavelet* juga terdapat fungsi penting lainnya yaitu fungsi *father wavelet* $\phi(t)$ atau bisa disebut fungsi skala. Fungsi $\phi(t)$ berfungsi untuk membangkitkan fungsi basis ortonormal yang menyusun ruang $L^2(R)$. Fungsi basis ortonormal tersebut bisa dituliskan sebagai:

$$\{\phi_{j_0,k}, \psi_{j,k}, j_0 \leq j, k \in Z\} \quad (4)$$

$\phi_{0,0}(t)$ (fungsi skala) berhubungan dengan fungsi $\psi(t)$, dimana $\{\phi_{j_0,k}, k \in Z\}$ akan membentuk anak ruang yang sama dengan $\{\psi_{j,k}, j_0 \leq j, k \in Z\}$ [17]. Contohnya *father wavelet* dari *wavelet* Haar:

$$\phi(t) = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{selainnya} \end{cases} \quad (5)$$

Matriks transformasi pada *wavelet* merupakan matriks yang orthogonal dimana didalam matriks tersebut terbentuk dari *father* $\phi(t)$ dan *mother wavelet* $\psi_{j,k}$ yang nilainya dapat diperoleh dari fungsi *father* (2.3) dan *mother* (2.7), dapat dituliskan dengan notasi matriks,

$$x = W^T d \quad (6)$$

Karena W orthogonal maka

$$d = Wx \quad (7)$$

dimana $d = (c_{0,0}, d_{0,0}, d_{1,0}, d_{1,1})$ [11]. Lalu reduksi dilakukan dengan mengambil $m < q$ dari koefisien-koefisien *wavelet* [11] sehingga,

$$D_{nxm}^* = X_{n \times p} W_{p \times m}^{*T} \quad (8)$$

dengan memberi nilai nol pada $(m+1)$ sampai dengan q dari matriks W^T , persamaan (2.11) mereduksi pengamatan variabel p menjadi m koefisien *wavelet* terpilih, dimana $m < p$ [11].

2.5 Klasifikasi Conjugate Gradient Fletcher Reeves

Metode conjugate gradient (CG) merupakan metode optimasi untuk meminimumkan fungsi yang arah pencariannya berdasarkan pada arah orthogonal [1]. Secara umum algoritma ini lebih cepat konvergen daripada metode penurunan tercepat. Conjugate gradien (CG) menggunakan vector tidak nol yang orthogonal dan bebas linier [1]. Dua vector orthogonal dan dikatakan orthogonal (Gconjugate) jika perkalian dalamnya bernilai nol (0), dapat ditulis sebagai berikut [1] :

$$d^T d_i = 0 \quad (9)$$

Berikut adalah algoritma pelatihan Conjugate Gradient Fletcher Reeves yang digunakan :

1. Definisikan masalah, misalkan matriks masukan(P) dan matriks target (T)
2. Inisialisasi, menentukan arsitektur jaringan, nilai batas MSE dan batas jumlah epoch sebagai kondisi berhenti, line search yang digunakan, membangkitkan bobot – bobot (W) dan bias (b) dengan bilangan acak kecil.
3. Jika kondisi (batas jumlah epoch) penghentian belum terpenuhi, lakukan langkah proses 2-13

Propagasi Maju

1. Hitung keluaran dari *hidden layer* y_j ($j = 1, 2, \dots, p$).

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

$$y_j = f(y_{-net_j}) = \frac{1}{1 + e^{-y_{-net_j}}} \quad (10)$$

2. Lanjutkan dengan menghitung keluaran dari output layer z_k ($j=1, 2, \dots, m$),

$$z_{-net_k} = w_{k0} + \sum_{j=1}^p y_j w_{kj}$$

$$z_k = f(z_{-net_k}) = \frac{1}{1 + e^{-z_{-net_k}}} \quad (11)$$

Propagasi Mundur

- 1 Hitung faktor kesalahan di unit output berdasarkan perbedaan(error) nilai actual dan nilai prediksi (output dari unit output)

$$\delta_k = (t_k - z_k) f'(z_{net_k}) = (t_k - z_k) z_k (1 - z_k) \quad (12)$$

- 2 Hitung faktor kesalahan di unit *hidden layer* berdasarkan factor sebelumnya

$$\delta_{net_k} = \sum_{k=1}^m \delta_k W_{kj}$$

$$\delta_j = \delta_{net_j} f'(y_{net_j}) = \delta_{net_j} y_j (1 - y_j) \quad (13)$$

- 3 Hitung *gradient* di unit output dari fungsi objektif yang sudah ditentukan

$$g_{k+1} = \frac{1}{N} \sum_{n=1}^p \delta_{nk} y_{nk} \quad (14)$$

- 4 Hitung *gradient* di unit *hidden layer*

$$g_{j+1} = \frac{1}{N} \sum_{n=1}^p \delta_{nj} z_{nj} \quad (15)$$

- 5 Hitung parameter β untuk semua *neuron* pada unit tersembunyi dan unit output. Parameter β untuk *Conjugate Gradient Fletcher Reeves* dapat dihitung dengan persamaan :

$$\beta_{k+1} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \quad (16)$$

dimana :

$$\beta_{k+1} = \beta \text{ nilai parameter pada iterasi saat } t+1$$

g_{k+1} = gradient pada iterasi saat t+1
 g_k = gradient pada iterasi sebelumnya

- 6 Hitung direction untuk semua *neuron* pada unit *hidden layer* dan unit output
- $$d_{t+1} = -g_{t+1}\beta_t d_t \quad (17)$$

dimana :

d_{t+1} = direction pada iterasi saat ini
 g_{t+1} = gradient pada iterasi saat ini
 β_t = nilai parameter β pada iterasi sebelumnya
 d_t = direction pada iterasi sebelumnya

untuk *direrction* awal : $d_1 = -g_1$

- 7 Hitung parameter α untuk semua neuron di unit *hidden layer* dan unit output dengan teknik *line search*.

Perubahan Bobot

Update bobot dilakukan dengan cara sebagai berikut :

$$w_{t+1} = w_t + \alpha_{t+1} d_{t+1} \quad (18)$$

dimana :

w_{t+1} = bobot yang akan diupdate
 w_t = bobot sebelumnya
 α_{t+1} = nilai alfa saat ini
 d_{t+1} = *direction* pada iterasi saat ini

Parameter α disini dapat dicari dengan menggunakan teknik *line search* sehingga dapat meminimumkan kinerja selama arah pencarian, karena parameter α merepresentasikan besar langkah yang diambil untuk setiap *direction* [10] Adapun Teknik *Line Search* yang digunakan

1. *Golden Section Search*, yakni pencarian linear sederhana tanpa harus menghitung gradient dari garis tersebut.
2. *Hybrid Bisection-Cubic Search*, yakni pencarian linear dengan menggunakan kombinasi antara metode biseksi dan interpolasi kubik.

2.6 Confusion Matrix

Pada pengukuran kinerja menggunakan *confusion matrix*, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN). Nilai *True Negative* (TN) merupakan jumlah data negatif yang terdeteksi dengan benar, sedangkan *False Positive* (FP) merupakan data negatif namun terdeteksi sebagai data positif. Sementara itu, *True Positive* (TP) merupakan data positif yang terdeteksi benar. *False Negative* (FN) merupakan kebalikan dari *True Positive*, sehingga data positif, namun terdeteksi sebagai data negatif.

Tabel 1 Confusion Matrix

Kelas	Terklasifikasi Positif	Terklasifikasi Negatif
Positif	(TP) True Positif	(FN) False Negatif
Negatif	(FP) False Positif	(TN) True Negatif

Berdasarkan nilai True Negative (TN), False Positive (FP), False Negative (FN), dan True Positive (TP) dapat diperoleh nilai akurasi, presisi dan recall. Nilai akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar. Dengan kata lain, nilai akurasi merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data. Nilai akurasi dapat diperoleh dengan Persamaan 1. Nilai presisi menggambarkan jumlah data kategori positif yang diklasifikasikan secara benar dibagi dengan total data yang diklasifikasi positif. Presisi dapat diperoleh dengan Persamaan 2. Sementara itu, recall menunjukkan berapa persen data kategori positif yang terklasifikasikan dengan benar oleh sistem. Nilai recall diperoleh dengan Persamaan 3.

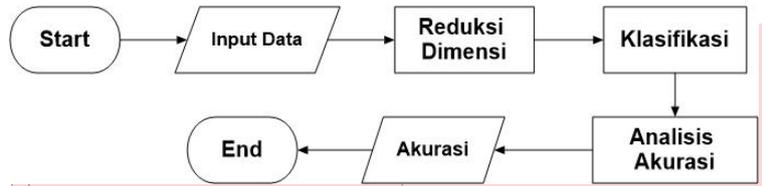
$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} * 100\% \tag{19}$$

$$\text{Presisi} = \frac{TP}{FP+TP} * 100\% \tag{20}$$

$$\text{Recall} = \frac{TP}{FN+TP} * 100\% \tag{21}$$

3. PERANCANGAN SISTEM

Pada penelitian ini, telah dibuat sistem penelitian yang akan dijalankan agar penelitian ini dapat berjalan dengan lancar. Berikut adalah gambaran umum dari sistem tersebut:



Gambar 1 Flowchart Gambaran Umum Sistem

Berdasarkan gambar diatas dikarenakan data microarray yang sangat besar diperlukan langkah Reduksi Dimensi untuk mengurangi data yang tidak relevan menggunakan metode Wavelet Transform. Selanjutnya tahap klasifikasi untuk menyelesaikan persamaan linear atau persamaan matriks secara iteratif. Hasil dari klasifikasi adalah analisi akurasi yang selanjutnya akan keluar akurasi yang didapatkan dengan menggunakan metode wavelet transform dan Modified Back Propagation (MBP) Flechter Reeves.

4. PENGUJIAN DAN ANALISIS

4.1 Pengujian

Pada pengujian ini data yang digunakan adalah data kanker kolon dan leukimia baik data yang tidak direduksi maupun data yang sudah tereduksi menggunakan Wavelet Haar. Skema klasifikasi menggunakan Modified Backpropagation (MBP).

Tabel 2 Spesifikasi Data

Data	Jumlah Kelas	Jumlah Record	Jumlah Dimensi
Kolon	2	62 (40 Cancer vs 22 Normal)	2.001
Leukimia	2	38 (27 Cancer vs 11 Normal)	7.130

Proses pengujian menggunakan beberapa skenario klasifikasi diantaranya yaitu Wavelet Haar dengan pembagian reduksi dimensi 25%, 40%, 50% dari data asli dan Modified Backpropagation (MBP) Flechter Reeves untuk membandingkan hasilnya dengan hasil sistem yang diusulkan. Setiap skema klasifikasi dilakukan percobaan masing- masing sebanyak 10 kali percobaan untuk membangun model agar mendapatkan hasil arsitektur terbaik kemudian dibandingkan dengan hasil akurasi.

Pada proses pengujian ini dilakukan perbandingan pada Hidden Neuron. Perbandingan dilakukan menggunakan 2, 4, dan 6 Neuron. Hal ini dilakukan untuk mengetahui jumlah Hidden Neuron yang digunakan pada model sistem yang dihasilkan menggunakan backpropagation termodifikasi dengan algoritma Conjugate Gradient Fletcher Reeves.

Tabel 1 Hasil Perbandingan Hidden Neuron

Skema Klasifikasi	Hidden Neuron Layer	Akurasi (%)						Rata _ Rata(%)
		Kanker Kolon			Kanker Leukimia			
		500	800	1000	1782	2851	3564	
MBP FLECHTER REEVES	2	70,6	64,7	58,8	88,9	88,9	66,7	73,1
	4	52,9	52,9	47,1	88,9	77,8	88,9	68,08
	6	64,7	58,8	47,1	88,9	66,7	77,8	67,33

Berdasarkan Tabel 1 didapatkan jumlah neuron terbaik pada hidden layer adalah sebanyak 2 neuron dengan akurasi sebesar 70.6% pada data kanker kolon dengan dimensi 500 dan akurasi sebesar 88.9% pada data

kanker leukimia dengan dimensi 1782. Akurasi tertinggi dianggap sebagai arsitektur terbaik yang selanjutnya digunakan kembali untuk diuji pada setiap data Kanker.

Tabel 4 Rata – Rata Akurasi Setiap Skema Klasifikasi

Skema Klasifikasi	Akurasi (%)						Rata-Rata Akurasi (%)
	Data Kanker Kolon			Data Kanker Leukimia			
	500	800	1000	1782	2851	3564	
MBP (Golden Section)	64,7	64,7	64,7	88,9	77,8	88,9	74,95
MBP (Hybrid Bisection Cubic)	52,9	64,7	52,9	77,8	77,8	88,9	69,16
BP STANDART	41,2	52,9	64,7	88,9	77,8	88,9	69,06

Berdasarkan Tabel 4 rata – rata akurasi untuk setiap skema klasifikasi dapat dilihat bahwa skema klasifikasi *Modified Backpropagation Flechter Reeves* dengan metode *Line Search Golden Section Search* memiliki rata-rata akurasi tertinggi yaitu sebesar 74,95%.

Tabel 5 Rata – Rata Training Time Setiap Skema Klasifikasi

Skema Klasifikasi	Training Time (s)						Rata-Rata Training Time(s)
	Data Kanker Kolon			Data Kanker Leukimia			
	500	800	1000	1782	2851	3564	
MBP (<i>Golden Section</i>)	1,76	1,34	1,24	2,42	1,84	2,29	1,81
MBP (<i>Hybrid Bisection Cubic</i>)	2,83	2,83	2,53	2,70	3,17	2,71	2,79
BP Standart	3,05	1,84	1,81	2,98	3,19	3,07	2,66

Berdasarkan Tabel 5 rata – rata *training time* untuk setiap skema klasifikasi dapat dilihat bahwa skema klasifikasi *Modified Backpropagation Flechter Reeves* dengan metode *Line Search Golden Section Search* memiliki rata-rata *training time* lebih singkat yaitu sebesar 1,81 detik. Hal ini karena pencarian linear pada *Modified Backpropagation Line Search Golden Section Search* sederhana tanpa harus menghitung *gradient*.

Tabel 6 Rata – Rata Akurasi Setiap Data

Data	Dimensi	Rata-Rata Akurasi(%)
Kanker Kolon	500	52,9
	800	60,7
	1000	60,7
Kanker Leukimia	1782	85,2
	2851	77,8
	3564	88,9

Berdasarkan Tabel 6 rata – rata akurasi untuk setiap jenis data dapat dilihat bahwa untuk jenis data Kanker Leukimia dengan dimensi 3564 memiliki rata-rata akurasi tertinggi yaitu sebesar 88,9%. Hal ini karena jumlah record kanker dan dimensi yang dimiliki oleh Kanker Kolon lebih banyak dibandingkan dengan data yang lainnya.

4.2 Analisis

Pada *Backpropagation* termodifikasi dengan *Conjugate Gradient Fletcher Reeves*, menghasilkan performansi sistem yang baik semakin banyak jumlah record kanker dan atribut yang diambil maka hasil akurasi yang didapatkan akan lebih baik. Hal ini dapat dibuktikan dari hasil rata rata akurasi yang didapatkan oleh jenis data kanker kolon. Dari hasil pengujian, skema *Wavelet Haar+Modified Backpropagation Line Search Golden Section Search* memiliki waktu *training* tercepat dari pada skema klasifikasi yang lainnya. *Wavelet Haar+Modified Backpropagation Line Search Golden Section Search* menghasilkan rentang waktu sekitar 1,2 – 2,4 detik untuk pelatihan masing – masing data.

Pada Tugas Akhir ini penulis merekomendasikan skema klasifikasi menggunakan *Wavelet Haar* dan *Modified Backpropagation Flechter Reeves* dengan teknik *line search golden section search* karena waktu

training yg dihasilkan sangat singkat dengan rata - rata waktu 1,81 detik. Jika dilihat dari akurasi yang didapatkan, skema klasifikasi dari *Wavelet Haar+Modified Backpropagation Flechter Reeves* dengan teknik *line search golden section search* memiliki akurasi yang tertinggi dari semua skema klasifikasi yaitu pada jenis data Kanker Leukimia dengan akurasi 88,9%.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan pengujian dan analisis yang telah dilakukan pada Tugas Akhir ini maka diperoleh beberapa hasil dan kesimpulan sebagai berikut :

Metode *Backpropagation* termodifikasi dan reduksi menggunakan *Wavelet Haar* dapat diimplementasikan dengan baik. Terlihat dari hasil pengujian *Wavelet Haar* dan *Modified Backpropagation Line Search Golden Section Search* menghasilkan rentang waktu sekitar 1,2 – 2,4 detik untuk pelatihan masing – masing data. Hal ini dapat menunjukkan bahwa *Wavelet Haar* dan *Modified Backpropagation Flechter Reeves* dapat menghasilkan waktu yang sangat singkat untuk pelatihan. Waktu yang dihasilkan paling minimum adalah 1,2 detik untuk pelatihan kanker kolon berdimensi 1000 dengan menggunakan teknik *line search golden section search*. Hal ini dikarenakan adanya pengurangan dimensi oleh *Wavelet Haar* sehingga waktu pelatihan yang dihasilkan lebih singkat. Jika dilihat dari akurasi, hasil pengujian skema klasifikasi *Wavelet Haar* dan *Modified Backpropagation Flechter Reeves* mendapatkan rata-rata akurasi sebesar 88,9% untuk teknik *line search Golden Section*. Untuk dapat menentukan *line search* dengan hasil terbaik, *Golden Section Search* merupakan *line search* yang terbaik. Selain itu Penggunaan *line search Golden Section Search* direkomendasikan oleh penulis. Disamping itu untuk *line search Hybrid Bisection Cubic Search* memberikan kesimpulan bahwa akurasi yang didapatkan dari hasil pelatihan oleh *line search Hybrid Bisection Cubic Search* tidak melebihi dari akurasi yang dihasilkan oleh *line search Golden Section Search*. Hal ini dikarenakan pada *Hybrid Bisection-Cubic Search* memiliki perhitungan yang banyak pada setiap langkahnya. Untuk dapat menentukan jumlah dimensi terbaik yang digunakan dalam menentukan *training time* dan akurasi, jumlah dimensi terbesar menghasilkan akurasi dan *training time* terbaik. Dalam penelitian ini data kanker leukimia yang mempunyai atribut lebih besar memiliki akurasi terbaik 88,9%, dibandingkan dengan data kolon yang memiliki akurasi terbaik 64,7%. Sedangkan dalam menentukan hasil terbaik *training time* data kanker kolon yang mempunyai atribut lebih kecil memiliki hasil terbaik 1,24 detik dibandingkan dengan data kanker leukimia yang memiliki hasil terbaik 1,8 detik. *Backpropagation standart* (BP) membutuhkan waktu yang relatif lama yaitu 1001 dengan mencapai batas maksimum epoch selain itu waktu *training* yg diperoleh memiliki rentang 1,3– 3,07 detik. Hal ini karena *Backpropagation Standart* (BP) menggunakan learning rate yang selalu tetap pada proses pelatihan dan update bobot hanya dipengaruhi faktor kesalahan atau error. Selain itu, tidak adanya peran conjugate gradient dan *line search* untuk meminimumkan arah pencarian dan parameter α dapat merepresentasikan besar langkah untuk memperbarui bobot yang optimal dan pada parameter α dan β yang dapat menghindari konvergensi lokal.

5.2 Saran

Saran yang dapat disampaikan untuk pengembangan Tugas Akhir ini adalah:

Diharapkan untuk kedepannya menggunakan proporsi data yang seimbang antara data normal dan data kanker. Menggunakan data *training* dan data *testing* yang lebih banyak untuk menghasilkan model arsitektur jaringan dengan baik

6. DAFTAR PUSTAKA

- [1] Nurfalaha, 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.
- [2] B. Anita, "Classification and clustering using intelligent techniques: application to microarray cancer data," *Master Thesis*, 2013.
- [3] P. A. Morettin, "From Fourier to Wavelet Analysis of Time Series," *Waves and Wavelets*, 2004.
- [4] P. S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*, Edinburgh, Napier University, 2002.
- [5] Adiwijaya, Maharani, M., Dewi, B.K., Yulianto, F.A. and Purnama, B., 2013. digital image compression using graph coloring quantization based on wavelet-SVD. In *Journal of Physics: Conference Series* (Vol. 423, No. 1, p. 012019). IOP Publishing..

- [6] A. A. Rohmawati, Transformasi Wavelet Diskret dan Partial Least Squares dalam Pemodelan Kalibrasi serta Implementasinya dengan OSS-R/Tugas Akhir, Surabaya, 2010.
- [7] P. Griffiths dan J. A. Hasseth, Fourier Transform Infrared Spectrometry, New York, John Wiley & Sons, 2007.
- [8] Percival, An Introduction to the Wavelet Analysis of Time Series, Seattle, Washington University, 2005.
- [9] S. Sunaryo, Model Kalibrasi dengan Transformasi Wavelet sebagai Metode Pra-pemrosesan [disertasi], Bogor: Sekolah Pascasarjana, Institut Pertanian Bogor, 2005.
- [10] Adiwijaya, N. W. U, Y. A dan M. F. D, "Teknik Line Search Pada Conjugate Gradient Dalam Proses Pelatihan Backpropagation Untuk Peramalan Temperatur Dan Kelembaban Udara," *Konferensi Nasional Sistem Informasi Makalah Nomor: KNSI-192*, 2013.
- [11] S. Kusumadewi, Membangun Jaringan Syaraf Tiruan Menggunakan Matlab & Excel Link, Yogyakarta: Graha Ilmu, 2004.
- [12] Adiwijaya, Wisesty, U.N. and Nhita, F., 2014. STUDY OF LINE SEARCH TECHNIQUES ON THE MODIFIED BACKPROPAGATION FOR FORECASTING OF WEATHER DATA IN INDONESIA. *Far East Journal of Mathematical Sciences*, 86(2), p.139.
- [13] Adiwijaya, Wisesty, U.N., Wirayuda, T.A.B., Baizal, Z.K.A. and Haryoko, U., 2013. AN IMPROVEMENT OF BACKPROPAGATION PERFORMANCE BY USING CONJUGATE GRADIENT ON FORECASTING OF AIR TEMPERATURE AND HUMIDITY IN INDONESIA. *Far East J. Math. Sci*, 1, pp.57-67.
- [14] A. V. Devi, D. Devaraj dan M. Venkatesulu, "Gene Expression Data Classification using Support Vector Machine and Mutual Information-based Gene Selection," *Far East Journal of Electronics and Communications*, 2015.