

DETEKSI KANKER BERDASARKAN KLASIFIKASI MICROARRAY DATA MENGGUNAKAN PRINCIPAL COMPONENT ANALYSIS DAN BACKPROPAGATION TERMODIFIKASI DENGAN CONJUGATE GRADIENT FLETCHER REEVES

Della Alfarydy Akbar¹, Adiwijaya², Danang Triantoro³

Ilmu Komputasi, Fakultas Informatika, Universitas Telkom

¹dellaalfarydy@gmail.com

²adiwijaya@telkomuniversity.ac.id

³danangtri@telkomuniversity.ac.id

Abstract

Cancer is the leading cause in death. In each year of cancer is expected to continue to increase due to the unhealthiness of life-style patterns. In recent decades microarray plays an important role in the diagnosis of cancer. Microarray is a technology that can store thousands of genes taken in some human cells as well. Thus the microarray has a very large data dimensions therefore, to improve the accuracy of cancer diagnosis is then carried out by means of reducing dimensions by using Principal Component Analysis (PCA) and Modified Back Propagation (MBP). MBP is a modification of Standard Backpropagation (BP) which apply the method Artificial Neural Network (ANN) algorithm Fletcher Reeves Conjugate Gradient to increase the training process. The test results were obtained from the Principal Component Analysis (PCA) and have modified Backpropagation performance span of about 2-3 seconds to train. The result minimum time is 2:21 seconds to training the Central Nervous using line search Hybrid Bisection Cubic technique. When viewed from accuracy, the test results classification scheme PCA + MBP get an average accuracy of 82.15% for the technique of line search of the Golden Section.

Keyword : *Cancer, Microarray, Principal Component Analysis (PCA), Modified Backpropagation (MBP), Conjugate Gradient Fletcher Reeves, Artificial Neural Network.*

Abstrak

Kanker merupakan penyebab utama dalam kematian. Dalam setiap tahun diperkirakan kanker akan terus meningkat karena tidak sehatnya gaya pola hidup. Pada beberapa dekade terakhir microarray berperan penting dalam diagnosis kanker. Microarray merupakan teknologi yang dapat menyimpan ribuan gen yang diambil dalam beberapa sel manusia sekaligus. Maka dari itu microarray memiliki dimensi data yang sangat besar oleh karena itu, untuk meningkatkan akurasi diagnosis kanker tersebut maka dilakukan dengan cara pengurangan dimensi dengan menggunakan *Principal Component Analysis* (PCA) dan *Modified Back Propagation* (MBP). MBP merupakan modifikasi dari Backpropagation Standart (BP) yang menerapkan metode Jaringan Syaraf Tiruan (JST) dengan algoritma *Conjugate Gradient Fletcher Reeves* untuk mempercepat proses pelatihan. Hasil pengujian yang diperoleh dari *Principal Component Analysis* (PCA) dan *Backpropagation* termodifikasi memiliki performansi rentang waktu sekitar 2 – 3 detik untuk pelatihan. Waktu yang dihasilkan paling minimum adalah 2.21 detik untuk pelatihan *Central Nervous* dengan menggunakan teknik *line search Hybrid Bisection Cubic*. Jika dilihat dari akurasi, hasil pengujian skema klasifikasi PCA+MBP mendapatkan rata-rata akurasi sebesar 82.15% untuk teknik *line search Golden Section*.

Kata Kunci : *Kanker, Microarray, Principal Component Analysis (PCA), Modified Backpropagation (MBP), Conjugate Gradient Fletcher Reeves, Jaringan Syaraf Tiruan.*

I. PENDAHULUAN

Menurut laporan data oleh Organisasi Kesehatan Dunia (WHO), Kanker adalah penyebab utama kematian di seluruh dunia, yaitu sekitar 8,2 juta kematian pada tahun 2012 dan diperkirakan akan meningkat setiap tahun karena tidak sehatnya gaya hidup. Kanker adalah istilah yang digunakan untuk penyakit di mana sel-sel abnormal membelah tanpa kontrol dan mampu menyerang sel jaringan lain. Sel-sel kanker dapat menyebar ke bagian lain dari tubuh melalui darah.

Dalam beberapa dekade terakhir, microarray data menjadi peran penting dalam mendiagnosis kanker dan untuk meningkatkan akurasi diagnosis kanker tersebut maka dibandingkanlah dengan teknik tradisional. *Microarray* dapat digunakan untuk melihat tingkat *gene expression* dalam suatu sampel sel tertentu untuk menganalisis ribuan gen secara bersamaan.

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.

Berdasarkan paper *Cancer Detection Based On Microarray Data Classification Using Principal Component Analysis (PCA) And Modified Back Propagation* yang diteliti oleh Adiyasa Nurfalah, Adiwijaya dan Arie Ardiyanti Suryani [5], mengusulkan kerangka kerja baru untuk mendeteksi kanker berdasarkan *microarray* data menggunakan perpaduan analisis komponen utama PCA dan metode *Conjugate Gradient BackPropagation* (MBP). MBP adalah modifikasi dari standar *back propagation* dengan mengimplementasikan algoritma *conjugate gradient* pada *backpropagation training*.

II. TINJAUAN PUSTAKA

A. PCA (Principal Component Analysis)

Principal Component Analysis (PCA) adalah sebuah teknik yang digunakan untuk menyederhanakan atau mereduksi dimensi, dimana sejumlah dimensi yang saling berkorelasi diubah menjadi sekumpulan dimensi yang tidak saling berkorelasi dengan ukuran data yang lebih kecil. PCA biasanya diaplikasikan pada image data dan data ekspresi gen.

Keluaran dari PCA itu sendiri nantinya menghasilkan sekumpulan dimensi baru yang biasa dinamakan *Principal Component* (PC). Untuk setiap PC merupakan kombinasi linier dari dimensi asli pada data asli [5].

Langkah pertama pada PCA adalah mencari data X^* dengan teknik zero-mean yaitu dengan mengurangi semua nilai $X_{i,j}$ pada matriks X dengan nilai rata – rata matriks tersebut

$$X^* = X - \bar{X} \quad (1)$$

Selanjutnya adalah dengan mencari kovarian dari matriks X^* , kovarian disimbolkan dengan C_x

$$C_x = \frac{1}{n-1} X^{*T} X^* \quad (2)$$

C_x adalah matriks kovariansi $n \times n$, dan n yaitu jumlah pengamatan. Pada matriks C_x hitung nilai eigen λ dan vector eigen v , dengan persamaan :

$$|C_x - \lambda I| = 0 \quad (3)$$

dan

$$(X - \lambda I) \mathbf{v} = 0 \quad (4)$$

vektor eigen yang didapatkan merupakan komponen utama untuk membentuk variable baru [11]. Variable baru tersebut merupakan perkalian antara vector eigen $\mathbf{v}$ terhadap matriks X^* .

Sedangkan persentase variansi yang dapat dijelaskan oleh variable baru tergantung kontribusi λ_i dari masing-masing nilai eigen [11], yang dihitung dengan rumus:

$$\lambda_i = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j} \times 100\% \quad (5)$$

Sedangkan penentuan jumlah variable baru yang digunakan tergantung kepada persentase kontribusi kumulatif variansi λ_i , dihitung dengan rumus :

$$K = \frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^n \lambda_i} \times 100\% \quad (6)$$

B. Jaringan Syaraf Tiruan

Jaringan Saraf Tiruan (*Artificial Neural Network*) atau disingkat JST adalah suatu algoritma proses pembelajaran yang meniru cara kerja jaringan saraf makhluk hidup. ANN terdiri dari beberapa kelompok unit pemroses kecil yang disebut dengan neuron [9].

Pada referensi [9], Pada umumnya arsitektur dari ANN terbagi menjadi empat, namun ada dua jenis arsitektur yang sering digunakan yaitu :

a. Single Layer Network

Arsitektur jst ini hanya terdiri dari *input* layer dengan node sumber yang terproyeksi ke *output* layer, tetapi tidak sebaliknya. Dengan kata lain, jaringan ini adalah jaringan jenis umpan maju (*feedforward*).

b. Multi Layer Network

Arsitektur layer tunggal dapat dikembangkan menjadi layer jamak dengan menambahkan satu atau lebih lapisan tersembunyi (*hidden layer*).

C. Algoritma Pelatihan Propagansi Mundur dengan Conjugate Gradient Fletcher Reeves

Algoritma conjugate gradient (CG) merupakan algoritma pencarian yang arahnya berdasarkan pada arah konjugasi. Secara umum algoritma ini lebih cepat *convergen* daripada metode penurunan tercepat. CG menggunakan vektor tidak nol yang orthogonal dan bebas linier [5]. Dua vektor $\mathbf{d}_i$ dan $\mathbf{d}_j$ dikatakan orthogonal jika perkalian dalamnya bernilai nol [5]:

$$\mathbf{d}_i \cdot \mathbf{d}_j = 0 \quad (7)$$

Peran CG dalam algoritma *backpropagation* yaitu untuk memperkecil *error* yang terdapat pada bobot-bobot yang terhubung diantara neuron sehingga dapat dituliskan fungsi objektif untuk meminimalkan *error*, yaitu:

$$E = \frac{1}{2N} \sum_i \sum_j (t_{ij} - \mathbf{w}_{ij} \cdot \mathbf{d}_j)^2 \quad (8)$$

Conjugate gradient merupakan metode untuk meminimasi fungsi turunan dengan menghitung pendekatan $\mathbf{w}_{k+1}$ secara iteratif berdasarkan :

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \alpha_k \mathbf{d}_k \quad (9)$$

$$\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \beta_k \mathbf{d}_k \quad (10)$$

Keterangan : α dan β = parameter momentum (untuk menghindari konvergensi lokal).

Sehingga algoritma pelatihannya menjadi sebagai berikut [5]:

1. Definisikan masalah, misalkan matriks masukan ($\mathbf{P}$) dan matrik target ($\mathbf{T}$).

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 penghentian belum terpenuhi, lakukan langkah proses 2-13

• **Propagasi maju**

4. Hitung keluaran dari *hidden layer* $z_j (j=1,2,3,...,p)$

$$y_{net_j} = v_{j0} + \sum_{i=1}^n x_i v_{ji} \quad (11)$$

$$y_j = f(y_{net_j}) = \frac{1}{1 + \exp^{-y_{net_j}}} \quad (12)$$

5. Lanjutkan dengan menghitung keluaran dari *output layer* $r_j (j=1,2,...,m)$,

$$z_{net_k} = w_{k0} + \sum_{j=1}^p z_j w_{kj} \quad (13)$$

$$z_k = f(z_{net_k}) = \frac{1}{1 + \exp^{-z_{net_k}}} \quad (14)$$

• **Propagasi mundur**

6. Hitung faktor kesalahan di unit *output* berdasarkan perbedaan (*error*) nilai aktual dan nilai prediksi (*output* dari unit *output*)

$$\delta_{r_j} = (r_j - \hat{r}_j) f'(z_{net_j}) = (r_j - \hat{r}_j) z_j (1 - z_j) \quad (15)$$

7. Hitung faktor kesalahan di unit *hidden layer* berdasarkan faktor sebelumnya

$$\delta_{z_j} = \sum_{k=1}^m \delta_{r_k} w_{kj} \quad (16)$$

$$\delta_{z_j} = \delta_{z_{net_j}} f'(z_{net_j}) = \delta_{z_{net_j}} z_j (1 - z_j) \quad (17)$$

8. Hitung gradien di unit *output* dari fungsi objektif yang sudah ditentukan

$$\delta_{r_{j+1}} = \frac{1}{N} \sum_{j=1}^N \delta_{r_j} \quad (18)$$

9. Hitung gradien di unit *hidden layer*

$$\delta_{z_{j+1}} = \frac{1}{N} \sum_{k=1}^p \delta_{z_k} w_{kj} \quad (19)$$

10. Hitung parameter β untuk semua neuron di unit *hidden layer* dan unit *output*. Parameter β untuk *Conjugate Gradient Fletcher Reeves* dapat dihitung dengan persamaan:

$$\beta_{k+1} = \frac{\delta_{z_{j+1}}^2}{\delta_{z_j}^2 + \delta_{z_{j+1}}^2} \quad (20)$$

dimana : $\beta_{k+1} = \beta$ nilai parameter pada iterasi saat ini

g_{k+1} = gradient pada iterasi saat ini

g_k = gradient pada iterasi sebelumnya

11. Hitung *direction* untuk semua *neuron* di unit *hidden layer* dan unit *output*.

$$d_{t+1} = -g_{t+1} + \beta_t d_t \quad (21)$$

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 *direction* awal: $d_t = -g_t$

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

- **Perubahan bobot**

13. *Update* bobot dilakukan dengan cara sebagai berikut:

$$w_{t+1} = w_t + \alpha_{t+1} d_{t+1} \quad (22)$$

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

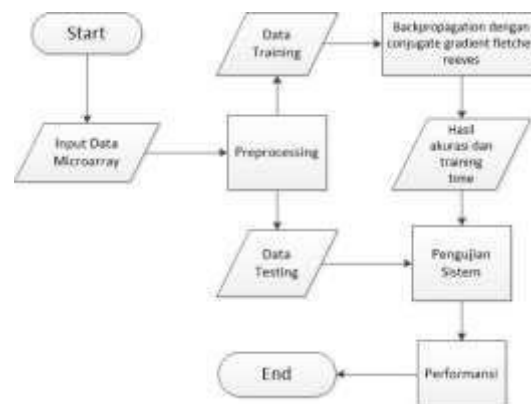
Parameter α disini dapat dicari dengan menggunakan teknik *line search* sehingga dapat meminimumkan kinerja selama arah pencarian, karena parameter α merepresentasikan besar langkah yang diambil untuk setiap *direction* [6]. Adapun Teknik *Line Search* yang digunakan

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.

III. METODOLOGI PENELITIAN

A. Desain Sistem

Desain sistem dengan *Backpropagation* termodifikasi menggunakan algoritma pelatihan *Conjugate Gradient Fletcher Reeves* dapat dilihat pada gambar 1.



Gambar 1 Desain Sistem

Berdasarkan Gambar 1 secara umum dijelaskan bahwa tahap pertama yaitu data microarray masuk ke dalam tahap *preprocessing*. Tujuan dari tahap *preprocessing* adalah agar data dapat diproses nantinya. Data hasil *preprocessing* kemudian di analisa menggunakan metode JST *backpropagation* termodifikasi dengan

algoritma *Conjugate Gradient Fletcher Reeves*. Setelah hasil dari analisa tersebut didapatkan maka langkah terakhir adalah menghitung performansinya.

B. Data

Data yang digunakan adalah Kanker Payudara, Tumor Usus, Kanker Paru-paru, Leukimia MLL, Kanker Ovarian, Kanker Prostat, dan Kanker Syaraf Pusat yg diambil dari *Kent-Ridge Biomedical Data Repository*, data yang digunakan pada skenario 1 menggunakan PCA, sedangkan pada skenario 2 tanpa menggunakan PCA. Proporsi data pada masing-masing skenario dapat dilihat pada Tabel 2 dan 3.

Tabel 2. Proporsi Data Training tanpa PCA

Data	Jumlah Kelas	Jumlah Record	Jumlah Atribut
Breast Cancer	2	78 (34 relapse vs 44 non-relapse)	24482
Central Nervous	2	48 (21 cancer vs 27 normal)	7130
Colon Tumor	2	50 (17 cancer vs 33 normal)	2001
Lung Cancer	2	32 (16 mesothelioma vs 16 ADCA)	12534
MLL Leukimia	3	57 (20 ALL vs 17 MLL vs 20 AML)	12583
Ovarian Cancer	2	202 (101 cancer vs 91 normal)	15155
Prostate Cancer	2	102 (52 Tumor vs 50 normal)	12601

Tabel 3. Proporsi Data Training dengan PCA

Data	Jumlah Kelas	Jumlah Record	Jumlah Atribut
Breast Cancer	2	78 (34 relapse vs 44 non-relapse)	96
Central Nervous	2	48 (21 cancer vs 27 normal)	59
Colon Tumor	2	50 (17 cancer vs 33 normal)	61
Lung Cancer	2	32 (16 mesothelioma vs 16 ADCA)	148
MLL Leukimia	3	57 (20 ALL vs 17 MLL vs 20 AML)	71
Ovarian Cancer	2	202 (101 cancer vs 91 normal)	252
Prostate Cancer	2	102 (52 Tumor vs 50 normal)	135

IV. HASIL DAN DISKUSI

Proses pengujian menggunakan beberapa skenario klasifikasi diantaranya yaitu PCA dan *Modified Backpropagation* (MBP), MPB (Tanpa PCA), dan BP (*Backpropagation Standart*) untuk membandingkan hasilnya dengan hasil sistem yang diusulkan. Setiap skema klasifikasi dilakukan percobaan masing-masing sebanyak 5 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 4. Hasil Perbandingan Neuron Hidden Layer

Skema Klasifikasi	Neuron Hidden Layer	Akurasi Ovarian Cancer (%)	Akurasi Breast Cancer (%)	Akurasi Lung Cancer (%)	Akurasi Prostate Cancer (%)	Rata – Rata (%)
MBP (Golden Section)	2	98.03	73.68	90.60	82.35	86.16
	4	96.04	47.36	85.23	76.47	76.27
	6	94.11	63.15	81.87	73.52	78.16
MBP (Hybrid Bisection)	2	96.07	63.15	87.91	97.05	86.04
	4	94.11	63.15	90.60	67.64	78.87
	6	86.77	57.89	73.03	76.47	73.54

Berdasarkan Tabel 4 didapatkan jumlah neuron terbaik pada hidden layer adalah sebanyak 2 neuron dengan akurasi sebesar 98.03%. Akurasi tertinggi dianggap sebagai arsitektur terbaik yang selanjutnya digunakan kembali untuk diuji pada setiap data Kanker.

Pada proses pengujian, proporsi data yang digunakan sebesar 80% untuk data *Training* dan 20% untuk data *Testing*. pembagian data mengacu pada *default* data yang didapat dari *Kent Ridge Biomedical Data Repository*.

Tabel 5. Hasil Pengujian

Jenis Data	Skema Klasifikasi	Akurasi (%)	Training Time (s)
Breast Cancer	PCA+MBP (Golden Section)	73.68	2.63
	PCA+MBP (Hybrid Bisection Cubic)	63.15	3.54
	PCA+BP	63.15	4.40
	MBP (Golden Section)	73.68	6.59
	MBP (Hybrid Bisection Cubic)	63.15	3.97
	BP	52.63	23.69
Central Nervous	PCA+MBP (Golden Section)	91.66	2.38
	PCA+MBP (Hybrid Bisection Cubic)	91.66	2.21
	PCA+BP	83.33	2.70
	MBP (Golden Section)	83.33	2.79
	MBP (Hybrid Bisection Cubic)	83.33	3.93
	BP	75	14.66
Colon Tumor	PCA+MBP (Golden Section)	75	2.62
	PCA+MBP (Hybrid Bisection Cubic)	75	2.56
	PCA+BP	83.33	6.66
	MBP (Golden Section)	75	2.70
	MBP (Hybrid Bisection Cubic)	66.67	2.70
	BP	75	9.33
Lung Cancer	PCA+MBP (Golden Section)	90.60	2.58
	PCA+MBP (Hybrid Bisection Cubic)	87.91	2.47
	PCA+BP	89.93	2.80
	MBP (Golden Section)	92.61	3.33
	MBP (Hybrid Bisection Cubic)	89.93	3.65
	BP	80.53	17.76
MLL Leukimia	PCA+MBP (Golden Section)	66.68	2.77
	PCA+MBP (Hybrid Bisection Cubic)	60	3.03
	PCA+BP	53.33	10.54
	MBP (Golden Section)	80	4.27

	MBP (Hybrid Bisection Cubic)	80	3.68
	BP	86.67	21.98
Ovarian Cancer	PCA+MBP (Golden Section)	98.03	2.76
	PCA+MBP (Hybrid Bisection Cubic)	96.07	2.68
	PCA+BP	94.11	5.94
	MBP (Golden Section)	96.07	7.35
	MBP (Hybrid Bisection Cubic)	96.07	8.24
	BP	91.97	68.77
Prostate Cancer	PCA+MBP (Golden Section)	82.35	2.44
	PCA+MBP (Hybrid Bisection Cubic)	73.53	2.74
	PCA+BP	73.52	10.77
	MBP (Golden Section)	97.05	4.76
	MBP (Hybrid Bisection Cubic)	97.05	4.89
	BP	85.29	13.69

Berdasarkan Tabel 4 *Backpropagation* termodifikasi dengan *Conjugate Gradient Fletcher Reeves*, menghasilkan performansi sistem yang baik. Semakin banyak proporsi data 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 *Ovarian* yaitu sebesar 98.03%. Dari hasil pengujian, skema klasifikasi PCA+MBP memiliki waktu training tercepat daripada skema klasifikasi yang lainnya. PCA+MBP menghasilkan rentang waktu sekitar 2 – 3 detik untuk pelatihan masing – masing data.

Pada makalah ini direkomendasikan skema klasifikasi menggunakan PCA+MBP dengan teknik *line search golden section search* karena waktu training yg dihasilkan sangat singkat dengan rata - rata waktu 2.60 detik. Jika dilihat dari akurasi yang didapatkan, skema klasifikasi dari PCA+MBP dengan teknik *line search golden section search* memiliki akurasi yang tertinggi dari semua skema klasifikasi yaitu pada jenis data *Ovarian Cancer*. Pada saat pengujian hasil akurasi pada jenis data *Ovarian* dengan menggunakan skema klasifikasi *Backpropagation Standart* (BP), akurasi mencapai 100% namun jika dilihat dari waktu training yg didapatkan cukup lama. Maka dari itu penulis memilih hasil akurasi tertinggi yaitu untuk skema klasifikasi dari PCA+MBP dengan teknik *line search golden section search* pada jenis data *Ovarian Cancer*

V. KESIMPULAN

Backpropagation termodifikasi dengan *Conjugate Gradient Fletcher Reeves*, menghasilkan performansi sistem yang baik. Hasil pengujian PCA+MBP memiliki rentang waktu sekitar 2 – 3 detik untuk pelatihan masing – masing data. Hal ini dapat menunjukan bahwa PCA+MBP dapat menghasilkan waktu yang sangat singkat untuk pelatihan. Waktu yang dihasilkan paling minimum adalah 2.21 detik untuk pelatihan *Central Nervous* dengan menggunakan teknik *line search Hybrid Bisection Cubic*. Hal ini dikarenakan adanya pengurangan dimensi oleh PCA sehingga waktu pelatihan yang dihasilkan lebih singkat. Jika dilihat dari akurasi, hasil pengujian skema klasifikasi PCA+MBP mendapatkan rata-rata akurasi sebesar 82.57% untuk teknik *line search Golden Section*.

DAFTAR PUSTAKA

- [1].Vanitha, C.D.A., Devaraj, D. and Venkatesulu, M., 2015. *Gene Expression Data Classification Using Support Vector Machine and Mutual Information-based Gene Selection*. *Procedia Computer Science*, 47, pp.13-21.
- [2].Siang, T.C., Soon, T.W., Kasim, S., Mohamad, M.S., Howe, C.W., Deris, S., Zakaria, Z., Shah, Z.A. and Ibrahim, Z., 2015. *A Review of Cancer Classification Software for Gene Expression Data*. *International Journal of Bio-Science and Bio-Technology*, 7(4), pp.89-108.
- [3].Bhasker, J. and Samad, T., 1991. *The clique-partitioning problem*. *Computers & Mathematics with Applications*, 22(6), pp.1-11.

- [4].Kochenberger, G., Glover, F., Alidaee, B. and Wang, H., 2005. *Clustering of microarray data via clique partitioning. Journal of Combinatorial Optimization*,10(1), pp.77-92.
- [5].Nurfalah, A. Adiwijaya. and Suryani, 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.
- [6].Adiwijaya, U.N., Wisesty, A., Yunanto. and , D.M, Tampubolon. 2013. Teknik *Line Search* Pada *Conjugate Gradient* Dalam Proses Pelatihan *Backpropagation* Untuk Peramalan Temperatur Dan Kelembaban Udara. Konferensi Nasional Sistem Informasi Makalah Nomor: KNSI-192 ,.
- [7].Adiwijaya, U.N., Wisesty, 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 Journal of Mathematical Sciences (FJMS)*, (Part I), pp.57-67.
- [8].Astuti, L.W. and Handayani, T., 2015, July. KLASIFIKASI MRI OTAK MENGGUNAKAN JARINGAN SYARAF TIRUAN BERDASARKAN DATA WAVELET YANG DIREDUKSI DENGAN NCMF. In *Seminar Nasional Informatika (SEMNASIF)* (Vol. 1, No. 3).
- [9].Suyanto, 2008, "Soft Computing: Membangun Mesin Ber-IQ Tinggi", Bandung : Informatika.
- [10].Shewchuk, Jonathan Richard, "An Introduction to the Conjugate Gradient Method Without the Agonizing Pain", Pittsburgh, School of Computer Science Carnegie Mellon University, 1994.
- [11].Susetyoko, R. and Purwantini, E., Teknik Reduksi Dimensi Menggunakan Komponen Utama Data Partisi Pada Pengklasifikasian Data Berdimensi Tinggi dengan Ukuran Sampel Kecil. *Dimensi*, 1, p.500.
- [12].Naes T, Isaksoon T, Fearn T, Davies T. 2002. A User Friendly Guide to Multivariate Calibration and Classification. NIR publication, UK.
- [13].Kusumadewi, Sri.,2004. "Membangun Jaringan Syaraf Tiruan Menggunakan Matlab & Excel Link", Yogyakarta : Graha Ilmu.
- [14].Adiwijaya, U. N. Wisesty and F. Nhita, Study of Line Search Techniques on the Modified Back propagation for Forecasting of Weather Data in Indonesia, *FarEast J. Math. Sci. (FJMS)* 86(2) (2014), 139-148.
- [15].Adiwijaya, 2014, "Aplikasi Matriks dan Ruang Vektor", Yogyakarta : Graha Ilmu