

LAPORAN PENELITIAN
SKEMA LANJUTAN



ANALISIS KLASIFIKASI DATA *BIOINFORMATICS*
MENGGUNAKAN *ARTIFICIAL NEURAL NETWORK*
Studi Kasus : Data Sel Kanker Payudara

PUSAT STUDI : SAINS DAN TEKNOLOGI

Tim:

Arif Setiawan, S.Kom, M.Cs 0623018201 Ketua Tim Pengusul

Pratomo Setiaji, M.Kom 0619067802 Anggota Tim Pengusul

FAKULTAS TEKNIK
UNIVERSITAS MURIA KUDUS
TAHUN 2017

PENGESAHAN LAPORAN PENELITIAN

1. **Judul Penelitian** : Analisis Klasifikasi Data *Bioinformatics*
Menggunakan *Artificial Neural Network*
Studi Kasus : Data Sel Kanker Payudara
2. **Skema Penelitian** : **Penelitian Lanjutan**
3. **Peneliti/pelaksana**
 - Nama Lengkap : Arif Setiawan, S.Kom, M.Cs
 - NIDN : 0623018201
 - Jabatan Fungsional : Lektor
 - Program Studi : Sistem Informasi
 - Anggota 1**
 - Nama Lengkap : Pratomo Setiaji, M.Kom
 - NIDN : 0619067802
 - Jabatan Fungsional : Lektor
 - Program Studi : Sistem Informasi
 - Anggota Teknisi/laboran** : 1 Orang
 - Anggota Mahasiswa** : 2 Orang
4. **Institusi Mitra (jika ada)**
 - Nama Mitra : -
 - Nama institusi mitra : -
 - Alamat : -
 - Nama Penanggungjawab : -
 - Jangka Waktu Pelaksanaan** : 6 Bulan
5. **Anggaran Biaya**
 - a. Biaya APB UMK : Rp. 4.000.000,-
 - b. Biaya sumber lain : Rp. 0
 - c. Biaya Keseluruhan : Rp. 0

Kudus, 13 November 2017

Mengesahui,
Fakultas Teknik

Fakultas Teknik
Arif Setiawan, ST, MT
NIP. 0610701000001141

Ketua Peneliti

Arif Setiawan, S.Kom, M.Cs
NIS. 0610701000001180

Menyetujui,
Ketua Lembaga Penelitian UMK

Lembaga Penelitian UMK
(Dr. M. Anik Indaryani, MS)
NIP/NIK/0610702010101010

DAFTAR ISI

	Halaman
Halaman Judul	1
Halaman Pengesahan	2
Daftar Isi	4
Daftar Gambar	5
Abstrak	6
BAB I PENDAHULUAN.....	7
BAB II LANDASAN TEORI.....	9
BAB III METODE PENELITIAN	12
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	13
BAB V KESIMPULAN DAN SARAN	16
DAFTAR PUSTAKA	17

DAFTAR GAMBAR

	Halaman
Gambar 3.1. Data Pelatihan DNA Manusia	13
Gambar 3.2. Data Pengujian DNA Manusia.....	14
Gambar 3.3. Representasi piksel pada citra digital	14
Gambar 3.4. Confusion Matrik Metode RF pada Proses Klasifikasi.....	15
Gambar 3.5. Beberapa pohon keputusan pada metode RF	15

Abstrak

Bioinformatics merupakan ilmu yang mempelajari penerapan teknik komputasi untuk mengelola dan menganalisis informasi hayati. Bidang ini mencakup penerapan metode – metode matematika, statistika, dan informatika untuk memecahkan masalah – masalah biologi, terutama yang terkait dengan data *bioinformatics*. Salah satu bidang ilmu *bioinformatics* adalah penggunaan *machine learning* untuk klasifikasi data *bioinformatics*. Untuk bisa mengaplikasikan teknik-teknik *machine learning*, hal yang pertama harus ada adalah data. Pada penelitian ini kegiatan yang dilakukan adalah menganalisa data *bioinformatics* menggunakan klasifikasi, metode yang digunakan adalah *artificial neural network* yaitu *backpropagation*, dengan studi kasus sel kanker payudara. Hasil dari penelitian ini adalah melakukan analisa klasifikasi data *bioinformatics* sel kanker payudara untuk stadium ringan, sedang dan akut, dengan menggunakan 5 node simpul dengan bobot yang diacak secara random, untuk nilai learning = 0,3 dan momentum atau bias = 0.2 serta proses iterasi atau perulangan yang digunakan sebanyak 500 kali

Kata Kunci : *bioinformatics, machine learning, klasifikasi, bioinformatics, artificial neural network, sel kanker payudara*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bioinformatics merupakan ilmu terapan yang lahir dari perkembangan teknologi informasi di bidang molekular. Pembahasan di bidang *bioinformatics* ini tidak terlepas dari perkembangan biologi molekular modern, salah satunya peningkatan pemahaman manusia dalam bidang genomic yang terdapat dalam molekul DNA. *Bioinformatics* merupakan ilmu yang mempelajari penerapan teknik komputasi untuk mengelola dan menganalisis informasi hayati. Bidang ini mencakup penerapan metode-metode matematika, statistika, dan informatika untuk memecahkan masalah-masalah biologi, terutama yang terkait dengan penggunaan sekuens DNA dan asam amino.

Metode informatika yang digunakan dalam teknik komputasi *bioinformatics* diantaranya adalah *Machine Learning*. *Machine Learning* cabang dari Kecerdasan Buatan, adalah disiplin ilmu yang mencakup perancangan dan pengembangan algoritma yang memungkinkan komputer untuk mengembangkan perilaku yang didasarkan pada data empiris, seperti dari sensor data basis data. Sistem pembelajar dapat memanfaatkan contoh (data) untuk menangkap ciri yang diperlukan dari probabilitas yang mendasarinya (yang tidak diketahui). Data dapat dilihat sebagai contoh yang menggambarkan hubungan antara variabel yang diamati. Fokus besar penelitian pembelajaran mesin adalah bagaimana mengenali secara otomatis pola kompleks dan membuat keputusan cerdas berdasarkan data. Kesukarannya terjadi karena himpunan semua perilaku yang mungkin, dari semua masukan yang dimungkinkan, terlalu besar untuk diliput oleh himpunan contoh pengamatan (data pelatihan). Karena itu pembelajar harus merampatkan (generalisasi) perilaku dari contoh yang ada untuk menghasilkan keluaran yang berguna dalam kasus-kasus baru.

Untuk bisa mengaplikasikan teknik-teknik *machine learning*, hal yang pertama harus ada adalah data. Tanpa data maka algoritma *machine learning* tidak dapat bekerja. Data yang dimiliki biasanya dibagi menjadi 2, yaitu data *training* dan data *testing*. Data *training* digunakan untuk melatih algoritma, ringkasan data *testing* dipakai untuk mengetahui performa algoritma yang sudah dilatih sebelumnya ketika menemukan data baru yang belum pernah dilihat sebelumnya. Ini biasanya disebut dengan generalisasi. Hasil dari pelatihan tersebut bisa disebut dengan model.

Dari model tersebut bisa dilakukan prediksi yang biasanya dibedakan menjadi dua tergantung tipe keluarannya. Jika hasil prediksinya bersifat diskrit, maka proses itu disebut klasifikasi.

1.2 Perumusan masalah

Berdasarkan latar belakang di atas maka perumusan masalah yang diambil adalah bagaimana menganalisa data *bioinformatics* menggunakan klasifikasi, metode yang digunakan adalah *artificial neural network* yaitu *backpropagation*, dengan studi kasus sel kanker payudara.

1.3 Tujuan Penelitian

1. Melakukan analisa klasifikasi data *bioinformatics*, dengan studi kasus sel kanker payudara
2. Mengembangkan bidang ilmu *Artificial Intelligence*, khususnya *Artificial Neural Network* dengan metode *Backpropagation*
3. Memahami konsep *machine learning* untuk *big data*, dengan metode *supervised learning*

1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Memberikan sumbangan pengetahuan yang berarti dalam perkembangan ilmu komputer, khususnya yang berkaitan *Artificial Intelligence*, *bioinformatics* dan metode *Backpropagation*.
2. Dapat membantu dalam mempelajari algoritma *Backpropagation* dengan langsung menerapkan metode yang ada pada kasus sel kanker payudara.
3. Sebagai alat bantu dalam menentukan kanker payudara di bidang informatika kesehatan

BAB II

TINJAUAN PUSTAKA

Bioinformatics adalah salah satu cabang baru ilmu biologi yang merupakan perpaduan antara biologi dan teknologi informasi. Menurut Altschul (1990) *bioinformatics* adalah manajemen dan analisis informasi biologis yang disimpan dalam *database*.

Ilmu ini mengajarkan aplikasi, analisis, dan mengorganisir miliaran bit informasi genetik dalam sel makhluk hidup. Studi *bioinformatics* terutama didukung oleh studi genomik, biologi komputasi, dan teknologi computer (Claverie, J. M., and Notredame, C 2007), genomik adalah studi yang berhubungan dengan pemetaan, sekuen, dan analisis genom. Walaupun belum jelas, secara umum Genomik bisa diartikan sebagai penggunaan informasi genom secara sistematis, dengan data eksperimental baru untuk menjawab permasalahan biologis, medis, maupun industri (Dominico, *et. al.*, 1992).

Bioinformatics sendiri mencakup kajian yang lebih mendalam dari genomik. Dalam studi *bioinformatics* digunakan komputer yang mampu menyimpan data dalam jumlah yang sangat banyak dan didukung berbagai macam software untuk menganalisis jutaan data yang berasal dari makhluk hidup.

Studi *Bioinformatics* mulai tumbuh sebagai akibat dari perkembangan berbagai metode sekuens baru yang menghasilkan data yang sangat banyak. Hal tersebut, secara kebetulan, didukung pula oleh teknologi penyimpanan, manajemen, dan pertukaran data melalui komputer. Inovasi dalam pemetaan dan sekuensing memiliki peran penting dalam proses pengambilan data biologis. Penggunaan *Yeast Artificial Chromosome (YAC)*, sangat membantu dalam konstruksi peta fisik *genome* kompleks secara lengkap (Fox.J.,2007). Untuk mengkloning fragmen-fragmen DNA besar (sekitar 150.000 pasangan basa) digunakan *bacterial Artificial Chromosome (BAC)*.

Perkembangan teknologi sekuensing dimulai dan semi-automatic sequencer yang pertama pada tahun 1987, dilanjutkan dengan Taq Cycle sequencing pada tahun 1990. Pelabelan Flourescen fragmen DNA dengan Sanger dideoxy Chain Termination Method, merupakan dasar bagi proyek sekuensing skala besar (Franca, *et. al.*, 2002).

Seluruh perkembangan tersebut sia-sia saja tanpa obyek yang diteliti, yang memiliki nilai komersil tinggi dan data yang berlimpah. Gampang ditebak, pasti Manusia melalui Human Genome Project. Selain perkembangan dalam bidang Genomik, *Bioinformatics* sangat dipengaruhi oleh perkembangan di bidang

teknologi informasi dan komputer. Pada fase awal sekitar tahun 80-an perkembangan yang paling signifikan adalah kapasitas penyimpanan data. Dari hanya beberapa puluh byte (Voet, D., *et al*, 2006), hingga mencapai Terabyte.

Setelah pembuatan *database*, selanjutnya dimulai perkembangan pemuatan perangkat lunak untuk mengolah data. Awalnya, metode yang digunakan hanya pencarian kata kunci, dan kalimat pendek. perkembangan selanjutnya berupa perangkat lunak dengan algoritma yang lebih kompleks, seperti penyandian nukleotida, menjadi asam-asam amino, kemudian membuat struktur proteinnya. Saat ini, perangkat lunak yang tersedia meliputi pembacaan sekuens nukleotida dari gel elektroforesis, prediksi kode protein, identifikasi primer, perbandingan sekuens, analisis kekerabatan, pengenalan pola dan prediksi struktur. Dengan perkembangan seperti diatas, ternyata masih belum cukup. Kurangnya pemahaman terhadap sistem biologis dan organisasi molekular membuat analisis sekuens masih mengalami kesulitan. Perbandingan sekuens antar spesies masih sulit akibat variabilitas DNA.

Usaha yang dilakukan saat ini, baru mencoba mempelajari teori-teori tersebut melalui proses inferensi, penyesuaian model, dan belajar dari contoh yang tersedia. Perkembangan perangkat keras komputer juga berperan sangat penting. Kecepatan prosesor, kapasitas RAM, dan VGA merupakan salah satu pendorong majunya *bioinformatics*. Terakhir perkembangan *bioinformatics* sangat dipengaruhi oleh pertumbuhan jaringan internet, mendukung kemudahan transfer data yang cepat dan mudah. Saat ini, telah tersedia sekitar 400 *database* biologis yang dapat diakses melalui internet. Beberapa aplikasi *bioinformatics*

1. Transformasi sekuen menjadi informasi genetik.

Intinya adalah menjual data, dalam bentuk gen komplet, atau fragmen, yang dapat digunakan oleh pihak lain untuk mencari potensi terhadap gen tersebut.

2. Pasien sebagai komoditas

Pasien dengan kecenderungan terhadap penyakit tertentu dapat diketahui, sehingga mudah sekali bagi perusahaan untuk menawarkan produknya.

3. Mencari potensi gen

Potensi dari sebuah gen sangat beragam, bergantung pada ekspresi gen tersebut. Aplikasi lebih lanjut dapat berupa *transgenik*, terapi genetik, atau berbagai rekayasa dan pemanfaatan genetik lainnya.

Penelitian-penelitian yang sudah dilakukan dan dijadikan sebagai acuan dalam penelitian ini adalah sebagai berikut :

1. *Clustering* Industri di Kabupaten Kudus Menggunakan *Fuzzy Possibilistik C-Means Standard* (Setiawan. A, 2012), penelitian ini membahas mengenai metode *clustering* menggunakan logika *fuzzy* untuk mengelompokkan skala industri dengan membentuk *class* baru untuk mempermudah dalam pengelolaan industri kecil menengah.
2. Implementasi Sistem Pengenalan Karakter Menggunakan Jaringan Syaraf Tiruan Untuk Pembacaan Dokumen Yang Rusak Karena Banjir (Setiawan. A, 2013), dalam penelitian ini membahas mengenai metode jaringan syaraf tiruan *Bacpropagation* untuk pengenalan karakter, dengan studi kasus dokumen yang rusak karena banjir.
3. Klasifikasi sel kanker payudara menggunakan metode *Perceptron* (Setiawan. A, 2015), penelitian ini membahas mengenai klasifikasi sel kanker payudara menggunakan metode *Perceptron*, penelitian ini menghasilkan pengelompokan kanker payudara di tingkat rendah, ringan dan tingkat akut.

BAB III

METODE PENELITIAN

3.1 Perencanaan analisis klasifikasi

Pada tahap ini dilakukan analisa terhadap kebutuhan rancang bangun sistem informasi klasifikasi sistem, kegiatan yang dilakukan adalah :

- a. Mengidentifikasi masalah yang ada pada sel penyakit kanker payudara
- b. Mengidentifikasi kebutuhan teknologi untuk klasifikasi sel kanker
- c. Pemilihan / kelayakan system, pada tahap ini memilih satu dari sekian metode *klasifikasi artificial neural network*

3.2 Observasi dan Pengolahan data primer yaitu data *bioinformatics*

Untuk mendapatkan data yang akurat maka dalam penelitian ini pengumpulan data dilakukan dengan cara :

- a. Observasi
Pada tahap ini pengumpulan data melalui pengamatan dan pencatatan terhadap sel kanker payudara dari Bankdata *EMBL*
- b. Studi Pustaka
Pada tahap ini mencari teori yang dapat digunakan sebagai landasan teori/kerangka dalam penelitian klasifikasi data *Bioinformatics*, jurnal penelitian yang membahas mengenai *bioinformatics*, untuk mencari metodologi yang sesuai dan membandingkan antara teori yang ada dengan fakta yang ada di lapangan.

3.3 Analisa Sistem

Tahapan Analisa ini dilakukan pada saat tahap perencanaan telah selesai. Pada tahapan ini melakukan penelitian lanjutan diperlukan untuk memperoleh data yang lebih terperinci (*detail investigation*), yang bertujuan untuk keperluan pengembangan sistem secara teknis.

Langkah-langkah yang perlu dilakukan dalam tahapan analisa sistem ini adalah sebagai berikut :

- a. Mengidentifikasi dan menganalisa masalah sistem yang berjalan (*problem analysis*), dalam hal ini dilakukan analisa mengenai kendala-kendala dalam proses klastering skala industri yang sudah digunakan selama ini.
- b. Mengidentifikasi dan menganalisa kriteria dan performance sistem yang dibutuhkan (*requirements analysis*), dalam hal ini dilakukan analisa mengenai kriteria sistem klastering skala industri yang dibutuhkan.
- c. Memberikan alternatif sistem yang diusulkan (*Generating Systems Alternatif*). Pada bagian ini dilakukan pemilihan sistem beserta input, proses dan output dari sistem aplikasi klastering skala industri yang disesuaikan dengan kebutuhan.
- d. Membuat usulan perancangan. Pada bagian ini dilakukan usulan perancangan yang akan digunakan untuk mendesain sistem aplikasi klastering skala industri, apabila usulan perancangan sistem disetujui, maka dapat diteruskan ke tahapan perancangan sistem.

3.4 Desain Machine Learning Menggunakan Backpropagation

Pada tahap ini dilakukan penyusunan rancang bangun klasifikasi data *bioinformatics* menggunakan *artificial neural network*

1. Perancangan Sistem

Perancangan secara umum atau perancangan secara konseptual (*conceptual design*). Tahapan ini adalah tindak lanjut dari analisa sistem, tahapan yang dilakukan untuk perancangan sistem aplikasi adalah :

- a. Memilih metode perancangan sistem, pada tahap ini disesuaikan dengan tahapan analisa sistem langkah terakhir.
 - b. Merancang pemodelan data yang sesuai dengan metode perancangan sistem.
- #### **2. Perancangan sistem secara terperinci (*detail design*)**
- a. Merancang konsep penyimpanan data atau basis data.
 - b. Merancang bagan alir sistem aplikasi.
 - c. Merancang hirarki input output dan antarmuka sistem aplikasi.

3.5 Implementasi Data Training

Pada tahap ini dilakukan pembuatan program klasifikasi data *bioinformatics* menggunakan *artificial neural network*. Pada tahap ini dilakukan penterjemahan analisa dan perancangan kedalam bahasa pemrograman. Pada tahap ini kegiatan yang dilakukan adalah :

1. Pengembangan analisa dan perancangan kedalam bahasa pemrograman (*coding*).
2. Melakukan pengujian sistem dengan menggunakan contoh data, dan membahas *step by step* perhitungan algoritma FPCM secara detail.
3. Mengevaluasi hasil pengujian sistem serta memperbaiki model sistem yang telah dikembangkan jika terjadi kesalahan pada saat uji coba sistem.

BAB IV HASIL DAN PEMBAHASAN

4.1 Perambatan-balik (Backpropagation)

Proses pelatihan perambatan balik meliputi tiga tahap, yaitu : prosedur umpan maju, perhitungan serta perambatan balik kesalahan, dan penyesuaian bobot. Algoritma perambatan balik diuraikan dalam langkah-langkah atau alur prosedur sebagai berikut :

1. Setiap unit masukan ($X_n, n = 1, \dots, n$) menerima sinyal-sinyal masukan x_n dan mengirimkan sinyal-sinyal ini ke unit-unit selanjutnya (unit-unit tersembunyi).
2. Setiap unit tersembunyi ($I_h, h = 1, \dots, h$) menjumlahkan sinyal-sinyal terbobotnya :

$$i_in_h = \theta_{im} + \sum_n x_n w_{hn} \quad (4)$$

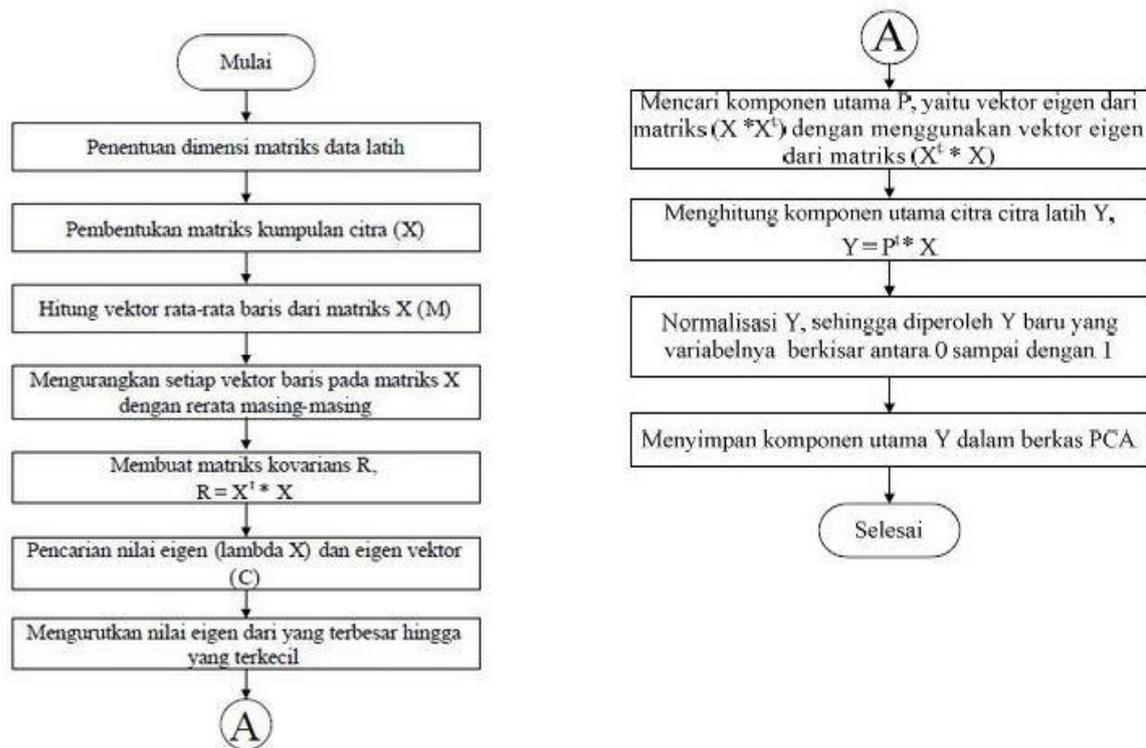
Kemudian menerapkan fungsi aktifasinya untuk menghitung sinyal keluarannya : $I_h = f(i_in_h)$ (5) lalu mengirimkannya pada semua unit lapis lapis keluaran.

3. Setiap unit keluaran ($O_k, k = 1, \dots, k$) menjumlahkan sinyal masukan terbobotnya :

$$o_in_k = \theta_{kh} + \sum_h i_h w_{kh} \quad (6)$$

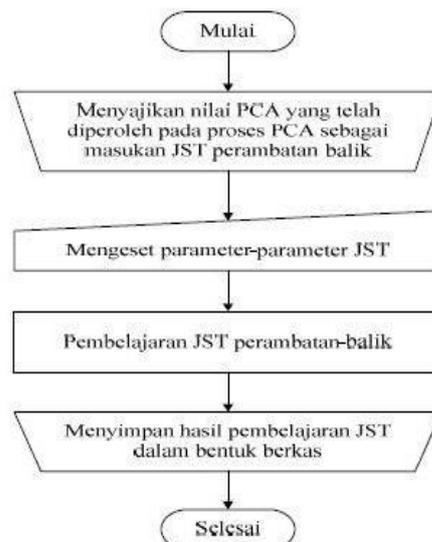
Kemudian menerapkan fungsi aktivasi untuk menghitung sinyal keluarannya.

4.2 Perancangan Sistem dan Perangkat Lunak



Gambar 1.1 Diagram alir proses PCA

4.3 Proses Jaringan Syaraf Tiruan



Gambar 4.2 Diagram alir proses JST

4.4 Proses Klasifikasi Sel Kanker Payudara



Gambar 4.3 Diagram alir Klasifikasi Sel Kanker Payudara

Proses pada Data Training

Training : 500, Learning : 0.3, Momentum 0.2

PerformanceVector

PerformanceVector:

accuracy: 97.33% +/- 3.27% (mikro: 97.33%)

ConfusionMatrix:

True: Cancer-ringani Cancer-medium Cancer-akut

Cancer-ringani: 50 0 0

Cancer-medium: 0 47 1

Cancer-akut: 0 3 49

kappa: 0.960 +/- 0.049 (mikro: 0.960)

ConfusionMatrix:

True: Cancer-ringani Cancer-medium Cancer-akut

Cancer-ringani: 50 0 0

Cancer-medium: 0 47 1

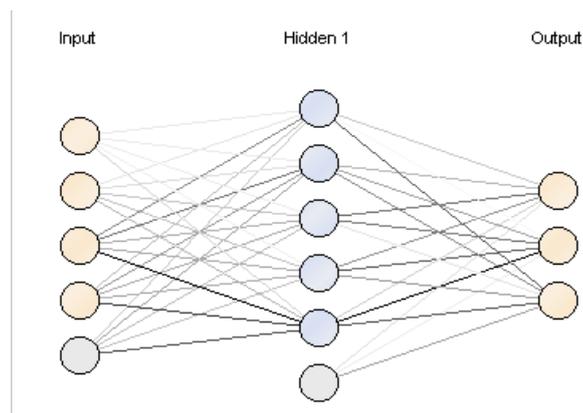
Cancer-akut: 0 3 49

accuracy: 97.33% +/- 3.27% (mikro: 97.33%)

	true Cancer- ringan	true Cancer- medium	true Cancer- akut	class precision
pred. Cancer- ringan	50	0	0	100.00%
pred. Cancer- medium	0	47	1	97.92%
pred. Cancer- akut	0	3	49	94.23%
class recall	100.00%	94.00%	98.00%	

kappa: 0.960 +/- 0.049 (mikro: 0.960)

	true Cancer- ringan	true Cancer- medium	true Cancer- akut	class precision
pred. Cancer- ringan	50	0	0	100.00%
pred. Cancer- medium	0	47	1	97.92%
pred. Cancer- akut	0	3	49	94.23%
class recall	100.00%	94.00%	98.00%	



Hidden 1

=====

Node 1 (Sigmoid)

Sepal-length: 0.840
 sepal-width: 0.995
 petal-length: -4.539
 petal-width: -2.527
 Bias: 2.067

Node 2 (Sigmoid)

Sepal-length: -0.743
 sepal-width: -1.335
 petal-length: 5.353
 petal-width: 3.921
 Bias: -3.557

Node 3 (Sigmoid)

Sepal-length: 0.676
sepal-width: -2.209
petal-length: 3.098
petal-width: 3.214
Bias: 2.532

Node 4 (Sigmoid)

Sepal-length: -0.604
sepal-width: 2.322
petal-length: -3.176
petal-width: -3.211
Bias: -2.377

Node 5 (Sigmoid)

Sepal-length: -1.581
sepal-width: -2.745
petal-length: 9.712
petal-width: 7.309
Bias: -6.625

Output

=====

Class 'Cancer-ringani' (Sigmoid)

Node 1: 2.104
Node 2: -2.864
Node 3: -5.578
Node 4: 4.152
Node 5: -1.792
Threshold: -1.183

Class 'Cancer-medium' (Sigmoid)

Node 1: 0.248
Node 2: -3.517
Node 3: 5.245
Node 4: -6.055
Node 5: -9.433
Threshold: 0.670

Class 'Cancer-akut' (Sigmoid)

Node 1: -6.313
Node 2: 4.396
Node 3: 1.019
Node 4: -3.824
Node 5: 6.297
Threshold: -3.556

(Fareed Akthar, 24th August 2012)

Pada proses training diatas kami menggunakan tool PCA, pada hidden layer kami berikan 5 node simpul dengan bobot yang diacak secara random, untuk nilai learning = 0,3 dan momentum atau bias = 0.2 serta proses iterasi atau perulangan yang digunakan sebanyak 500 kali. Hasil yang didapat pada proses diatas menunjukkan bahwa akurasi yang didapat pada proses tersebut adalah 97,3% dimana pada clas Cancer-ringannya dapat diprediksi keakuratannya sebesar 100% ringannya untuk prediksi Cancer-medium masih ada failed 1 pada cancer akut. Dan prediksi pada akut terdapat failed 3 record pada Cancer-medium. (Larose, 2005)

Proses BPN pada Data Testing

Training : 500, Learning : 0.3, Momentum 0.2

PerformanceVector

PerformanceVector:

accuracy: 100.00% +/- 0.00% (mikro: 100.00%)

ConfusionMatrix:

True: Cancer-ringannya Cancer-medium Cancer-akut

Cancer-ringannya: 8 0 0

Cancer-medium: 0 13 0

Cancer-akut: 0 0 17

kappa: 1.000 +/- 0.000 (mikro: 1.000)

ConfusionMatrix:

True: Cancer-ringannya Cancer-medium Cancer-akut

Cancer-ringannya: 8 0 0

Cancer-medium: 0 13 0

Cancer-akut: 0 0 17

accuracy: 100.00% +/- 0.00% (mikro: 100.00%)

	true Cancer-ringannya	true Cancer-medium	true Cancer-akut	class precision
pred. Cancer-ringannya	8	0	0	100.00%
pred. Cancer-medium	0	13	0	100.00%
pred. Cancer-akut	0	0	17	100.00%
class recall	100.00%	100.00%	100.00%	

kappa: 1.000 +/- 0.000 (mikro: 1.000)

	true Cancer-ringannya	true Cancer-medium	true Cancer-akut	class precision
pred. Cancer-ringannya	8	0	0	100.00%
pred. Cancer-medium	0	13	0	100.00%
pred. Cancer-akut	0	0	17	100.00%
class recall	100.00%	100.00%	100.00%	

Hidden 1
=====

Node 1 (Sigmoid)

Sepal-length: -0.381
sepal-width: 1.528
petal-length: -1.899
petal-width: -2.279
Bias: -1.772

Node 2 (Sigmoid)

Sepal-length: -0.232
sepal-width: -2.121
petal-length: 4.430
petal-width: 6.212
Bias: -3.220

Node 3 (Sigmoid)

Sepal-length: -0.307
sepal-width: -2.295
petal-length: 4.796
petal-width: 6.671
Bias: -3.443

Node 4 (Sigmoid)

Sepal-length: -0.596
sepal-width: -1.365
petal-length: 1.546
petal-width: 2.404
Bias: -0.111

Node 5 (Sigmoid)

Sepal-length: 0.438
sepal-width: -2.052
petal-length: 2.356
petal-width: 2.886
Bias: 2.726

Output
=====

Class 'Cancer-ringan' (Sigmoid)

Node 1: 2.702
Node 2: -2.382
Node 3: -2.019
Node 4: -3.592
Node 5: -3.999
Threshold: 1.457

Class 'Cancer-medium' (Sigmoid)

```
-----  
Node 1: -4.200  
Node 2: -5.649  
Node 3: -6.588  
Node 4: 0.873  
Node 5: 5.221  
Threshold: 0.128  
  
Class 'Cancer-akut' (Sigmoid)  
-----  
Node 1: -3.255  
Node 2: 5.158  
Node 3: 5.265  
Node 4: 2.337  
Node 5: -0.142  
Threshold: -6.229
```

Pada proses testing diatas pada hidden layer kami berikan 5 node simpul dengan bobot yang diacak secara random, untuk nilai learning = 0,3 dan momentum atau bias = 0.2 serta proses iterasi atau perulangan yang digunakan sebanyak 500 kali sama seperti pada training. Hasil yang didapat pada proses diatas menunjukkan bahwa akurasi yang didapat pada proses tersebut adalah 100% untuk semua prediksi klass yang ada. (Kadir, 2010) Dengan demikian tingkat akuratan pada pola 5 node pada hidden layer dengan 500 iterasi serta menggunakan learning 0.3 dan bias 0.2 mendekati tingkat kesempurnaan dengan terbukti tingkat error yang sedikit dan akurasi data testing maupun training yang lebih dari 90% (Larose, 2005).

BAB V

KESIMPULAN DAN REKOMENDASI

5.1 Kesimpulan

Berdasarkan proses yang telah dilakukan pada penelitian ini ini, , dapat disimpulkan beberapa hal, antara lain :

1. Penggabungan dua metode yaitu analisis komponen utama (PCA) dan jaringan syaraf tiruan perambatan-balik yang diterapkan dalam sistem pengenalan wajah memberikan hasil yang sangat baik. Hal ini ditunjukkan dengan tingkat pengenalan selama pengujian, yaitu dengan tingkat pengenalan rata-rata 85 %.
2. Jumlah citra latih yang digunakan pada penelitian ini berbanding lurus dengan kinerja sistem klasifikasi kanker, artinya semakin banyak citra latih yang digunakan untuk pelatihan, hasil pengenalannya akan semakin baik.

5.2 Rekomendasi

Beberapa saran yang bisa diberikan untuk selanjutnya antara lain :

1. Perlu dilakukan penelitian lanjutan untuk mengembangkan sistem klasifikasi sel kanker payudara.
2. Perlu dilakukan pengembangan lebih lanjut mengenai pengenalan stadium cancer yang lebih spesifik dengan parameter yang lebih luas.

DAFTAR PUSTAKA

- Altschul SF., W. Gish., W. Miller, 1990, Basic Local Alignment Search Tool. *J. Mol. Biol.* 215, 403–410. The first BLAST paper, a classic in the field of bioinformatics.
- Claverie, J. M. and Notredame, C., 2007, *Bioinformatics for Dummies*, Second Edition, Wiley Publishing Inc., New York.
- Dominico, P., Marx, J., Scoch, P. and Cunha, B., 1992, Rapid Plasmid DNA Isolation From Mucoid Gram-Negative Bacteria, *J. Clin. Microbiol.*, 2859- 2863.
- Fox, J., 2007, *BLAST*, Bioinformatics, Michael Smith Laboratories, UBC
- Franca, L., Carrilho, E. and Kist, T., 2002, A Review of DNA Sequencing Techniques, *Rev. Biophys.*, 169-200
- Fareed Akthar, C. H. (24th August 2012). *RapidMiner 5 Operator Reference*. Rapid-I www.rapid-i.com.
- Fisher. (1936). *Cancer Data Set* . Dipetik November 19, 2014, dari UCI Machine Learning Repository: <http://archive.ics.uci.edu/ml/datasets/Cancer>
- Kadir, A. (2010). *Identifikasi Tiga Jenis Bunga Cancer Menggunakan Anfis*, 6.
- Larose, D. T. (2005). *DISCOVERING KNOWLEDGE IN DATA An Introduction to Data Mining*. A JOHN WILEY & SONS, INC., PUBLICATION.
- Setiawan. A. Setiaji. P., 2012, *Clustering Industri di Kabupaten Kudus Menggunakan Fuzzy Possibilistik C-Means Standard*, *KNASTIK*, UKDW, Yogyakarta,
- Setiawan. A, 2013 *Implementasi Sistem Pengenalan Karakter Menggunakan Jaringan Syaraf Tiruan Untuk Pembacaan Dokumen Yang Rusak Karena Banjir*, *SAINTECH*, UMK, Kudus
- Setiawan. A, 2015, *Klasifikasi sel kanker payudara menggunakan metode Perceptron*, *SNATIF*, UMK, Kudus
- Voet, D., Voet, J.G. and Pratt, C.W., 2006, *Fundamentals of Biochemistry Life at The Moleculer Level*, 2nd ed., John Wiley and Sons Inc USA

Lampiran

Daftar Riwayat Hidup Ketua Peneliti

A. Identitas Diri

1	Nama Lengkap	Arif Setiawan, S.Kom, M.Cs
2	Jabatan Fungsional	Lektor
3	Jabatan Struktural	-
4	NIP/NIK/Identitas Lainnya	0610 7010 0000 1180
5	NIDN	0623018201
6	Tempat dan Tanggal Lahir	Semarang, 23 Januari 1982
7	Alamat Rumah	Perum Grand Regency No B3, Purwosari Kudus
8	Nomor Telepon/Faks/HP	08156672762
9	Alamat Kantor	Gondangmanis, Po. Box. 53 Bae, Kudus,
10	Nomor Telepon/Faks/HP	(0291)438229
11	Alamat e-mail	arif.setiawan@umk.ac.id

B. Riwayat Pendidikan

	S-1	S-2
Nama Perguruan Tinggi	Universitas Dian Nuswantoro	Universitas Gadjah Mada
Bidang Ilmu	Teknik Informatika	Ilmu Komputer
Tahun Masuk-Lulus	2000-2004	2007-2009
Judul Skripsi/Thesis/Desertasi	Sistem Pendukung Keputusan Pengendalian Narkotika Dan Psikotropika Pada Dinas Kesehatan Jawa Tengah	Klastering Skala Industri Di Kabupaten Kudus Menggunakan Fuzzy Possibilistic C-Means Standar
Nama Pembimbing/Promotor	Dr. Abdul Syukur, M.M	Drs.RetantyoWardoyo, M.Sc.,Ph.D

C. Pengalaman Penelitian

No	Nama Kegiatan	Program	Tahun
1	Implementasi Enkripsi Database Dengan Algorithma Kriptografi	APBU	2004
2	Implementasi Sistem Informasi Geografis Untuk Pengolahan Sumber Daya Air Tanah di Kabupaten Kudus	Dosen Muda, Dikti	2005
3	Instalasi Tower Internet untuk RT/RW Net	APBU	2005
4	Implementasi Sistem Informasi Penjualan Jenang On Line Di Kabupaten Kudus	Dosen Muda, Dikti	2006
5	Membina Wirausaha Baru IBC (Bisnis Laundry)	Kementrian Perindustrian	2007
6	Instalasi Tower Internet untuk RT/RW Net	APBU	2008
7	Clustering Industri di Kabupaten Kudus	Tesis UGM	2009

	Menggunakan Fuzzy Possibilistik C-Means Standar		
8	Penerapan Teknologi Untuk Desa Binaan dengan produk TILASA (Teknologi Layak Desa)	APBU	2009
9	Membina Wirausaha Baru Mahasiswa PKMK (Bisnis Warnet)	Kementerian Perindustrian	2010
10	Implementasi Sistem Pengenalan Karakter Menggunakan Jaringan Syaraf Tiruan Untuk Pembacaan Dokumen Yang Rusak Karena Banjir	APBU	2010
11	Membina Wirausaha Baru IBC (Home Industri Konveksi)	Kementerian Perindustrian	2011
12	Sistem Absensi Karyawan Menggunakan Sistem Deteksi Suara	UMK	2011
13	Penerapan Teknologi Produksi Dan Pemasaran Kerajinan Bordir di Paguyuban Bordir Lestari Kudus Sebagai Produk Unggulan Kabupaten Kudus Berbasis Teknologi Informasi	IPTEKDA LIPI	2012
14	Pengembangan Teknologi Desain Grafis untuk Peningkatan Kapasitas dan Mutu Bordir Bermesin Bordir Manual di Paguyuban Bordir Kabupaten Kudus Jawa Tengah	Kementerian Ristek	2012

D. Publikasi

No	Judul	Keterangan
1	Nama : Arif Setiawan (2007) Judul : “ Menumbuhkan Grup Riset Jaringan Komputer di Universitas Muria Kudus”.	Dipublikasikan dalam MAWAS Universitas Muria Kudus, ISSN: 0853-0335, Vol 18, Edisi No. 1/Juni/2007.
2	Nama : Arif Setiawan (2009) Judul : “Analisis Klasifikasi Suara Berdasarkan Gender Dengan Format Wav Menggunakan Algoritma k-means”	Dipublikasikan dalam Majalah Ilmiah JURNAL SAINS DAN TEKNOLOGI Universitas Muria Kudus, ISSN: 1979-6870, Vol 2, Edisi No. 2/Des/2009.
3	Nama : Arif Setiawan, Diana Laily Fitri, Nanik Susanti (2010) Judul : “Analisa Pengenalan Karakter Menggunakan Jaringan Syaraf Tiruan Untuk Pembacaan Dokumen Yang Rusak Karena Banjir”	Dipublikasikan dalam Majalah Ilmiah JURNAL SAINS DAN TEKNOLOGI Universitas Muria Kudus, ISSN: 1979-6870, Vol 3, Edisi No. 1/Juni/2010).
4	Nama : Arif Setiawan, Pratomo Setiaji, Andy Prasetyo Utomo (2011) Judul “Sistem Informasi Perwalian Pada	Dipublikasikan dalam Majalah Ilmiah JURNAL SAINS DAN TEKNOLOGI Universitas Muria Kudus, ISSN: 1979-

	Program Studi Sistem Informasi Universitas Muria Kudus”	6870, Vol 4, Edisi No. 1/Juni/2011).
5	Nama : Arif Setiawan, Pratomo Setiaji (2011) Judul “Klasifikasi Suara Berdasarkan Gender Dengan Algoritma K-Means”	Dipublikasikan melalui seminar Nasional Pengaplikasian Telematika SINAPTIKA 2011, Prosiding SINAPTIKA 2011 Universitas Mercu Buana,ISSN: 2086-8251, 23/Juli/2011.
6	Nama : Pratomo Setiaji, Arif Setiawan (2011) Judul “Sistem Informasi Geografis Obyek Wisata Kabupaten Kudus”	Dipublikasikan melalui seminar Nasional Pengaplikasian Telematika SINAPTIKA 2011, Prosiding SINAPTIKA 2011 Universitas Mercu Buana,ISSN: 2086-8251, 23/Juli/2011.

Kudus, 13 November 2017
Ketua Tim Pengusul

Arif Setiawan, S.Kom, M.Cs
NIP. 0610 7010 00001180

Daftar Riwayat Hidup Anggota Peneliti

A. Identitas Diri

1	Nama Lengkap (dengan gelar)	Pratomo Setiaji, S.Kom., M.Kom
2	Jabatan Fungsional	Lektor
3	Jabatan Struktural	-
4	NIP/NIK/Identitas Lainnya	0610 7010 0000 1193
5	NIDN	0619067802
6	Tempat dan Tanggal Lahir	Semarang, 19 Juni 1978
7	Alamat Rumah	Jl. Muria Raya II/34 Kudus
8	Nomor Telepon/Faks/HP	08122885008
9	Alamat Kantor	Gondangmanis, Po. Box. 53 Bae, Kudus,
10	Nomor Telepon/Faks/HP	(0291)438229
11	Alamat e-mail	pratomo.setiaji@yahoo.com

B. Riwayat pendidikan

	S-1	S-2
Nama Perguruan Tinggi	Universitas Dian Nuswantoro	Universitas Diponegoro
Bidang Ilmu	Sistem Informasi	Sistem Informasi
Tahun Masuk-Lulus	1996-2001	2008-2011
Judul Skripsi/Thesis/Desertasi	Sistem Informasi Pengisian Krs Pada Sekolah Tinggi Ilmu Ekonomi (Stie) Dharmaputra Semarang	Penerapan AHP Dan GIS Untuk Pendeteksian Kondisi Jembatan Di Kabupaten Kudus
Nama Pembimbing/Promotor	Drs. Agus Prayitno, M.M	Prof. Drs. Mustafid M. Eng., Ph.D

C. Pengalaman Penelitian Dalam 5 Tahun Terakhir

No	Tahun	Judul Penelitian	Pendanaan	
			Sumber*	Jml (Juta Rp)
1	2005	Sistem Informasi Laboratorium Pada Progdi Teknik Mesin UMK	UMK	2
2	2006	Analisa dan Perancangan Sistem Informasi Penjualan Jenang On Line Di Kabupaten Kudus	Penelitian Dosen Muda, DIKTI	10
3	2007	Sistem Informasi Perwalian Pada Progdi Sistem Informasi	UMK	2
4	2008	Analisa dan Perancangan Sistem Informasi Pemasaran Kerajinan Kain Bordir Berbasis Web di Desa Samirejo	Penelitian Dosen Muda, DIKTI	10

		Kecamatan Gebog Kabupaten Kudus		
5	2011	Sistem Informasi Geografi Obyek Wisata Di Kabupaten Kudus	UMK	2.5

D. Pengalaman penulisan artikel Ilmiah Dalam Jurnal Dalam 5 Tahun Terakhir

No	Judul Artikel Ilmiah	Volume/ Nomor/Tahun	Nama Jurnal
1	Sistem Informasi Geografis (SIG) Universitas Muria Kudus	ISSN: 0853-0335 Vol 21 Edisi No 2 Desember 2008 2008 hal : 47 - 52	MAWAS
2	Analisis Klasifikasi Suara Berdasarkan Gender Dengan Format Wav Menggunakan Algoritma k-means	ISSN: 1979-6870, Vol 2, Edisi No. 2/Des/2009	JURNAL SAINS DAN TEKNOLOGI
3	Sistem Informasi Geografis Obyek Wisata Kabupaten Kudus	ISSN: 2086-8251, 23/Juli/2011	Prosiding Sinaptika

E. Pengalaman Penyampaian Makalah Secara Oral Pada Pertemuan / Seminar Ilmiah Dalam 5 Tahun Terakhir

No	Nama Pertemuan Ilmiah / Seminar	Judul Artikel ilmiah	Waktu dan tempat
1	Seminar Nasional Teknik Informatika & Sistem Informasi SETISI	Sistem Informasi Geografis Tataletak Pengairan Sungai dan Irigasi Pada Balai PSDA SELUNA	Universitas Kristen Maranatha 24 September 2011
2	International Conference On Information Systems For Business Competitiveness ICISBC	Geographic Information System tourism in Kudus	8-9 Desember 2011

Kudus, 9 November 2018

Anggota Peneliti



Pratomo Setiaji, S.Kom, M.Kom

NIP. 0610 7010 0000 1193

Mahasiswa yang menjadi tim penelitian :

1. Nama : Moh Radiksa Indriyanto
Nim : 201253012
Program Studi : Sistem Informasi
2. Nama : Andi agus Setyawan
Nim : 201253021
Program Studi : Sistem Informasi
3. Nama : Muhammad Syaifudin
Nim : 201253023
Program Studi : Sistem Informasi
4. Nama : Aris Suryawan
Nim : 201153005
Program Studi : Sistem Informasi
5. Nama : Brahma Ari Prabowo
Nim : 201253001
Program Studi : Sistem Informasi

Alokasi Anggaran Penelitian

ANGGARAN	Jumlah
Penerimaan : APBU Universitas Muria Kudus 2016/2017	Rp. 4.000.000
Pengeluaran :	
1. Pembuatan Proposal	Rp. 100.000
2. Analisa Sistem & kebutuhan informasi	Rp 1.400.000
3. Perancangan sistem	Rp. 1.000.000
4. Pembuatan <i>Prototype</i> Sistem	Rp. 1.000.000
5. Pembuatan Laporan	Rp. 100.000
6. Dokumentasi	Rp. 100.000
7. Sewa alat	Rp. 300.000
Jumlah	Rp. 4.000.000