

**KAJIAN PENERAPAN ALGORITMA C4.5, MLP, DAN NAIVE BAYES
UNTUK PREDIKSI *SOFTWARE DEFECT***

TESIS



Oleh:

Aditiya Hermawan

1111600209

PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)

PROGRAM PASCASARJANA

UNIVERSITAS BUDI LUHUR

JAKARTA

2013

**ADITIYA
HERMAWN
1111600209**

**KAJIAN PENERAPAN ALGORITMA C4.5, MLP, DAN NAIVE BAYES UNTUK PREDIKSI
SOFTWARE DEFECT**



**TESIS
AGUSTUS
2013**

**KAJIAN PENERAPAN ALGORITMA C4.5, MLP, DAN NAIVE
BAYES UNTUK PREDIKSI *SOFTWARE DEFECT***

TESIS



Oleh :
Aditiya Hermawan
1111600209

**PROGRAM STUDI : MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCA SARJANA
UNIVERSITAS BUDI LUHUR**

**JAKARTA
2013**

**KAJIAN PENERAPAN ALGORITMA C4.5, MLP, DAN NAIVE
BAYES UNTUK PREDIKSI *SOFTWARE DEFECT***

TESIS

Diajukan untuk memenuhi salah satu persyaratan
memperoleh gelar Magister Ilmu Komputer (MKOM)



Oleh :
Aditiya Hermawan
1111600209

**PROGRAM STUDI : MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCA SARJANA
UNIVERSITAS BUDI LUHUR**

**JAKARTA
2013**



PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

LEMBAR PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama Mahasiswa : Aditiya Hermawan
Nomor Induk Mahasiswa : 1111600209
Program Studi : Magister Ilmu Komputer
Konsentrasi : Teknologi Sistem Informasi
Fakultas/Program : Pascasarjana

menyatakan bahwa Tesis yang berjudul:

***KAJIAN PENERAPAN ALGORITMA C4.5, MLP, DAN NAIVE BAYES UNTUK
PREDIKSI SOFTWARE DEFECT***

1. Merupakan hasil karya tulis ilmiah sendiri dan bukan merupakan karya yang pernah diajukan untuk memperoleh gelar akademik oleh pihak lain,
2. Saya ijin untuk dikelola oleh Universitas Budi Luhur sesuai dengan norma hukum dan etika yang berlaku.

Pernyataan ini saya buat dengan penuh tanggung jawab dan saya bersedia menerima konsekuensi apapun sesuai aturan yang berlaku apabila di kemudian hari pernyataan ini tidak benar.

Jakarta, 27 Agustus 2013

(Aditiya Hermawan)



PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

LEMBAR PENGESAHAN

Nama Mahasiswa : Aditiya Hermawan
Nomor Induk Mahasiswa : 1111600191
Konsentrasi : Teknologi Sistem Informasi
Judul Tesis : KAJIAN PENERAPAN ALGORITMA C4.5, MLP,
DAN NAIVE BAYES UNTUK PREDIKSI *SOFTWARE*
DEFECT

Jakarta, 27 Agustus 2013

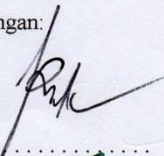
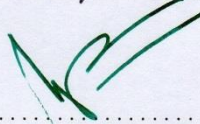
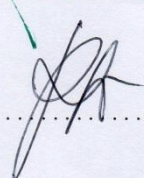
Tim Penguji:

Tanda Tangan:

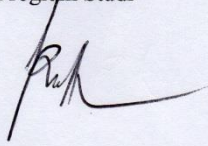
Ketua,
(Dr., Ir. Nazori Az, M.T)

Anggota,
(Dr., Ir. Wendi Usino, M.Sc., M.M)

Pembimbing Utama,
(Dr., Ir. Prabowo Pudjo Widodo, MS)


.....

.....

.....

Ketua Program Studi


(Dr., Ir. Nazori Az, M.T)

ABSTRAK

Perusahaan atau institusi diberbagai bidang memerlukan software untuk membantu proses bisnis mereka agar dapat berjalan dengan cepat, tepat, efektif dan efisien. Tentunya software yang digunakan haruslah mempunyai standar kualitas yang baik agar tujuan dari perusahaan atau institusi tersebut dapat terpenuhi. Oleh karena itu diperlukan *software-software* yang tidak memiliki kesalahan / error (*defect*). Kesalahan biasanya disebabkan oleh kesalahan manusia. Sudah banyak peneliti lain melakukan pembuatan model untuk pengembangan *software*, yang nantinya akan digunakan untuk para pengembang untuk membuat software. Model yang diusulkan diharapkan agar dapat menghasilkan *software* yang berkualitas yang tanpa *defect*. Dari beberapa penelitian sebelumnya belum ada model yang akurat untuk prediksi *Software Defect* karena jumlah variabel yang banyak dan beragam mengakibatkan prediksi yang kurang akurat. Model yang cukup baik untuk melakukan prediksi *software defect* adalah C4.5, MLP dan Naïve Bayes. Beberapa peneliti lain juga mencoba meningkatkan akurasi yang ada dengan melakukan pemilihan variable yang digunakan. Pada Penelitian ini *Genetic Algorithm* akan diterapkan untuk pemilihan variable pada metode C4.5, MLP dan Naïve Bayes dengan menggunakan data dari *NASA Dataset*. Setelah itu akan dilakukan pengujian dengan Kurva ROC dan Confusion Matrix untuk mencari model mana yang menghasilkan tingkat akurasi paling tinggi dalam prediksi *software defect*. Hasil akurasi yang diperoleh membuktikan bahwa Naïve Bayes memiliki tingkat akurasi yang lebih tinggi dibandingkan C4.5 dan MLP. Naïve Bayes dengan *Genetic Algorithm* menghasilkan persentase akurasi 88,25% dan dengan nilai AUC (*Area Under Curve*) sebesar 0.772. dengan demikian algoritma Naïve Bayes dioptimisasi dengan *Genetic Algorithm* dapat memprediksi *Software Defect* dengan lebih baik.

Kata Kunci: *Software Defect, C4.5, MLP, Naive Baye , Genetic Algorithm*

ABSTRACT

Companies and institutions in various fields require software to help their business processes in order to run fast, precise, effective and efficient. Surely the software used must have good quality standards for the purpose of a company or institution can be met. Thus the necessary software that does not have a fault / error (defect). Errors are usually caused by human error. There have been many other researchers doing modeling for software development, which will be used for the development of fatherly making software. The proposed model is proposed is expected to produce quality software is without defect. Of several previous studies there has been no accurate model for prediction of Software Defect due to the number of variables are many and varied which resulted in less accurate predictions. Model is good enough to do software defect prediction is C4.5, MLP and Naïve Bayes. Several other researchers also tried to improve the accuracy of existing by variable selection are used. The research on genetic algorithm will be applied to the selection of variables in methods of C4.5, MLP and Naïve Bayes using data from the NASA dataset. After that will be tested with ROC curves and Confusion Matrix to find which model produces the highest level of accuracy in the prediction of software defects. Accuracy results obtained prove that the Naïve Bayes has a higher degree of accuracy than C4.5 and MLP. Naïve Bayes with Genetic Algorithm produces an accuracy percentage of 88.25% and the value of AUC (Area Under Curve) of 0.772. thus Naïve Bayes algorithm optimized with a genetic algorithm to predict the Software Defect better.

Keywords: Software Defect, C4.5, MLP, Naive Bayes, Genetic Algorithm

LEMBAR PERSEMBAHAN

Perubahan yang baik itu bertahap.

(Aditiya Hermawan)

Dengan mengucap puji syukur kepada Tuhan Yang Maha Esa, thesis ini kupersembahkan untuk :

- 1 Bapak. Ouw Kim Tji dan Ibu. Yo Cian Nio, tercinta yang telah membesarkan aku dan selalu membimbing, mendukung, memotivasi, memberi apa yang terbaik bagiku serta selalu mendoakan aku untuk meraih kesuksesanku.
- 2 Kakak dan Adikku (Kel. Ouw) yang telah menjadi curahan hatiku, yang telah memberiku semangat, aku selalu sayang kalian.
- 3 Yunia Oktari, yang selalu setia dan memberikan semangat.

KATA PENGANTAR

Dengan memanjatkan Puji dan syukur kehadiran Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan hidayahnya kepada penulis, sehingga penulis bisa menyelesaikan tesis yang berjudul *Kajian Penerapan Algoritma C4.5, MLP, dan Naïve Bayes* untuk *Prediksi Software Defect*. Tujuan dari penulisan tesis ini adalah sebagai salah satu syarat untuk memperoleh gelar Magister Ilmu Komputer (MKOM) pada Program Studi Magister Ilmu Komputer, Universitas Budi Luhur Jakarta.

Penulis sungguh sangat menyadari, bahwa dalam menyusun tesis ini tidak akan terwujud tanpa adanya dukungan dan bantuan dari pelbagai pihak. Maka, dalam kesempatan ini penulis menghaturkan penghargaan dan ucapan terima kasih yang sebesar - besarnya kepada :

1. Bapak DR. Ir. Prabowo Pudjo Widodo, MS selaku dosen pembimbing tesis yang telah membimbing dan memotivasi penulis dalam mengerjakan tesis ini.
2. Orang Tua, Bpk. Ouw Kim Tji dan Ibu Yo Cian Nio terima kasih atas dukungan doa, semangat dan kasih sayang yang senantiasa diberikan kepada penulis.
3. Keluarga Besar Ouw terima kasih atas dukungan doa, semangat dan kasih sayang yang senantiasa diberikan kepada penulis.
4. Yunia Oktari tercinta yang selalu memberikan semangat, dukungan, kasih sayang dan doa bagi penulis.
5. Bapak dan Ibu Dosen pengampu mata kuliah di Program Studi Magister Ilmu Komputer Universitas Budi Luhur yang telah dengan sabar memberikan ilmu pengetahuan, pencerahan, dan bimbingan dalam belajar.
6. Rekan-rekan mahasiswa MKOM Universitas Budi Luhur kelas XA Semester 1 dan 2 dan MKOM Semester 3 konsentrasi Teknologi Sistem Informasi, terima kasih atas kebersamaan, kerja keras dan dukungan semangatnya.
7. Teman seperjuangan Yusuf Kurnia dan Andre yang telah berbagi suka dan duka mulai dari awal perkuliahan sampai selesainya tesis ini.

8. Teman –teman dari keluarga besar STMIK Buddhi dan semua pihak yang tidak dapat disebutkan satu persatu.

Penulis menyadari, sebagai makhluk Tuhan yang jauh dari kesempurnaan, bahwa masih banyak kekurangan dari tesis ini, oleh karena itu penulis mengharapkan kritik dan saran yang bersifat membangun demi kesempurnaan penelitian tesis ini. Semoga tesis ini masih dapat memberikan manfaat dari keterbatasannya.

Jakarta, Juli 2013

Aditiya Hermawan

DAFTAR ISI

	Halaman
ABSTRAK	i
ABSTRACT	ii
LEMBAR PERSEMBAHAN	iii
KATA PENGANTAR	iv
DAFTAR ISI	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL	xi
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Masalah Penelitian	4
1.2.1 Identifikasi Masalah	4
1.2.2 Pembatasan Masalah	4
1.2.3 Rumusan Masalah	4
1.3 Tujuan dan Manfaat Penelitian	5
1.3.1 Tujuan Penelitian	5
1.3.2 Manfaat Penelitian	5
1.4 Tata-Urut Penulisan	6
1.5 Daftar Pengertian	7
BAB II LANDASAN TEORI DAN KERANGKA PEMIKIRAN	9
2.1 Tinjauan Pustaka	9
2.1.1 Data Mining	9
2.1.2 Algoritma Klasifikasi Data Mining	19
2.1.3 Algoritma C4.5	23
2.1.4 Naïve Bayes	31
2.1.5 Multi Layer Perceptron	35
2.1.6 Pemilihan Variabel	44
2.1.7 Evaluasi dan Validasi Metode Klasifikasi Data Mining	44
2.1.8 <i>Software Quality Assurance</i>	46
2.2 Tinjauan Studi	48
2.2.1 Penelitian oleh Tim Menzies, Jeremy Greenwald, dan Art Frank	48
2.2.2 Penelitian oleh Stefan Lessmann, Bart Baesens, Christophe Mues, dan Swantje Pietsch	50
2.2.3 Penelitian oleh Atac Deniz Oral, Ayse Basar Bener	51
2.2.4 Penelitian oleh N. Gayatri, S. Nickolas, dan A. V. Reddy	52
2.2.5 Penelitian oleh Qinbao Song, Zihan Jia, Martin Shepperd, Shi Ying, dan Jin Liu	54
2.2.6 Penelitian oleh Tracy Hall	56
2.2.7 Rangkuman penelitian terkait	57
2.3 Tinjauan Obyek Penelitian	58
2.3.1 Prediksi Software Defect	58
2.3.2 <i>Nasa Dataset</i>	61
2.4 Pola Pikir	63

2.4.1	Kerangka Konsep	63
2.4.2	Kerangka Pemikiran Pemecahan Masalah	64
2.5	Hipotesis.....	65
BAB III METODOLOGI DAN DESAIN PENELITIAN		66
3.1	Jenis Penelitian.....	66
3.2	Metode Pengumpulan Data	67
3.2.1	Data Sekunder	67
3.2.2	Metode Pemilihan Sampel.....	67
3.2.3	Studi Kepustakaan.....	67
3.3	Langkah-Langkah Penelitian.....	67
3.4	Desain Penelitian.....	69
3.4.1	<i>Business/Research Understanding Phase</i>	69
3.4.2	<i>Data Understanding Phase</i>	70
3.4.3	<i>Data Preparation Phase</i>	72
3.4.4	<i>Modeling Phase</i>	73
3.4.5	<i>Evaluation Phase</i>	80
3.4.6	<i>Deployment Phase</i>	80
3.5	Jadwal Penelitian.....	81
BAB IV PEMBAHASAN HASIL PENELITIAN.....		83
4.1	Pengukuran Penelitian.....	83
4.1.1	Hasil Penelitian.....	83
4.1.2	Pehitungan Pemilihan variable (<i>Genetic Algorithm</i>).....	83
4.1.3	Algoritma C4.5.....	95
4.1.4	Naive Bayes.....	97
4.1.5	Multi Layer Perceptron	99
4.2	Evaluasi dan Validasi	101
4.2.1	<i>Confusion Matrix</i> dan <i>ROC Curve</i> Algoritma C4.5 Optimasi <i>Genetic Algorithm</i>	101
4.2.2	<i>Confusion Matrix</i> dan <i>ROC Curve</i> Algoritma MLP Optimasi <i>Genetic Algorithm</i>	114
4.2.3	<i>Confusion Matrix</i> dan <i>ROC Curve</i> Algoritma Naïve Bayes Optimasi <i>Genetic Algorithm</i>	127
4.3	Hasil Komparasi.....	140
4.4	Penerapan Algoritma Terpilih.....	142
4.5	Implikasi Penelitian.....	147
4.5.1	Aspek Sistem.....	147
4.5.2	Aspek Manajerial	148
4.5.3	Aspek Penelitian Lanjut	148
BAB V PENUTUP.....		149
5.1	Kesimpulan.....	149
5.2	Saran.....	149
DAFTAR PUSTAKA		151
LAMPIRAN-LAMPIRAN.....		155
RIWAYAT HIDUP SINGKAT		177

DAFTAR GAMBAR

Gambar	Halaman
II-1 Langkah-langkah KDD ^[Maimon dan Rokach 2010]	10
II-2 Proses CRISP-DM	13
II-3 Klasifikasi Data Mining	15
II-4 Langkah-langkah dalam proses Data Mining ^[Maimon 2010]	16
II-5 Cabang pohon dibuat berdasarkan variasi nilai atribut	24
II-6 Cabang pohon keputusan yang dibuat dari nilai kontinyu	24
II-7 Cabang pohon yang terbentuk dari nilai biner	25
II-8 Pohon keputusan dengan cabang dari atribut A5 tidak konsisten.....	27
II-9 Pohon keputusan setelah di pruning.....	28
II-10 Keputusan setelah dicabangkan dengan <i>age</i>	30
II-11 Pemilihan titik perpecahan dalam atribut.....	30
II-12 Biological Neural Network	36
II-13 Fungsi Aktivasi Sigmoid.....	39
II-14 Neural Net yang dihasilkan oleh Multilayer Perceptron.....	40
II-15 Model yang diusulkan Tim Menzies.....	49
II-16 Model yang diusulkan Stefan Lessmann.....	51
II-17 Model yang diusulkan Gayatri	53
II-18 Framework yang diusulkan Qinqiao Song	55
II-19 Flowchart Software Development.....	58
II-20 Kerangka Pemikiran Pemecahan Masalah	64
III-1 Tahap CRISP-DM (Cross Standard Industries Process for Data Mining)....	68
III-2 Proses perbandingan metode dengan teknik optimisasi	76
III-3 Model yang diusulkan pada Rapid Miner 5.....	79
III-4 Cross Validation pada Rapid Miner 5	79
III-3-5 Metode pada cross validation pada Rapid Miner 5	79
IV-1 Pohon Keputusan Dengan <i>Dataset</i> JM1	97
IV-2 Model Multi Layer Perceptron dengan Rapid Miner	100
IV-3 Nilai AUC dengan Metode C4.5 pada <i>Dataset</i> CM1	102
IV-4 Nilai AUC dengan Metode C4.5 pada <i>Dataset</i> JM1.....	103

IV-5 Nilai AUC dengan Metode C4.5 pada Dataset KC1	104
IV-6 Nilai AUC dengan Metode C4.5 pada Dataset KC3	105
IV-7 Nilai AUC dengan Metode C4.5 pada Dataset MC1	106
IV-8 Nilai AUC dengan Metode C4.5 pada Dataset MC2	107
IV-9 Nilai AUC dengan Metode C4.5 pada Dataset MW1	108
IV-10 Nilai AUC dengan Metode C4.5 pada Dataset PC1	109
IV-11 Nilai AUC dengan Metode C4.5 pada Dataset PC2	110
IV-12 Nilai AUC dengan Metode C4.5 pada Dataset PC3	111
IV-13 Nilai AUC dengan Metode C4.5 pada Dataset PC4	112
IV-14 Nilai AUC dengan Metode C4.5 pada Dataset PC 5	113
IV-15 Nilai AUC dengan Metode MLP pada Dataset CM1	115
IV-16 Nilai AUC dengan Metode MLP pada Dataset JM1	116
IV-17 Nilai AUC dengan Metode MLP pada Dataset KC1	117
IV-18 Nilai AUC dengan Metode MLP pada Dataset KC3	118
IV-19 Nilai AUC dengan Metode MLP pada Dataset MC1	119
IV-20 Nilai AUC dengan Metode MLP pada Dataset MC2	120
IV-21 Nilai AUC dengan Metode MLP pada Dataset MW1	121
IV-22 Nilai AUC dengan Metode MLP pada Dataset PC1	122
IV-23 Nilai AUC dengan Metode MLP pada Dataset PC2	123
IV-24 Nilai AUC dengan Metode MLP pada Dataset PC3	124
IV-25 Nilai AUC dengan Metode MLP pada Dataset PC4	125
IV-26 Nilai AUC dengan Metode MLP pada Dataset PC5	126
IV-27 Nilai AUC dengan Metode Naive Bayes pada <i>Dataset</i> CM1	128
IV-28 Nilai AUC dengan Metode Naive Bayes pada Dataset JM1	129
IV-29 Nilai AUC dengan Metode Naive Bayes pada Dataset KC1	130
IV-30 Nilai AUC dengan Metode Naive Bayes pada Dataset KC3	131
IV-31 Nilai AUC dengan Metode Naive Bayes pada Dataset MC1	132
IV-32 Nilai AUC dengan Metode Naive Bayes pada Dataset MC2	133
IV-33 Nilai AUC dengan Metode Naive Bayes pada Dataset MW1	134
IV-34 Nilai AUC dengan Metode Naive Bayes pada Dataset PC1	135
IV-35 Nilai AUC dengan Metode Naive Bayes pada Dataset PC2	136

IV-36 Nilai AUC dengan Metode Naive Bayes pada Dataset PC3	137
IV-37 Nilai AUC dengan Metode Naive Bayes pada Dataset PC4	138
IV-38 Nilai AUC dengan Metode Naive Bayes pada Dataset PC5	139
IV-39 <i>Graphical User Interface (GUI) Aplikasi Software Defect Prediction....</i>	145

DAFTAR TABEL

Tabel	Halaman
II-1 Data Pembelian Komputer	29
II-2 Contoh data prediksi dengan Naive Bayes.....	34
II-3 Rata-rata dan variance dengan Naive Bayes	34
II-4 Nilai Awal untuk Input, Bobot, dan Bias.....	40
II-5 Perhitungan untuk bobot dan bias terbaru.....	43
II-6 Model Confusion Matrix ^[Bramer 2007]	45
II-7 Hasil penelitian Tim Menzies pada metode Naive Bayes.....	49
II-8 Performace Klasifikasi	52
II-9 Rangkuman tinjauan Studi	57
II-10 Variabel-variabel dalam NASA <i>dataset</i>	62
II-11 Komparasi <i>dataset</i> NASA MDP dengan PROMISE ^[Shepperd, et al. 2011]	63
III-1 Spesifikasi <i>dataset</i> yang digunakan.....	70
III-2 Variabel Setiap Dataset.....	71
III-3 Confusion Matrix.....	77
III-4 Deskripsi Confusion Matrix.....	77
III-5 Metric of Software Quality Assurance (SQA).....	80
III-6 Jadwal Penelitian	81
IV-1 Daftar Data perhitungan	86
IV-2 Pembobotan Variabel	87
IV-3 Pemilihan Variabel dengan Metode C4.5 + <i>Genetic Algorithm</i>	89
IV-4 Pemilihan Variabel dengan Metode Multi Layer Perceptron + <i>Genetic Algorithm</i>	91
IV-5 Pemilihan Variabel dengan Metode Naïve Bayes + <i>Genetic Algorithm</i>	93
IV-6 Hasil nilai info dan gain untuk menentukan simpul akar dengan data training	95
IV-7 Perhitungan rata-rata dan standart deviasi masing-masing atribut.....	98
IV-8 Probabilitas Posterior Data	98
IV -9 Confusion Matrix CM1 dengan metode C4.5	101
IV-10 Confusion Matrix JM1 dengan metode C4.5	102

IV-11 Confusion Matrix KC1 dengan metode C4.5	104
IV-12 Confusion Matrix KC3 dengan metode C4.5	105
IV-13 Confusion Matrix MC1 dengan metode C4.5	106
IV-14 Confusion Matrix MC2 dengan metode C4.5	107
IV-15 Confusion Matrix MW1 dengan metode C4.5	108
IV-16 Confusion Matrix PC1 dengan metode C4.5.....	109
IV-17 Confusion Matrix PC2 dengan metode C4.5.....	110
IV-18 Confusion Matrix PC3 dengan metode C4.5.....	111
IV-19 Confusion Matrix PC4 dengan metode C4.5.....	112
IV-20 Confusion Matrix PC5 dengan metode C4.5.....	113
IV-21 Confusion Matrix CM1 dengan metode Multi Layer Perceptron	114
IV-22 Confusion Matrix JM1 dengan metode Multi Layer Perceptron.....	115
IV-23 Confusion Matrix KC1 dengan metode Multi Layer Perceptron	117
IV-24 Confusion Matrix KC3 dengan metode Multi Layer Perceptron	118
IV-25 Confusion Matrix MC1 dengan metode Multi Layer Perceptron	119
IV-26 Confusion Matrix MC 2 dengan metode Multi Layer Perceptron	120
IV-27 Confusion Matrix MW1 dengan metode Multi Layer Perceptron	121
IV-28 Confusion Matrix PC1 dengan metode Multi Layer Perceptron.....	122
IV-29 Confusion Matrix PC2 dengan metode Multi Layer Perceptron.....	123
IV-30 Confusion Matrix PC3 dengan metode Multi Layer Perceptron.....	124
IV-31 Confusion Matrix PC4 dengan metode Multi Layer Perceptron.....	125
IV-32 Confusion Matrix PC5 dengan metode Multi Layer Perceptron.....	126
IV-33 Confusion Matrix CM1 dengan metode Naive Bayes.....	127
IV-34 Confusion Matrix JM1 dengan metode Naive Bayes.....	128
4-35 Confusion Matrix KC1 dengan metode Naive Bayes.....	130
IV-36 Confusion Matrix KC3 dengan metode Naive Bayes	131
IV-37 Confusion Matrix MC1 dengan metode Naive Bayes.....	132
IV-38 Confusion Matrix MC2 dengan metode Naive Bayes.....	133
IV-39 Confusion Matrix MW1 dengan metode Naive Bayes.....	134
IV-40 Confusion Matrix PC1 dengan metode Naive Bayes	135
IV-41 Confusion Matrix PC2 dengan metode Naive Bayes	136

IV-42 Confusion Matrix PC3 dengan metode Naive Bayes	137
IV-43 Confusion Matrix PC4 dengan metode Naive Bayes	138
IV-44 Confusion Matrix PC5 dengan metode Naive Bayes	139
IV-45 Perbandingan <i>Accuracy</i> Algoritma C4.5, MLP dan Naive Bayes dengan Optimisasi GA.....	140
IV-46 Perbandingan Nilai AUC Algoritma C4.5, MLP dan Naive Bayes dengan Optimisasi GA.....	141
IV-47 Data Baru Untuk Penerapan Algoritma Terpilih.....	143
IV-48 Confussion Matrix data baru dengan algoritma Naive Bayes	144
IV-49 Metric of Software Quality Assurance (SQA)	145
IV-51 Spesifikasi Hardware dan Software.....	147

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Contoh Data Set CM 1 (10 Data).....	155
2 Contoh Data Set JM 1 (10 Data).....	157
3. Contoh Data Set KC 1 (10 Data).....	158
4. Contoh Data Set KC 3 (10 Data).....	159
5. Contoh Data Set MC 1 (10 Data).....	161
6. Contoh Data Set MC 2 (10 Data).....	163
7. Contoh Data Set MW 1 (10 Data).....	165
8. Contoh Data Set PC 1 (10 Data)	167
9. Contoh Data Set PC 2 (10 Data)	169
10. Contoh Data Set PC 3 (10 Data)	171
11. Contoh Data Set PC 4 (10 Data)	173
12. Contoh Data Set PC 5 (10 Data)	175

DAFTAR PUSTAKA

- [Anbarasi 2010] Anbarasi, M., Anupriya, E., dan Iyengar, N. *Prediction of Heart Disease with Feature Subnet Selection using Genetic Algorithm*. International Journal Of engineering Science and Technology, 2010
- [Alpaydin 2010] Alpaydin, Ethem. *Introduction to Machine Learning 2nd*. London: The MIT Press, 2010.
- [Bramer 2007] Bramer, Max. *Principles of Data Mining*. London: Springer, 2007.
- [Challagulla] Challagulla, Venkata U. B., Farock B. Bastani, dan I Ling Yen. "Empirical Assessment of Machine Learning based Software Defect Prediction Techniques." *Proceedings of the 10th IEEE International Workshop on Object-Oriented Real-Time Dependable Systems (WORDS'05)*. Computer Society, 2005.
- [Chang 2007] Chang, Chingpao Pao, dan Chu Chih Ping. "An Action-Based Approach for Software Process Measurement." *Journal of Systems and Software*, 2007.
- [Chemuturi 2011] Chemuturi, Murali. *Mastering Software Quality Assurance*. Florida: J.Ross Publishing, 2011.
- [Dick 2005] Dick, S., dan A. Kandel. *Computational Intelligence in Software Quality Assurance*. World Scientific Publishing, 2005.
- [Fenton 1999] Fenton, N.E. "A Critique of Software Defect Prediction Models." *IEEE*, 1999: 1.
- [Fenton 1997] Fenton, N.E., dan S.L. Pfleeger. *Software Metrics, A Rigorous & Practical Approach*. London: International Thomson Computer Press, 1997.