

**KAJIAN PENERAPAN MODEL C45, *SUPPORT VECTOR MACHINE (SVM)*, DAN *NEURAL NETWORK* DALAM
PREDIKSI KENAIKAN KELAS:**

Studi Kasus SMK BINA TAQWA Cimanggis - Depok

TESIS



Oleh :

LUSI ARIYANI

1211600869

**PROGRAM STUDI : MAGISTER ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR
JAKARTA
2014**

**KAJIAN PENERAPAN MODEL C45, *SUPPORT VECTOR MACHINE (SVM)*, DAN *NEURAL NETWORK* DALAM
PREDIKSI KENAIKAN KELAS:**

Studi Kasus SMK BINA TAQWA Cimanggis - Depok

TESIS

**Diajukan sebagai salah satu persyaratan
memperoleh Gelar Magister Ilmu Komputer (M.Kom)**



**Oleh :
LUSI ARIYANI**

1211600869

**PROGRAM STUDI : MAGISTER ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR
JAKARTA
2014**



PROGRAM STUDI MAGISTER ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

LEMBAR PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : Lusi Ariyani

NIM : 1211600869

Program Studi : Magister Ilmu Komputer

Konsentrasi : Teknologi Sistem Informasi

Menyatakan bahwa TESIS yang berjudul :

KAJIAN PENERAPAN MODEL C45, SUPPORT VECTOR MACHINE (SVM) DAN NEURAL NETWORK DALAM PREDIKSI KENAIKAN KELAS :

Studi Kasus SMK BINA TAQWA Cimanggis - Depok

Merupakan hasil karya tulis ilmiah sendiri dan bukan merupakan karya yang pernah diajukan untuk memperoleh gelar akademi oleh pihak lain,

1. 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, 21 Agustus 2014





PROGRAM STUDI MAGISTER ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

LEMBAR PENGESAHAN

Nama : Lusi Ariyani
NIM : 1211600869
Program Studi : Magister Ilmu Komputer
Konsentrasi : Teknologi Sistem Informasi
Jenjang : Strata 2 (S2)
berjudul : Kajian Penerapan Model C45, *Support Vector Machine*
(SVM), Dan *Neural Network* Dalam Prediksi Kenaikan Kelas :
Studi Kasus SMK BINA TAQWA Cimanggis - Depok

Jakarta, Agustus 2014

Tim Penguji

Tanda tangan :

Ketua,
Dr. Moedjiono, M.Sc

Anggota,
Dr. Ir. Nazori AZ, MT

Pembimbing,
Dr. Ir. Prabowo Pudjo Widodo, MS

Ketua Program Studi

Dr. Ir. Nazori AZ, MT

ABSTRAK

Penilaian hasil belajar merupakan prediksi kenaikan kelas bagi siswa Sekolah Menengah Kejuruan (SMK) untuk melanjutkan ke tingkatan kelas berikutnya. Banyak mata pelajaran yang diujikan. Dari hasil mata pelajaran yang diujikan, maka sekolah dapat melihat rata-rata sehingga dapat diketahui siswa yang naik kelas atau tidak. Dengan cara tersebut siswa maupun guru atau wali kelas dapat memprediksi mata pelajaran dan nilai yang mempengaruhi kenaikan kelas. Selama ini Sekolah Menengah Kejuruan (SMK) Bina Taqwa belum memiliki pola-pola prediksi kenaikan kelas sebagai acuan untuk memprediksi jumlah siswa naik kelas dan yang tidak. Prediksi kenaikan kelas yang dilakukan saat ini masih manual, data yang diambil dari nilai semester siswa di akhir tahun. Prediksi hampir sama dengan klasifikasi yang akan terjadi di masa mendatang. Sehingga akan menghambat tentang peringkat sekolah dalam mengatasi kenaikan kelas siswa. Kendala tersebut dapat diatasi dengan analisis yang diuji menggunakan 3 buah metode algoritma yaitu algoritma C4.5, algoritma *Support Vector Mechine* dan *Neural Network*. Dari hasil pengujian dengan mengukur kinerja ketiga algoritma tersebut diketahui bahwa algoritma C45 memiliki nilai *accuracy* paling tinggi. Sehingga dapat diterapkan untuk permasalahan prediksi kenaikan kelas.

Kata Kunci : Siswa, Kenaikan Kelas, Algoritma C4.5, Algoritma *Support Vector Mechine*, *Neural Network*.

ABSTRACT

Evaluation of the result from student's studies could be an expectation for the student to go to the next step to continue the next grade at vocational high school. Too many subject ate to be done by students. From the result of the subject which being tested, school can get the average, then school will decided their students can continue to the next grade or not. The prediction for decided about students can go to the next grade or not till this time still in manual and data takes by the result from the end of semester. All predection almost the same with classification which will happen in the future it can be a constraint for the school to manage the rank to solve how to decided about the rank level for the student. The constraint can be solved with analysis which using 3 algorithm C45, algorithm Support Vector Machine and Neural Network. From the result of the research with analysis three of them we'll know that algorithma Support Vector Machine have high in accuration. Then we can use in class to solve the predection problem abaout students up to the next grade.

Keywords: *the students, will go to the next grade with, Algorithm C45, Algorithm Support Vector Mechine, Neural Network.*

KATA PENGANTAR

Puji syukur, penulis panjatkan kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan berkat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan tesis ini tepat pada waktunya. Dimana tesis ini penulis sajikan dalam bentuk buku yang sederhana. Adapun judul tesis, yang penulis ambil sebagai berikut “Kajian Penerapan Model C45, *Support Vector Machine*, Dan *Neural Network* Dalam Prediksi Kenaikan Kelas : Studi Kasus SMK Bina Taqwa Cimanggis – Depok”.

Tujuan penulisan tesis ini dibuat sebagai salah satu untuk mendapatkan gelar Magister Ilmu Komputer (M.Kom) pada Magister Ilmu Komputer Program Pascasarjana Budi Luhur.

Tesis ini diambil berdasarkan hasil penelitian atau riset mengenai “Kajian Penerapan Model C45, *Support Vector Machine*, Dan *Neural Network* Dalam Prediksi Kenaikan Kelas : Studi Kasus SMK Bina Taqwa Cimanggis – Depok”.

. Penulis juga lakukan mencari dan menganalisa berbagai macam sumber referensi, baik dalam bentuk jurnal ilmiah, buku-buku literatur, *internet*, dan lain-lain yang terkait dengan pembahasan pada tesis ini.

Penulis menyadari bahwa tanpa bimbingan dan dukungan dari semua pihak dalam pembuatan tesis ini, maka penulis tidak dapat menyelesaikan tesis ini tepat pada waktunya. Untuk itu izinkanlah penulis kesempatan ini untuk mengucapkan ucapan terima kasih yang sebesar-besarnya kepada :

1. Bapak Dr. Ir. Prabowo Pudjo Widodo, MS selaku pembimbing tesis yang telah menyediakan waktu, pikiran dan tenaga dalam membimbing penulis dalam menyelesaikan tesis ini.
2. Bapak Dr. Ir Nazori AZ, MT selaku Ketua Program Studi Pascasarjana Magister Ilmu Komputer Universitas Budi Luhur.
3. Bapak Dr. Moedjiono, M.Sc selaku Direktur Pascasarjana Magister Ilmu Komputer Universitas Budi Luhur.

4. Bapak Arya Sukarya, SPd selaku Kepala Sekolah SMK Bina Taqwa Cimanggis - Depok yang telah mengizinkan penulis melakukan riset untuk mendapatkan data atau informasi yang penulis butuhkan.
5. Achmad Sarwandianto M.Kom suami tercinta yang telah memberikan motivasi kepada saya.
6. Seluruh staf pengajar (dosen) Magister Ilmu Komputer Program Pascasarjana Budi Luhur yang telah memberikan pelajaran yang berarti bagi penulis selama menempuh studi.
7. Seluruh staf dan karyawan Magister Ilmu Komputer Program Pascasarjana Budi Luhur yang telah melayani penulis dengan baik selama kuliah.
8. Orang tua dan keluarga tercinta yang telah memberikan dukungan material dan moral kepada penulis.
9. Seluruh teman-teman kuliah Magister Ilmu Komputer Program Pascasarjana Budi Luhur yang selalu bekerja sama dan saling membantu penulis selama menempuh studi.
10. Serta semua pihak yang terlalu banyak untuk penulis sebutkan satu persatu sehingga terwujudnya penulisan tesis ini.

Penulis menyadari bahwa penulisan tesis ini masih jauh masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan karya ilmiah yang penulis hasilkan untuk yang akan datang.

Akhir kata semoga tesis ini dapat bermanfaat bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, 21 Agustus 2014

Lusi Ariyani

Penulis

DAFTAR ISI

	Halaman
LEMBAR PERNYATAAN	ii
LEMBAR PENGESAHAN	iii
ABSTRAK.....	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Masalah Penelitian	3
1.2.1 Identifikasi Masalah	3
1.2.2 Batasan Masalah	4
1.2.3 Rumusan Masalah	5
1.3 Tujuan dan Manfaat Penelitian	5
1.3.1 Tujuan	5
1.3.2 Manfaat	5
1.4 Sistematika Penulisan	6
1.5 Daftar Pengertian	7
BAB II LANDASAN TEORI DAN KERANGKA PEMIKIRAN	9
2.1 Tinjauan Pustaka	9
2.1.1 Prediksi	9
2.1.2 Pengertian Data Mining.....	9
2.1.3 Algoritma Klasifikasi Data Mining	13
2.1.4 Algoritma C4.5	13
2.1.4.1 Konsep Konsep Algoritma C4.5	13
2.1.4.2 Studi Kasus Algoritma C4.5	15
2.1.5 <i>Support Vector machine</i>	21
2.1.5.1 Contoh Penerapan <i>Support Vector Machine</i>	24
2.1.5.2 Evaluasi dan Validasi	25
2.1.6 <i>Neural Network</i>	28
2.1.6.1 Konsep <i>Neural Network</i>	28
2.6.1.2 Studi Kasus <i>Neural Network</i>	32
2.1.7 Evaluasi dan Validasi Metode Klasifikasi Data Mining..	35
2.1.8 Rapid Miner 5	37
2.2 Tinjauan Studi yang Relevan	38
2.3 Tinjauan Objek Penelitian	45

2.3.1 Sekolah Menengah Kejuruan Cimanggis – Depok.....	45
2.3.2 Struktur Organisasi SMK Bina Taqwa.....	45
2.3.3 Dukungan Sistem Informasi KBM SMK Bina Taqwa....	47
2.3.4 Sarana dan Prasarana SMK Bina Taqwa.....	47
2.4 kerangka Pemikiran	48
2.5 Hipotesis Penelitian	50
BAB III METODOLOGI dan RANCANGAN PENELITIAN	51
3.1 Jenis Penelitian	51
3.2 Metode Pemilihan Sampel	52
3.2.1 Populasi	52
3.2.2 Sampel	52
3.3 Metode Pengumpulan Data	53
3.4 Instrumentasi Penelitian	53
3.5 Teknik Analisis Data	54
3.6 Jadwal Penelitian	59
BAB IV HASIL ANALISA DAN PEMBAHASAN	61
4.1 Hasil Penelitian	61
4.2 Evaluasi dan Validasi	70
4.3 Perancangan Sistem	75
4.4 Implikasi Penelitian	76
BAB V PENUTUP	78
DAFTAR PUSTAKA	80
LAMPIRAN – LAMPIRAN	82
RIWAYAT HIDUP SINGKAT	87

DAFTAR GAMBAR

Gambar	Halaman
II-1 Proses CRISP-DM	10
II-2 Pohon Keputusan Hasil Perhitungan Node 1	18
II-3 Pohon Keputusan Hasil Perhitungan Node 1.1	19
II-4 Pohon Keputusan Pada Node 1.1.2	21
II-5 SVM Berusaha Menemukan <i>Hyperplane</i> Terbaik Yang Memisahkan Kedua Class Negatif Dan Positif 2	22
II-6 <i>Neural Network</i>	31
II-7 MLP	33
II-8 Tampilan layar <i>RapidMiner</i>	38
II-9 Struktur Organisasi SMK Bina Taqwa	46
II-10 Kerangka Pemikiran	49
III-1 Diagram Korelasi Antara Siswa/i Dengan Nilai	54
IV-1 simpul akar kepribadian	64
IV-2 simpul akar tanggungan orang tua	66
IV-3 Pohon Keputusan hasil perhitungan dengan metode C4.5	66
IV-4 Model Algoritma <i>Support Vector Machine</i> (SVM)	67
IV-5 Model Algoritma <i>Support Vector Machine</i> (SVM)	67
IV-6 Desain Model <i>Support Vector Machine</i>	68
IV-7 <i>Neural Network</i> yang dihasilkan dengan metode <i>neural network</i>	68
IV-8 Kurva ROC dengan algoritma C4.5	73
IV-9 Kurva ROC dengan Metode SVM	73
IV-10 Kurva ROC dengan Metode <i>Neural Network</i>	74
IV-11 Tampilan implementasi model Prediksi Kenaikan Kelas	76

DAFTAR TABEL

Tabel	Halaman
2.1 Keputusan Main Tennis	17
2.2 Perhitungan Node 1	18
2.3 Perhitungan Node 1. 1	20
2.4 Perhitungan Node 1.1.2	21
2.5 AND Problem	26
2.6 Matrik Klasifikasi untuk Model 2 Class	27
2.7 Nilai Awal untuk Input, Bobot, dan Bias	28
2.8 Perhitungan untuk bobot dan bias terbaru	37
2.9 Tabel Model Confision Matrix	38
2.10 Tinjauan Studi yang Relevan	43
3.1 Sampel	54
3.2 Atribut dan Nilai Kategori	57
3.3 Data Training	60
3.4 Jadwal Penelitian	62
4.1 nilai entropy dan gain untuk menentukan simpul akar	65
4.2 nilai entropy dan gain untuk menentukan akar 1.1	66
4.3 nilai bobot akhir untuk <i>hidden layer</i>	72
4.4 Nilai Bobot Akhir untuk <i>Output Layer</i>	71
4.5 Model confusion Matrix untuk C45	73
4.6 Model confusion matrik untuk metode SVM	74
4.7 Model Confusion Matrix untuk metode neural network	74
4.8 Komparasi Nilai Accuracy, Precision, dan Recall	75
4.9 Komparasi Nilai AUC	77
4.10 Komparasi Nilai Accuracy dan AUC	77
4.11 Tabel Nilai Akurasi dan Persisi Data Baru Algoritma SVM	78

DAFTAR PUSTAKA

- [Bramer. M. 2007] Bramer. M. 2007. *Principles of Data Mining*. London: Springer
- [Burges C. J. 1998]Burges C. J. 1998. *A Tutorial On Support Vector Machines For Pattern Recognition*. Boston: Kluwer Academic Publishers.
- [Dong. Y., Xia. Z., Tu. M., dan Xing. G. 2007] Dong. Y., Xia. Z., Tu. M., dan Xing. G. 2007. An Optimization Method For Selecting Parameters In Support Vector Machines. *Sixth International Conference On Machine Learning And Applications* , 1.
- [Gorunescu F. 2011] Gorunescu F. 2011. *Data Mining Concepts, Models and Techniques*. Berlin Heidelberg: Springer Verlag.
- [Han. J., dan Kember. M. 2006] Han. J., dan Kember. M. 2006. *Data Mining Concepts adn Techniques*. San Fransisco: Morgan Kauffman
- [Harry, Dhika. 2012] Harry, Dhika. 2012. Kajian Komparasi Penerapan Algoritma C4.5, Naïve Bayes, Dan Neural Network Dalam Pemilihan Mitra Kerja Penyedia Jasa Transportasi: Studi Kasus Cv. Viradi Global Pratama. Tesis., STIMIK Nusa Mandiri, 2012.
- [Huang. K., Yang. H., King. I., dan Lyu. M. 2008] Huang. K., Yang. H., King. I., dan Lyu. M. 2008. *Machine Learning Modeling Data Locally And Globally*. Berlin Heidelberg: Zhejiang University Press, Hangzhou And Springer-Verlag Gmbh.
- [Kusrini, dan Luthfi. T. E. 2009] Kusrini, dan Luthfi. T. E. 2009. *Algoritma Data Mining*. Yogyakarta: Penerbit Andi.
- [Larose, Daniel. T. 2005.]Larose, Daniel. T. 2005. *Discovering Knowledge in Data*. New Jersey: John Willey & Sons, Inc.
- [Liao 2007] Liao 2007, *Recent Advances in Data Mining of Enterprise Data: Algorithms and Application*, World Scientific Publishing, Singapore.
- [Maimon, Oded. 2005] Maimon, Oded. 2005, “*Data Mining and Knowledge discovery Handbook*”, Springer, Newyork.