

**APLIKASI PERBANDINGAN ALGORITMA PREDIKSI
KELULUSAN MAHASISWA MENGGUNAKAN C4.5 DAN
*SUPPORT VECTOR MACHINE (SVM)***

TUGAS AKHIR

Oleh:

RINTO MANURUNG (121110057)

YANTI SUSANNA SIREGAR (121111788)

IYEN LABENA SITUMORANG (121113185)



**PROGRAM STUDI TEKNIK INFORMATIKA
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
MIKROSKIL
MEDAN
2016**

**COMPARISON APPLICATION OF GRADUATION
PRECEDITION ALGORITHM USING C4.5 AND *SUPPORT
VECTOR MACHINE (SVM)***

FINAL RESEARCH



RINTO MANURUNG (121110057)

YANTI SUSANNA SIREGAR (121111788)

IYEN LABENA SITUMORANG (121113185)



**STUDY PROGRAM OF INFORMATICS ENGINEERING
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
MIKROSKIL
MEDAN
2016**

LEMBARAN PENGESAHAN

APLIKASI PERBANDINGAN ALGORITMA PREDIKSI
KELULUSAN MAHASISWA MENGGUNAKAN
C4.5 DAN *SUPPORT VECTOR MACHINE* (SVM)

TUGAS AKHIR

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Sarjana Strata Satu
Program Studi Teknik Informatika

Oleh:

RINTO MANURUNG (121110057)
YANTI SUSANNA SIREGAR (121111788)
IYEN LABENA SITUMORANG (121113185)

Disetujui oleh:

Dosen Pembimbing I,



Kristian Telaumbanua S.T., M.T.

Dosen Pembimbing II,



Andrew Sagitta Jauhari S.Kom., M.T.

Medan, 2 Agustus 2016
Diketahui dan disahkan Oleh:

Ketua Program Studi
Teknik Informatika

Hardy S.Kom., M.Sc.

ABSTRAK

Kelulusan mahasiswa adalah salah satu elemen penting dalam penilaian akreditasi perguruan tinggi. Salah satu teknik memprediksi kelulusan yang dapat diterapkan untuk mengetahui mahasiswa berpotensi lulus tepat waktu atau tidak terlambat adalah dengan menggunakan *Machine Learning*.

Dua algoritma dalam *machine learning* yaitu C4.5 dan *Support Vector Machine* (SVM) akan dibandingkan untuk memprediksi atribut yang sama dari data mahasiswa STMIK Mikroskil Medan periode 4 tahun terakhir yaitu Indeks Prestasi Sementara (IPS) semester 1 hingga semester 6. Algoritma C4.5 akan menghasilkan pohon keputusan yang mampu mengatasi atribut bersifat *continue*, mampu melakukan pemangkasan pohon, sementara *Support Vector Machine* akan menghasilkan *hyperplane* terbaik yang dapat memisahkan dua set data dari dua kelas berbeda.

Berdasarkan pengujian yang telah dilakukan secara *K-fold* dan *incremental*, algoritma C4.5 dan *Support Vector Machine* (SVM) memiliki tingkat akurasi yang berbeda. Algoritma C4.5 menghasilkan tingkat akurasi 78,43% dan *Support Vector Machine* (SVM) menghasilkan tingkat akurasi 83,65%. Dengan demikian, disimpulkan bahwa algoritma *Support Vector Machine* (SVM) memiliki tingkat akurasi yang lebih tinggi daripada C4.5.

Kata Kunci: *Machine Learning, Prediksi, C4.5, Support Vector Machine, Akurasi.*

UNIVERSITAS MIKROSKIL

KATA PENGANTAR

Puji dan syukur kepada Tuhan Yang Maha Esa, penulis dapat menyelesaikan Tugas Akhir yang berjudul *APLIKASI PERBANDINGAN ALGORITMA PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN C4.5 DAN SUPPORT VECTOR MACHINE (SVM)*, sesuai dengan yang direncanakan.

Untuk itu, penulis mengucapkan terimakasih kepada:

1. Bapak Kristian Telaumbanua S.T., M.T., selaku pembimbing I, yang telah membimbing penulis dalam menyelesaikan Tugas Akhir ini.
2. Bapak Andrew Sagitta Jauhari S.Kom., M.T., selaku pembimbing II, yang telah membimbing penulis dalam menyelesaikan Tugas Akhir ini.
3. Bapak Dr.Mimpin Ginting, M.S., selaku ketua STMIK Mikroskil Medan.
4. Bapak Djoni, S.Kom., M.T.I., selaku Wakil Ketua I STMIK Mikroskil Medan.
5. Bapak Hardy, S.Kom., M.Sc., selaku Ketua Program Studi Teknik Informatika STMIK Mikroskil Medan.
6. Seluruh Dosen pengajar Program Studi Teknik Informatika yang telah memberikan ilmu dan pengetahuan yang sangat berguna dalam menyelesaikan Tugas Akhir ini..
7. Terkhusus kepada kedua orang tua dan keluarga yang telah memberikan semangat, bimbingan dan dukungan baik moril maupun material selama penulis mengikuti pendidikan hingga terselesaikannya Tugas Akhir ini.
8. Seluruh teman-teman mahasiswa STMIK Mikroskil Medan.

Tugas akhir ini dibuat untuk melengkapi persyaratan guna memperoleh gelar Sarjana Strata Satu pada Program Studi Teknik Informatika, STMIK Mikroskil Medan. Semoga hasil dari Tugas Akhir ini berguna bagi pembaca dan yang berkepentingan.

Medan, Agustus 2016

Penulis

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR	ii
DAFTAR ISI.....	iii
DAFTAR GAMBAR	v
DAFTAR TABEL.....	viii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan dan Manfaat	2
1.4. Ruang Lingkup.....	2
1.5. Metodologi Penelitian	3
BAB II TINJAUAN PUSTAKA	6
2.1. Prediksi Kelulusan Mahasiswa.....	6
2.2. Prediksi.....	6
2.2.1. Prediksi Kualitatif	6
2.2.2. Prediksi Kuantitatif	7
2.3. Pohon Keputusan.....	7
2.3.1. Algoritma C4.5	10
2.3.2. Kelebihan Pohon Keputusan.....	12
2.4. <i>Supervised Learning</i>	13
2.5. <i>Support Vector Machine (SVM)</i>	13
2.5.1. Klasifikasi Data Linear Separable	15
2.5.2. Klasifikasi Data Linear <i>Non-Separable</i>	18
2.5.3. Metode Kernel	19
2.5.4. Karakteristik <i>Support Vector Machine</i>	20
2.5.5. Kelebihan <i>Support Vector Machine (SVM)</i>	21
2.6. Konsep Klasifikasi	22
2.6.1. Model	22
2.6.2. Pengukuran Kinerja Klasifikasi	23
2.6.3. <i>K-Fold Cross Validation</i>	24

BAB III ANALISIS DAN PERANCANGAN	26
3.1 Analisis.....	26
3.1.1. Analisis Proses	26
3.1.2. Analisis Kebutuhan.....	59
3.2 Perancangan	64
3.2.1. Perancangan Tampilan.....	64
3.2.2. Perancangan Basis Data.....	72
BAB IV HASIL DAN PENGUJIAN.....	74
4.1. Hasil	74
4.2. Pengujian.....	86
BAB V KESIMPULAN DAN SARAN	98
5.1. Kesimpulan.....	98
5.2. Saran.....	98
DAFTAR PUSTAKA	99
DAFTAR RIWAYAT HIDUP	51

UNIVERSITAS MIKROSKIL

DAFTAR GAMBAR

Gambar 2.1 Struktur <i>decision tree</i> (Pramudiono, 2008).....	8
Gambar 2.2 Margin <i>hyperplane</i> (Prasetyo, 2012)	15
Gambar 2.3 SVM linear separable (Prasetyo 2012)	16
Gambar 2.4 SVM non-linear separable (Prasetyo, 2012)	19
Gambar 3.1 <i>Flowchart</i> algoritma C4.5.....	26
Gambar 3.2 Pohon keputusan <i>node</i> 1	36
Gambar 3.3 Pohon keputusan <i>node</i> 2	41
Gambar 3.4 Perhitungan <i>entropy</i> dan <i>gain</i> <i>Node</i> IPS (B+)	43
Gambar 3.5 Pohon keputusan <i>node</i> 3 pada IPS2(B).....	46
Gambar 3.6 Hasil pohon keputusan	48
Gambar 3.7 <i>Flowchart</i> alur Kerja SVM	49
Gambar 3.8 <i>Use case</i> sistem prediksi kelulusan mahasiswa	60
Gambar 3.9 <i>Activity diagram</i> <i>training</i> data.....	61
Gambar 3.10 <i>Activity diagram</i> <i>testing</i> data	62
Gambar 3.11 Simpan hasil prediksi	63
Gambar 3.12 <i>Form</i> halaman utama	65
Gambar 3.13 <i>Form</i> halaman Awal.....	66
Gambar 3.14 <i>Form</i> halaman <i>training</i>	66
Gambar 3.15 <i>Form</i> <i>mock up</i> <i>testing</i>	67
Gambar 3.16 Tampilan hasil pohon keputusan C4.5.....	69
Gambar 3.17 Tampilan hasil <i>testing</i> <i>Support Vector Machine</i>	69
Gambar 3.18 Tampilan <i>about</i>	70
Gambar 3.19 <i>Form</i> halaman prediksi	71
Gambar 4.1 Halaman awal.....	74
Gambar 4.2 Tampilan halaman <i>incremental</i>	75
Gambar 4.3 Tampilan halaman <i>training</i> data	75
Gambar 4.4 Tampilan halaman <i>training</i> berhasil	76
Gambar 4.5 Tampilan halaman <i>testing</i> data	76

Gambar 4.6 Tampilan halaman <i>Input data testing</i>	77
Gambar 4.7 Tampilan halaman hasil <i>testing</i>	77
Gambar 4.8 Tampilan utama	78
Gambar 4.9 Tampilan halaman prediksi	78
Gambar 4.10 Tampilan <i>input data training</i>	79
Gambar 4.11 Tampilan pilih file yang sesuai dengan format	79
Gambar 4.12 Tampilan halaman format data tidak sesuai	80
Gambar 4.13 Tampilan proses <i>training</i> berhasil	80
Gambar 4.14 Tampilan <i>input data testing</i>	81
Gambar 4.15 Tampilan halaman proses <i>testing</i> berhasil	81
Gambar 4.16 Tampilan halaman proses prediksi	82
Gambar 4.17 Tampilan halaman simpan hasil prediksi	83
Gambar 4.18 Tampilan halaman simpan hasil prediksi berhasil	83
Gambar 4.19 Tampilan halaman hasil <i>training</i>	84
Gambar 4.20 Tampilan halaman hasil <i>training algoritma C4.5</i>	84
Gambar 4.21 Tampilan hasil <i>training Support Vector Machine</i>	85
Gambar 4.22 Tampilan halaman <i>about</i>	85
Gambar 4.23 Hasil <i>k-fold</i> pengujian 1	86
Gambar 4.24 Hasil <i>k-fold</i> pengujian 2	87
Gambar 4.25 Hasil <i>k-Fold</i> pengujian 3	87
Gambar 4.26 Hasil <i>k-fold</i> pengujian 4	88
Gambar 4.27 Hasil <i>k-fold</i> pengujian 5	89
Gambar 4.28 Hasil <i>k-fold</i> pengujian 6	89
Gambar 4.29 Hasil <i>k-fold</i> pengujian 7	90
Gambar 4.30 Hasil <i>k-fold</i> pengujian 8	90
Gambar 4.31 Hasil <i>incremental</i> pengujian 1	92
Gambar 4.32 Hasil <i>incremental</i> pengujian 2	92
Gambar 4.33 Hasil <i>incremental</i> pengujian 3	93
Gambar 4.34 Hasil <i>incremental</i> pengujian 5	93

Gambar 4.35 Hasil <i>incremental</i> pengujian 5	94
Gambar 4.36 Hasil <i>incremental</i> pengujian 6	95
Gambar 4.37 Hasil <i>incremental</i> pengujian 7	95
Gambar 4.38 Hasil <i>incremental</i> pengujian 8	96
Gambar 4.39 Hasil <i>incremental</i> pengujian 9	96
Gambar 4.40 Hasil <i>incremental</i> pengujian 10	97



DAFTAR TABEL

Tabel 1.1 Pengujian dengan <i>k-fold</i>	4
Tabel 1.2 Pengujian secara <i>incremental</i>	5
Tabel 2.1 <i>Matriks confusi</i> 1	23
Tabel 2.2 Pengujian <i>k-fold validation</i>	25
Tabel 3.1 Tabel <i>dataset</i> algoritma C4.5	27
Tabel 3.2 Kategorikal data.....	29
Tabel 3.3 Data hasil <i>pre-processing</i>	29
Tabel 3.4 Perhitungan nilai <i>entropy</i> dan <i>gain node</i> 1.....	34
Tabel 3.5 <i>Dataset</i> IPS1 dengan kategori data B	36
Tabel 3.6 Perhitungan <i>entropy</i> dan <i>gain node</i> 2	39
Tabel 3.7 Kategorikal data B+ pada IPS2.....	41
Tabel 3.8 Perhitungan <i>entropy</i> dan <i>gain node</i> 3 IPS2(B+).....	42
Tabel 3.9 <i>Dataset</i> IPS2 dengan kategorikal data B	44
Tabel 3.10 Perhitungan <i>entropy</i> dan <i>gain node</i> 3.....	45
Tabel 3.11 <i>Dataset</i> IPS2 kategorikal C.....	46
Tabel 3.12 Perhitungan <i>node</i> 3 IPS2(C)	47
Tabel 3.13 Tabel data <i>training Support Vector Machine (SVM)</i>	50
Tabel 3.14 Data <i>training</i> setelah <i>pre-processing</i>	50
Tabel 3.15 Perhitungan faktor kuadrat.....	50
Tabel 3.16 Lanjutan perhitungan faktor kuadrat.....	51
Tabel 3.17 Narasi <i>use case training</i> data.....	60
Tabel 3.18 Narasi <i>use case testing</i> data	61
Tabel 3.19 Narasi simpan hasil kelulusan.....	63
Tabel 3. 20 Tabel data <i>training</i>	73
Tabel 4.1 Hasil pengujian secara <i>k-fold</i>	91
Tabel 4.2 pengujian dengan data 25% <i>testing</i>	91
Tabel 4.3 pengujian dengan 50% data <i>testing</i>	94
Tabel 4.4 Hasil pengujian secara <i>incremental</i>	97