

**KOMPARASI METODE KLASIFIKASI *DATA MINING*
ALGORITMA DECISION TREE C4.5 DAN *SVM* UNTUK
PREDIKSI PENENTUAN BEASISWA KIP KULIAH : (STUDI
KASUS UNIVERSITAS MUHAMADIYAH PRINGSEWU)**

TESIS



Disusun Oleh:

**Agung Purwanto
1921211001**

**PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA
PROGRAM PASCASARJANA INSTITUT INFOTMATIKA &
BISNIS DARMAJAYA BANDAR LAMPUNG
2022**

**KOMPARASI METODE KLASIFIKASI *DATA MINING*
ALGORITMA DECISION TREE C4.5 DAN *SVM* UNTUK
PREDIKSI PENENTUAN BEASISWA KIP KULIAH : (STUDI
KASUS UNIVERSITAS MUHAMADIYAH PRINGSEWU)**

TESIS

**Diajukan sebagai Salah Satu Syarat untuk Mecapai Gelar
MAGISTER
Pada Program Studi Teknik Informatika
IIB Darmajaya Bandar Lampung**



Disusun Oleh:

**Agung Purwanto
1921211001**

**PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA
PROGRAM PASCASARJANA INSTITUT INFOTMATIKA &
BISNIS DARMAJAYA BANDAR LAMPUNG
2022**



Saya yang bertanda tangan di bawah ini, menyatakan bahwa Tesis yang saya ajukan ini adalah hasil karya saya sendiri, tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi atau karya yang pernah ditulis atau diterbitkan orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka. Karya ini adalah milik saya dan pertanggungjawaban sepenuhnya berada di pundak saya.

Bandar Lampung, 17 Januari 2023

Agung Purwanto
NPM. 1921211001

LEMBAR PERSETUJUAN TESIS

Judul Tesis : Komparasi Metode Klasifikasi *Data Mining*
Algoritma Decision Tree C4.5 Dan *Svm* Untuk Prediksi
Penentuan Beasiswa Kip Kuliah : (Studi Kasus
Universitas Muhamadiyah Pringsewu)

Nama Mahasiswa : Agung Purwanto

NPM : 1921211001

Program Studi : Magister Teknik Informatika

Telah diperiksa dan disetujui untuk diajukan dalam ujian tesis guna memperoleh gelar Magister Teknik Informatika pada program pascasarjana IIB Darmajaya

Bandar Lampung, 17 Januari 2023

Disetujui oleh:

Dosen Pembimbing

Ketua Program Studi

Dr. Handoyo W.N., S.Kom.,M.T.I
NIDN.

Joko Triloka, M.T.,Ph.D
NIDN. 0227077301

HALAMAN PENGESAHAN

Telah diuji dan dipertahankan didepan Tim Penguji Tugas Akhir
Program Studi Magister Teknik Informatika IIB Darmajaya dan dinyatakan
diterima untuk memenuhi syarat guna memperoleh gelar
Magister

Mengesahkan,

1. Tim Penguji

Tanda Tangan

Penguji I : **Dr. Sutedi, S.Kom., M.T.I.**

Penguji II : **Dr. Sri Lestari, S.Kom.,M.Cs**

2. Dekan Fakultas Ilmu Komputer

Dr. Sutedi, S.Kom., M.T.I.

NIK.00600303

Tanggal Lulus Ujian Tesis : 17 Januari 2023

DAFTAR RIWAYAT HIDUP

Data Pribadi

Nama : Agung Purwanto
Tempat, Tanggal Lahir : Palembang, 22 September 1987
Jenis Kelamin : Laki-laki
Agama : Islam
Kewarganegaraan : Indonesia
Alamat : Jl. Perwira II No.14 Tanjung Baru, Kedamaian,
Bandar Lampung
E-mail : purwanto.agung@gmail.com

Riwayat Pendidikan

- SD Negeri
- SMP Negeri
- SMA 2 Palembang
- S1 Sistem Informasi Teknokrat Lampung

Pengalaman Pekerjaan

- STIKes Muhammadiyah Pringsewu sekarang Menjadi Universitas Muhammadiyah Pringsewu

PERSEMBAHAN

Alhamdulillah, atas rahmat dan hidayah-Nya, saya dapat menyelesaikan Tesis ini dengan baik. Karya sederhana ini penulis persembahkan untuk:

1. Bapak dan Ibu tercinta, Bapak Bambang Listiono dan Ibu Marsini yang selalu memberikan do'a dan dukungan, memberi motivasi dan semangat serta memberikan kasih sayang yang besar yang tak mungkin bisa terbalas dengan apapun.
2. Apri Sulistianingsih istri tercinta serta anakku tersayang yang selalu menguatkan saat mulai melemah semangatnya.
3. Pak Handoyo Widi Nugroho selaku dosen yang telah membimbing sampai dengan proses selesainya tesis ini.
4. Teman Seperjuangan MTI

MOTTO

“Gagal itu adalah hal yang biasa, tapi kegagalan yang sesungguhnya adalah saat kita menyerah dan berhenti untuk mencoba.”

(Agung Purwanto)

PRAKATA

Dengan memanjatkan puji dan syukur ke hadirat Allah SWT atas segala rahmat dan karuniaNya penyusunan tesis yang berjudul "**Komparasi Metode Klasifikasi *Data Mining Algoritma Decision Tree C4.5 Dan Svm Untuk Prediksi Penentuan Beasiswa KIP Kuliah : (Studi Kasus Universitas Muhamadiyah Pringsewu)***" dapat saya selesaikan. Tesis ini ditulis dalam rangka memenuhi salah satu persyaratan memperoleh gelar Magister Teknik Informatika di program Pascasarjana IIB Darmajaya Lampung.

Dengan selesainya Tesis ini, saya mengucapkan terima kasih yang setulus-tulusnya kepada:

1. Bapak Dr. Ir. Firmansyah YA, M.B.A., M.Sc selaku Rektor IIB Darmajaya
2. Bapak Dr. Sutedi, S.Kom., MTI Selaku Dekan Fakultas Ilmu Komputer IIB Darmajaya
3. Bapak Joko Triloka, S.Kom., MT., Ph.D selaku ketua program studi Magister Teknik Informatika IIB Darmajaya.
4. Bapak Dr. Handoyo W.N. S.Kom.,M.T.I selaku pembimbing yang telah membimbing dan mengarahkan penulis selama penyusunan tesis ini dari awal hingga selesai.
5. Para Dosen Pascasarjana IIB Darmajaya yang telah membimbing saya dengan ilmu-ilmu yang bermanfaat dan memotivasi dalam penyelesaian tesis ini.

Saya berharap tesis ini mudah-mudahan dapat memberikan sumbangsih bagi pendidikan yang selalu menghadapi tantangan seiring dengan tuntutan zaman.

Bandar Lampung, 08 Juli 2022
Penyusun

Agung Purwanto

ABSTRAK

KOMPARASI METODE KLASIFIKASI DATA MINING ALGORITMA DECISION TREE C4.5 DAN SVM UNTUK PREDIKSI PENENTUAN BEASISWA KIP KULIAH : (STUDI KASUS UNIVERSITAS MUHAMADIYAH PRINGSEWU)

Disusun Oleh:

**Agung Purwanto
1921211001**

Program beasiswa Indonesia Pintar Kuliah diadakan untuk meringankan beban mahasiswa yang memiliki masalah biaya khususnya yang terdampak pandemi Covid-19 ini untuk melanjutkan studi ke tingkat pendidikan tinggi. Walaupun pihak pemerintah sudah membantu dengan menerbitkan Kartu Indonesia Pintar (KIP) Kuliah melalui program ini, namun kenyataannya tidak semua mahasiswa kurang mampu memiliki KIP Kuliah. Oleh sebab itu, pemberian beasiswa ini haruslah dilakukan secara selektif sesuai dengan kriteria atau faktor bobot penilaian yang sesuai dengan jenis beasiswa yang diadakan. Salah satunya beasiswa yang diberikan untuk mahasiswa yang berasal dari keluarga kurang mampu yang haruslah sesuai kemampuan ekonomi yang sebenarnya sehingga tidak kesulitan dalam pembiayaan perkuliahan. Sebelumnya masih dilakukan secara manual dengan menggunakan sorting pada *Microsoft excel* yang masih memiliki kekurangan dalam ketepatan hasil, sehingga beasiswa yang diberikan kurang tepat sasaran. *Data mining* adalah salah satu teknik yang bisa digunakan sebagai alternatif untuk membantu pengambilan keputusan pemberian beasiswa. Teknik *data mining* yang diterapkan adalah Klasifikasi. Metode Klasifikasi yang digunakan *Support Vector Machine* (SVM) dan *Decision tree* (pohon keputusan) dengan algoritma C4.5. Terdapat tiga parameter uji yang digunakan sebagai evaluasi sistem yaitu *accuracy*, *precision* dan *recall*. Dari hasil perbandingan tersebut antara metode SVM atau *Decision Tree*, Diduga yang memiliki akurasi yang lebih baik adalah *Decision tree* Algoritma C4.5.

Kata Kunci: Beasiswa, KIP Kuliah, *Data mining*, Algoritma *Decision tree* C4.5, SVM

DAFTAR ISI

Halaman

HALAMAN JUDUL	i
PERNYATAAN KEASLIAN LAPORAN TESIS	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
DAFTAR RIWAYAT HIDUP.....	v
HALAMAN PERSEMBAHAN	vi
MOTO.....	vii
PRAKATA	viii
INTISARI	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	ivx
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Batasan Masalah.....	4
1.3 Rumusan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Data mining.....	8
2.2 KDD.....	13

2.3 Klasifikasi	14
2.4 Algoritma C4.5.....	15
2.5 <i>Support Vector Machine</i>	18
2.6 <i>Evaluasi dan Validasi Metode Klasifikasi Data mining</i>	19
2.7 Tinjauan Studi.....	22
2.8 Tinjauan Obyek Penelitian.....	27
2.9 Beasiswa.....	36
2.10 Kerangka Konsep / Pola Pikir Pemecahan.....	36
2.11 Hipotesis.....	38
BAB III METODOLOGI PENELITIAN.....	31
3.1 Metode Penelitian.....	39
3.2 Deskripsi Data.....	41
3.3 Tahap Tahap Penelitian.....	42
3.4 Teknik Analisis, Desain, dan Pengujian	56
BAB IV HASIL DAN PEMBAHASAN	58
4.1 Bussiness Understanding atau Pemahaman Bisnis	58
4.2 Data Understanding atau Pemahaman Data.....	58
4.3 Data Preparation atau Persiapan Data	59
4.4 Pemodelan	60
4.5 Evaluasi	66
BAB V KESIMPULAN DAN SARAN.....	68
5.1 Kesimpulan	68
5.2 Saran.....	68

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR GAMBAR

	Halaman
Gambar 2.1. Prinsip <i>Data mining</i>	10
Gambar 2.2. Tahapan Metodologi KDD	13
Gambar 2.3 Plot Contoh Data	19
Gambar 2.4. Kerangka Konsep/Pola Pikir	37
Gambar 3.1 Tahapan Metodologi CRISP-DM	40
Gambar 3.2 Langkah-langkah Penelitian	44
Gambar 3.3 Bagan Alir Algoritma <i>SVM</i>	47
Gambar 3.4 Bagan Alir Algoritma <i>Decision tree</i>	51
Gambar 3.5 Sampel dataset	52
Gambar 3.6 Pohon Keputusan Root Node	54
Gambar 3.7 Pohon Keputusan Root Node	56
Gambar 4.1 Potongan Dataset	59
Gambar 4.2 Model Klasifikasi Algoritma C4.5	60
Gambar 4.3 Pohon Keputusan	61
Gambar 4.4 Text View pohon keputusan	61
Gambar 4.5 Confusion Matrix Hasil klasifikasi Algoritma C4.5	62
Gambar 4.6 Model Klasifikasi Algoritma SVM	63
Gambar 4.7 Hasil Pengujian Support Vector Machine	63
Gambar 4.8 Confusion Matrix Hasil klasifikasi Algoritma SVM	64
Gambar 4.9 Kernel Model Support Vector Machine	65
Gambar 4.10 Kurva AUC C4.5	65
Gambar 4.11 Kurva AUC SVM	66

DAFTAR TABEL

	Halaman
Tabel 2.1. Model <i>Confusion Matrix</i>	20
Tabel 2.2. Ringkasan Tinjauan Studi	25
Table 3.1 Atribut Data yang digunakan	42
Tabel 3.2 Data sample model klasifikasi.....	48
Tabel 3.3 Data sample baru.....	50
Tabel 3.4 Nilai Tiap Atribut.....	53
Tabel 3.5 Nilai Tiap Atribut.....	55
Tabel 4.1 Atribut Dataset	58
Tabel 4.2 Perbandingan Hasil Akurasi Algoritma C4.5.....	62
Tabel 4.3 Perbandingan Hasil Akurasi Algoritma SVM.....	64
Tabel 4.4 Perbandingan Hasil Penelitian	67