

SKRIPSI

**IMPLEMENTASI ALGORITMA C5.0 UNTUK PREDIKSI KETEPATAN
KELULUSAN MAHASISWA PENERIMA KIP-KULIAH MERDEKA
(STUDI KASUS UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA)**



DEWI WULANSARI

NIM : 195410031

**PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA
2023**

SKRIPSI

**IMPLEMENTASI ALGORITMA C5.0 UNTUK PREDIKSI KETEPATAN
KELULUSAN MAHASISWA PENERIMA KIP-KULIAH MERDEKA
(STUDI KASUS UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA)**

Diajukan sebagai salah satu syarat untuk menyelesaikan studi

Program Sarjana

Program Studi Informatika

Fakultas Teknologi Informasi

Universitas Teknologi Digital Indonesia

Yogyakarta

Disusun Oleh

DEWI WULANSARI

NIM : 195410031

PROGRAM STUDI INFORMATIKA

PROGRAM SARJANA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA

YOGYAKARTA

2023

HALAMAN PERSETUJUAN

UJIAN SKRIPSI

Judul : Implementasi Algoritma C5.0 untuk Prediksi Ketepatan Kelulusan Mahasiswa Penerima KIP-Kuliah Merdeka (Studi Kasus Universitas Teknologi Digital Indonesia)

Nama : Dewi Wulansari

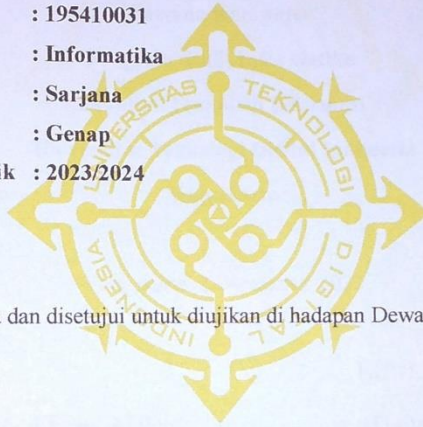
NIM : 195410031

Program Studi : Informatika

Program : Sarjana

Semester : Genap

Tahun Akademik : 2023/2024



Telah diperiksa dan disetujui untuk diujikan di hadapan Dewan Penguji Skripsi

Yogyakarta, _____

Dosen Pembimbing,

Ariesta Damavanti, S.Kom., M.Cs.

NIDN: 0020047801

HALAMAN PENGESAHAN

SKRIPSI

IMPLEMENTASI ALGORITMA C5.0 UNTUK PREDIKSI KETEPATAN
KELULUSAN MAHASISWA PENERIMA KIP-KULIAH MERDEKA
(STUDI KASUS UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA)

Telah dipertahankan di depan Dewan Penguji Skripsi dan dinyatakan
diterima untuk memenuhi sebagian persyaratan guna memperoleh Gelar
Sarjana Komputer

Program Studi Informatika
Fakultas Teknologi Informasi
Universitas Teknologi Digital Indonesia
Yogyakarta

Yogyakarta, 1 September 2023

Dewan Penguji



1. Maria Mediatrix, S.Kom., M.Eng.
2. Ariesta Damayanti, S.Kom., M.Cs.

NIDN

0514089101

0020047801

Tandatangan

Mengetahui

Ketua Program Studi Informatika


Dini Fakta Sari, S.T., M.T.
NIDN. 0507108401

PERNYATAAN KEASLIAN SKRIPSI

Dengan ini saya menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 21 Agustus 2023



Dewi Wulansari

NIM: 195410031

HALAMAN PERSEMBAHAN

Puji Syukur penulis panjatkan kepada Allah SWT, yang telah memberikan kesehatan, rahmat dan hidayah, sehingga penulis masih diberikan kesempatan untuk menyelesaikan skripsi ini sebagai salah satu syarat untuk mendapatkan gelar kesarjanaan. Shalawat serta salam semoga tetap tercurahkan kepada Nabi Muhammad SAW yang telah membawa kita dari zaman kebodohan menuju zaman hidayah seperti saat ini. Akhirnya terselesaikan juga tugas akhir ini dan untuk itu penulis ingin mempersembahkannya untuk orang-orang yang penulis cintai dan sayangi, yaitu:

1. Allah SWT yang melimpahkan rahmat dan hidayah-nya sehingga penulis dapat menyelesaikan skripsi ini.
2. Keluarga tercinta, bapak, ibu, dan adik penulis yang telah memberikan do'a, dukungan dan nasihat yang selalu menyertai setiap langkah hingga saat ini.
3. Teman-teman kos Ajib yang telah menjadi teman yang baik selama perkuliahan.
4. Keluarga besar WAMIKA, HIMAFORKA, dan Departemen Healing yang telah memberikan pengalaman, pelajaran, dan teman selama perkuliahan.
5. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah membantu terselesaikannya skripsi ini.

HALAMAN MOTTO

“Mulailah semuanya dengan BISMILLAH”

“Just do it!. And you can do!!!.”

KATA PENGANTAR

Segala puji syukur kami panjatkan kehadirat Tuhan Yang Maha Kuasa, karena oleh anugerah-Nya, kemurahan dan kasih setiaNya yang besar akhirnya penulis dapat menyelesaikan penulisan skripsi dengan judul: “Implementasi Algoritma C5.0 untuk Prediksi Ketepatan Kelulusan Mahasiswa Penerima KIP-Kuliah Merdeka (Studi Kasus Universitas Teknologi Digital Indonesia)”.

Penulis menyadari sepenuhnya bahwa skripsi ini masih jauh dari kesempurnaan karena menyadari segala keterbatasan yang ada. Untuk itu demi sempurnanya skripsi ini, penulis sangat membutuhkan dukungan dan sumbangsih pikiran yang berupa kritik dan saran yang bersifat membangun.

Penyusunan skripsi ini tidak lepas dari bantuan berbagai pihak, karena itu penulis menyampaikan ucapan terima kasih kepada:

1. Bapak Ir. Totok Suprawoto, M.M., M.T. Sebagai Rektor Universitas Teknologi Digital Indonesia.
2. Bapak Ir. Muhammad Guntara, M.T. selaku Dekan Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia.
3. Ibu Dini Fakta Sari, S.T., M.T dan Ibu Femi Dwi Astuti, S.Kom., M.Cs. Selaku Ketua dan Sekretaris Program Studi Informatika.
4. Ibu Ariesta Damayanti, S.Kom., M.Cs. Selaku pembimbing yang telah meluangkan waktunya untuk membimbing dan mengarahkan penulis.
5. Bapak Drs. Tri Prabawa, M.Kom. dan Ibu Maria Mediatrice, S.Kom., M.Eng. selaku Dosen Penguji yang memberikan saran, masukan dan arahannya.
6. Keluarga besar UKM WAMIKA dan Departemen Healing yang telah menjadi rumah dan memberikan banyak pelajaran dan pengalaman.

7. Sahabat dan teman-teman mahasiswa Universitas Teknologi Digital Indonesia yang tidak bisa penulis sebutkan satu persatu yang telah mendukung dan terlibat dalam membantu terselesaikannya skripsi ini.
8. Terimakasih kepada diri saya sendiri yang telah bertahan, berusaha tidak menyerah dan bismillah. *I am grateful that God gave me the strength not to give up.*

Kiranya skripsi ini dapat memberikan manfaat dan masukan bagi pembaca. Terima Kasih.

Yogyakarta, 21 Agustus 2022



Dewi Wulansari

NIM: 195410031

DAFTAR ISI

	Hal
SKRIPSI.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN	v
HALAMAN MOTTO	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
DAFTAR LAMPIRAN	xiii
INTISARI.....	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Ruang Lingkup.....	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI.....	7
2.1 Tinjauan Pustaka	7
2.2 Dasar Teori.....	11
2.2.1 Data Mining	11
2.2.2 Metode Prediksi Data Mining	11
2.2.3 Algoritma C5.0.....	12
2.2.4 Confusion Matrix	15
BAB III METODE PENELITIAN.....	17
3.1 Bahan/Data	17
3.2 Analisis Kebutuhan	17
3.2.1 Kebutuhan Perangkat Keras	17
3.2.2 Kebutuhan Perangkat Lunak	17
3.2.3 Kebutuhan Input.....	18

3.2.4	Kebutuhan Proses	18
3.2.5	Kebutuhan Output	18
3.3	Prosedur Pengumpulan Data	19
3.3.1	Sumber Data	19
3.3.2	Jenis Data	19
3.4	Analisis dan Rancangan Sistem	20
3.4.1	Flowchart	20
3.4.2	Use Case Diagram	29
3.4.3	Activity Diagram	31
3.4.4	Perancangan Antarmuka	32
BAB IV IMPLEMENTASI DAN PEMBAHASAN		34
4.1	Implementasi Sistem	34
4.1.1	Kode Program Algoritma C5.0	34
4.1.2	Kode program <i>Preprocessing</i> Data	36
4.1.3	Kode Program Prediksi	38
4.2	Uji Coba Sistem	39
4.3	Hasil dan Pembahasan Sistem	41
BAB V PENUTUP		44
5.1	Kesimpulan	44
5.2	Saran	44
DAFTAR PUSTAKA		46
LAMPIRAN		49

DAFTAR GAMBAR

	Hal
Gambar 2. 1 Confusion matrix dua kelas	16
Gambar 3. 1 Flowchart Proses Pelatihan	20
Gambar 3. 2 Pohon Keputusan.....	27
Gambar 3. 3 Flowchart Proses Pengujian	29
Gambar 3. 4 Use Case Diagram	30
Gambar 3. 5 Activity Diagram.....	32
Gambar 3. 6 Tampilan Halaman Prediksi Data	33
Gambar 3. 7 Tampilan Halaman Prediksi Algoritma C5.0.....	33
Gambar 4. 1 Kode Program Split Data	34
Gambar 4. 2 Kode Program Membuat Objek Dan Melatih Model.....	34
Gambar 4. 3 Kode Program Decision Tree.....	35
Gambar 4. 4 Kode Program Test Model	36
Gambar 4. 5 Kode Program Menghapus Kolom.....	36
Gambar 4. 6 Kode Program Mengganti Nilai NaN.....	37
Gambar 4. 7 Kode Program Normalisasi Data.....	37
Gambar 4. 8 Kode Program Clustering Data	38
Gambar 4. 9 Kode Program Pengolahan Prediksi Data	39
Gambar 4. 10 Uji Coba Prediksi Data.....	40
Gambar 4. 11 Tampilan Hasil Prediksi Data	40
Gambar 4. 12 Hasil Confusion Matrix Data Latih.....	41
Gambar 4. 13 Hasil Confusion Matrix Data Uji	42

DAFTAR TABEL

	Hal
Tabel 2. 1 Tinjauan Pustaka	9
Tabel 3. 2 Tabel Data Latih	23
Tabel 3. 3 Tabel Data Uji	23
Tabel 3. 4 Tabel Hasil Perhitungan.....	25
Tabel 3. 5 Deskripsi Use Case	31
Tabel 4. 1 Hasil Pengujian Prediksi Pada Data Uji.....	42

DAFTAR LAMPIRAN

	Hal
Lampiran 1 Tabel Data Latih	49
Lampiran 2 Tabel Data Uji	53
Lampiran 3 Tabel Hasil Prediksi	54
Lampiran 4 Catatan Seminar Pra Skripsi	59
Lampiran 5 Ketentuan Pendadaran.....	60
Lampiran 6 Catatan Pendadaran	61
Lampiran 7 Keputusan Hasil Ujian	62
Lampiran 8 ACC Revisi	63

INTISARI

Prediksi merupakan proses perkiraan atau estimasi terhadap nilai atau kejadian di masa depan berdasarkan data atau informasi yang tersedia saat ini. Salah satu algoritma yang dapat digunakan untuk melakukan Prediksi adalah algoritma C5.0. Algoritma C5.0 adalah pengembangan dari algoritma C4.5 dan memiliki keunggulan dalam tingkat akurasi yang lebih tinggi, kecepatan pengambilan keputusan, serta kemampuan mengatasi data hilang dan data dalam jumlah besar.

Data yang digunakan dalam penelitian ini adalah dataset mahasiswa penerima KIP-Kuliah Merdeka dari tahun 2020/2021 dan 2021/2022. Jumlah keseluruhan data adalah 195 dimana data ini digunakan sebagai data latih 80% dan data uji 20% berdasarkan atribut SKS, IPS1, IPS2, IPS3, IPS4 dan Label. Prediksi ini akan menentukan dua kriteria yaitu lulus tepat waktu dan tidak lulus tepat waktu.

Hasil akurasi yang didapatkan dari sistem Prediksi ini menggunakan Algoritma C5.0 sebesar 97% yang diukur menggunakan *confusion matrix*.

Kata Kunci: *Algoritma C5.0, Data Mining, Kelulusan, Prediksi, UTDI.*

ABSTRACT

Prediction is the process of estimating or estimating future values or events based on currently available data or information. One of the algorithms that can be used to make predictions is the C5.0 algorithm. The C5.0 algorithm is a development of the C4.5 algorithm and has advantages in a higher level of accuracy, speed of decision making, and the ability to overcome missing data and large amounts of data.

The data used in this study is a dataset of students who received KIP-Kuliah Merdeka from 2020/2021 and 2021/2022. The total amount of data is 195 where this data is used as 80% training data and 20% test data based on SKS, IPS1, IPS2, IPS3, IPS4 and Label attributes. This prediction will determine two criteria, namely passing on time and not passing on time.

The accuracy results obtained from this prediction system using the C5.0 algorithm is 97% which is measured using the confusion matrix.

Keywords: *C5.0 Algorithm, Data Mining, Graduation, Prediction, UTDI.*