

**PENERAPAN ALGORITMA DECISION TREE C4.5  
UNTUK PREDIKSI DATA PENERIMAAN SISWA  
BARU SMA TAMAN HARAPAN 1**

**SKRIPSI**

**Oleh :**

**YORI RICHARD SIAHAAN**

**201710225146**



**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS BHAYANGKARA JAKARTA RAYA  
2021**

## LEMBAR PERSETUJUAN PEMBIMBING

Judul Skripsi : PENERAPAN ALGORITMA DECISION TREE  
C4.5 UNTUK PREDIKSI DATA PENERIMAAN  
SISWA BARU SMA TAMAN HARAPAN 1

Nama Mahasiswa : Yori Richard Siahaan

Nomor Pokok Mahasiswa : 201710225146

Program Studi/Fakultas : Informatika / Ilmu Komputer

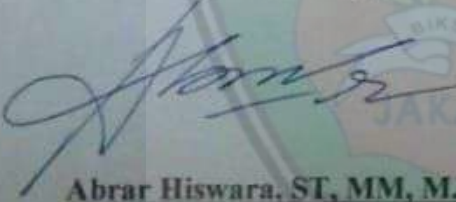
Tanggal Lulus Ujian Skripsi : 14 Juli 2021

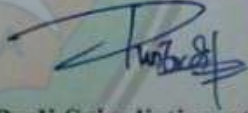
Bekasi, 19 Juli 2021

MENYETUJUI,

Pembimbing I

Pembimbing II

  
Abrar Hiswara, ST, MM, M.Kom  
NIDN. 0324028101

  
Dwi Budi Srisulistiowati, S.Kom, MM  
NIDN. 0323057701

## LEMBAR PENGESAHAN

Judul Skripsi : PENERAPAN ALGORITMA DECISION TREE  
C4.5 UNTUK PREDIKSI DATA PENERIMAAN  
SISWA BARU SMA TAMAN HARAPAN 1

Nama Mahasiswa : Yori Richard Siahaan

Nomor Pokok Mahasiswa : 201710225146

Program Studi / Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 14 Juli 2021

Bekasi, 19 Juli 2021

Mengesahkan,

Ketua Tim Penguji : Dian Hartanti, S.Kom., M.M.S.I  
NIDN. 0329098303

Penguji (I) : Mayadi, S.Kom., M.Kom  
NIDN. 0408087802

Penguji (II) : Abrar Hiswara, ST, MM, M.Kom  
NIDN. 0324028101

Mengetahui,

Ketua Program Studi  
Informatika

Dekan  
Fakultas Ilmu Komputer

Rakhmat Purnomo, S.Pd., S.Kom., M.Kom.  
NIDN. 0322108201

Herlawati, S.Si., M.M., M.Kom.  
NIDN. 0311097302



LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Yori Richard Siahaan  
NPM : 201710225146  
Program Studi : Informatika  
Fakultas : Ilmu Komputer  
Judul Tugas Akhir : PENERAPAN ALGORITMA DECISION TREE C4.5  
UNTUK PREDIKSI DATA PENERIMAAN SISWA  
BARU SMA TAMAN HARAPAN 1

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Bekasi, 6 Juni 2021

Penulis



Yori Richard Siahaan

## ABSTRAK

**Yori Richard Siahaan. 201710225146. PENERAPAN ALGORITMA DECISION TREE C4.5 UNTUK PREDIKSI DATA PENERIMAAN SISWA BARU SMA TAMAN HARAPAN 1**

Penerimaan siswa baru pada suatu institusi pendidikan merupakan sebuah kegiatan yang selalu dilaksanakan setiap tahun ajaran baru dimana calon siswa SMA Taman Harapan 1 Kota Bekasi tersebut selalu meningkat dari tahun ke tahun. Adapun masalah penelitian ini yaitu hasil penyeleksian yang tidak konsisten, kecurangan dan penilaian yang bersifat subjektif. Metode yang digunakan pada penelitian ini menggunakan algoritma C4.5 untuk memprediksi data penerimaan siswa baru. Langkah-langkah yang dilakukan penelitian ini untuk membuat klasifikasi terdiri dari normalisasi dataset, *crossvalidation*, pembangunan model *decision tree* C4.5, dan evaluasi. Penelitian ini menghasilkan tingkat akurasi pada algoritma C4.5 yang ada di pemrograman *python*.

Kata Kunci (*sentence case*): siswa baru, Algoritma C4.5, Decision Tree.

## ABSTRACT

**Yori Richard Siahaan. 201710225146. APPLICATION OF THE DECISION TREE C4.5 ALGORITHM FOR DATA PREDICTION OF NEW STUDENT ADMISSIONS IN SMA TAMAN HARAPAN 1**

*Acceptance of new students at an educational institution is an activity that is always carried out every new school year where prospective students of SMA Taman Harapan 1 Bekasi City always increase from year to year. The problems of this research are inconsistent selection results, fraud and subjective assessments. The method used in this study uses the C4.5 algorithm to predict new student admissions data. The steps taken in this study to make a classification consist of dataset normalization, crossvalidation, building a C4.5 decision tree model, and evaluation. This research produces an accuracy level on the C4.5 algorithm in python programming.*

*Keywords (sentence case): new students, Algorithm C4.5, Decision Tree.*

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI  
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK**

---

---

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Yori Richard Siahaan  
NPM : 201710225146  
Program Studi : Informatika  
Fakultas : Ilmu Komputer  
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (Non-Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul :

**PENERAPAN ALGORITMA DECISION TREE C4.5 UNTUK PREDIKSI  
DATA PENERIMAAN SISWA BARU SMA TAMAN HARAPAN 1**

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi  
Pada tanggal : 5 Juni 2021  
Yang Menyatakan

Yori Richard Siahaan

## KATA PENGANTAR

Segala Puji dan Syukur dengan tulus kami persembahkan ke hadirat Tuhan Yang Maha Esa, karena hanya dengan petunjuk hidayah-Nya peneliti mampu menyelesaikan tugas akhir yang berjudul PENERAPAN ALGORITMA DECISION TREE C4.5 UNTUK PREDIKSI DATA PENERIMAAN SISWA BARU SMA TAMAN HARAPAN 1.

Segala serta salam peneliti pada Puji dan Syukur yang memberikan motivasi bagi umat kristen, khususnya bagi peneliti untuk selalu berproses menuju insan yang memiliki intelektual tinggi dan berakhlak mulia.

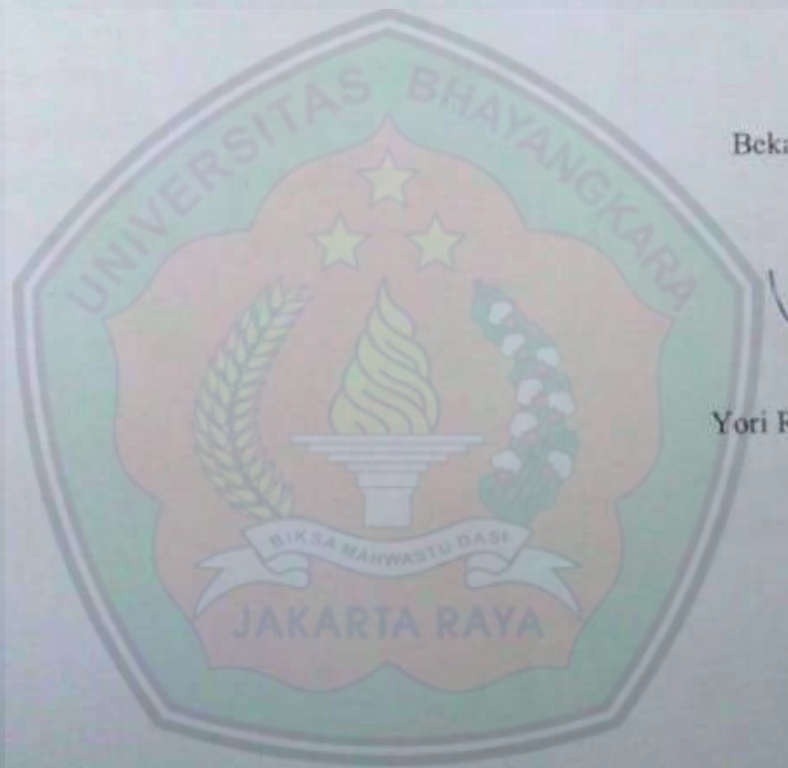
Penyelesaian skripsi ini merupakan suatu pekerjaan sangat berat bagi peneliti yang fakir ilmu, namun berkat dan Tuhan Yang Maha Esa dan bantuan dari berbagai pihak baik berupa materiil maupun moril, akhirnya tugas akhir ini dapat terselesaikan dengan baik. Oleh karena itu peneliti menyampaikan rasa hormat, ungkapan terima kasih serta penghargaan setinggi-tingginya kepada:

1. Irjen Pol. (Purn) Dr. H. Bambang Karsono, Drs., S.H, M.M. selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Herlawati, S.Si., M.M., M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Rakhmat Purnomo, S.Pd., S.Kom., M.Kom. selaku Program Studi Informatika Universitas Bhayangkara Jakarta Raya.
4. Abrar Hiswara, S.T, M.M, M.Kom. selaku dosen pembimbing I yang sabar memberikan arahan, saran dan motivasi pada peneliti sehingga skripsi ini dapat berjalan dengan lancar dan baik.
5. Dwi Budi Srisulistiowati, S.Kom, M.M. selaku dosen pembimbing II yang telah membimbing penulis dalam menyelesaikan skripsi ini.
6. Ibu Riama Nainggolan dan Bapak Johannes Siahaan yang telah waktu mendoakan peneliti demi kelancaran dan kesuksesan penulisan skripsi ini hingga tercapai cita-cita anaknya.



7. Seluruh dosen yang telah mengajarkan banyak hal dan selalu memberikan semangat untuk berproses hingga akhir perkuliahan peneliti.

Skripsi ini tentunya tiada terbebas dari kekurangan dan kelemahan. Oleh karena itu peneliti mengharapkan kritik dan saran penyempurnaan untuk perbaikan dimasa mendatang. Penulis berharap semoga skripsi ini dibaca oleh banyak orang, terutama civitas akademika Universitas Bhayangkara Jakarta Raya. Selain itu peneliti berharap semoga skripsi ini dapat memberikan nilai guna baik peneliti maupun bagi pembaca. Aamiin



Bekasi, 3 Juni 2021

Peneliti,

Yori Richard Siahaan

## DAFTAR ISI

	Halaman
<b>LEMBAR PERSETUJUAN PEMBIMBING .....</b>	<b>i</b>
<b>LEMBAR PENGESAHAN.....</b>	<b>ii</b>
<b>LEMBAR PERNYATAAN BUKAN PLAGIASI.....</b>	<b>iii</b>
<b>ABSTRAK.....</b>	<b>iv</b>
<b>ABSTRACT .....</b>	<b>v</b>
<b>LEMBAR PERNYATAAN PUBLIKASI.....</b>	<b>vi</b>
<b>KATA PENGANTAR .....</b>	<b>vii</b>
<b>DAFTAR ISI .....</b>	<b>ix</b>
<b>DAFTAR TABEL .....</b>	<b>xii</b>
<b>DAFTAR GAMBAR .....</b>	<b>xiii</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xv</b>
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah.....	3
1.3 Rumusan Masalah.....	3
1.4 Batasan Masalah .....	3
1.5 Tujuan Penelitian .....	3
1.6 Manfaat Penelitian .....	4
1.7 Tempat dan Waktu Penelitian.....	4
1.8 Metode Penelitian .....	4
1.9 Sistematika Penulisan .....	5
<b>BAB II LANDASAN TEORI.....</b>	<b>7</b>
2.1 Tinjauan Pustaka.....	7

2.2	Landasan Teori .....	15
2.3	Kerangka Pemikiran .....	39
2.4	Flowchart .....	40
2.5	Unified Modeling Language (UML) .....	44
<b>BAB III METODOLOGI PENELITIAN.....</b>		<b>52</b>
3.1	Pendekatan dan Jenis Pendekatan.....	52
3.2	Objek Penelitian.....	54
3.3	Spesifikasi.....	54
3.4	Analisa Permasalahan .....	55
3.5	Flowchart Berjalan dan Usulan .....	57
3.6	Metode Pengumpulan data .....	59
3.7	Teknik Analisis Data .....	60
3.8	Visi dan Misi serta Tujuan SMA Taman Harapan 1 Kota Bekasi.....	65
3.9	Analisa Data.....	66
<b>BAB IV PERANCANGAN SISTEM DAN IMPLEMENTASI.....</b>		<b>67</b>
4.1	Prosedur Sistem Usulan.....	67
4.2	Bussiness Understanding .....	70
4.3	Data Understanding .....	71
4.4	Data Preparation .....	72
4.5	Perhitungan Manual Algoritma C4.5.....	74
4.6	Hasil dan Pembahasan <i>Python</i> Algoritma C4.5.....	89
4.7	Confusion Matrix.....	99
<b>BAB V PENUTUP .....</b>		<b>101</b>
5.1	Kesimpulan .....	101
5.2	Saran .....	101

<b>DAFTAR PUSTAKA .....</b>	<b>102</b>
<b>LAMPIRAN .....</b>	<b>106</b>



## DAFTAR TABEL

	Halaman
Tabel 2.1 Penelitian Sebelumnya .....	7
Tabel 2.2 Data Buys Computer .....	26
Tabel 2.3 Simbol-Simbol Flowchart .....	42
Tabel 2.4 Simbol Use Case.....	45
Tabel 2.5 Simbol Activity Diagram .....	47
Tabel 2.6 Simbol Sequence Diagram .....	49
Tabel 2.7 Simbol Class Diagram.....	50
Tabel 3.1 Atribut Data Penelitian.....	62
Tabel 3.2 Atribut yang akan digunakan.....	64
Tabel 4.1 Data awal.....	71
Tabel 4.2 Data yang sudah di Transformasi.....	73
Tabel 4.3 Dataset Penerimaan Siswa Baru SMA Taman Harapan 1.....	74
Tabel 4.4 Dataset Penerimaan Siswa Baru.....	89
Tabel 4.5 Pengujian Confusion Matrix.....	100

## DAFTAR GAMBAR

	Halaman
Gambar 1.1 Diagram Model yang diusulkan.....	5
Gambar 2.1 Tahapan Data Mining .....	16
Gambar 2.2 Konsep <i>Decision Tree</i> .....	21
Gambar 2.3 Model Pohon Keputusan .....	23
Gambar 2.4 <i>Decision Tree</i> Node 1 (root node) .....	31
Gambar 2.5 <i>Decision Tree</i> Node 1 (root node dan leaf node).....	31
Gambar 2.6 <i>Decision Tree</i> Node 1.1 .....	34
Gambar 2.7 <i>Decision Tree</i> Node dan Leaf Node 1.1 .....	34
Gambar 2.8 <i>Decision Tree</i> Node 1.2 .....	37
Gambar 2.9 <i>Decision Tree</i> final .....	37
Gambar 2.10 Kerangka Pemikiran .....	39
Gambar 2.11 Diagram UML .....	44
Gambar 3.1 Sekolah SMA Taman Harapan 1 .....	54
Gambar 3.2 Use Case Penerimaan Siswa Baru .....	56
Gambar 3.3 Flowchart berjalan dan usulan .....	57
Gambar 3.4 Flowchart Sistem Usulan .....	59
Gambar 4.1 Use Case Diagram pada Penerimaan Siswa Baru.....	67
Gambar 4.2 Activity Diagram pada Penerimaan Siswa Baru .....	68
Gambar 4.3 Sequence Diagram pada Penerimaan Siswa baru .....	69
Gambar 4.4 Class Diagram pada Penerimaan Siswa Baru .....	70
Gambar 4.5 Hasil Algoritma C4.5 (Manual).....	89
Gambar 4.6 Proses Import Data dalam bentuk .csv .....	95
Gambar 4.7 Menentukan Format Data .csv .....	96
Gambar 4.8 Melakukan label encoder .....	96
Gambar 4.9 Hasil <i>Decision Tree</i> .....	97
Gambar 4.10 Mengimport package dan mengimport data .csv .....	98
Gambar 4.11 Menampilkan tingkat akurasi pada algoritma C4.5 .....	98
Gambar 4.12 Mengambil tabel yang pertama .....	98

Gambar 4.13 Melihat tingkat akurasi kembali pada algoritma C4.5.....99

Gambar 4.14 Perbandingan keterangan hasil dan prediksi pada Python.....99



## DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Hasil Penelitian dengan kuesioner .....	119
Lampiran 2 Surat Penelitian SMA Taman Harapan 1 .....	120
PLAGIARISME .....	121
BIODATA MAHASISWA .....	122
KARTU BIMBINGAN SKRIPSI .....	123

