



**PENERAPAN ALGORITMA C4.5 DAN LINEAR
REGRESSION UNTUK MEMREDIKSI DINI GANGGUAN
AUTISME**

LAPORAN SKRIPSI

HAFID GHULAM SHARWAR

41519010020

**UNIVERSITAS
MERCU BUANA**

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS

ILMU KOMPUTER

UNIVERSITAS MERCU BUANA JAKARTA

2023



**PENERAPAN ALGORITMA C4.5 DAN LINEAR
REGRESSION UNTUK MEMPREDIKSI DINI GANGGUAN
AUTISME**

LAPORAN SKRIPSI

HAFID GHULAM SHARWAR

41519010020

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS
ILMU KOMPUTER
UNIVERSITAS MERCU BUANA JAKARTA
2023**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan dibawah ini:

Nama : Hafid Ghulam Sharwar
NIM : 41519010020
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Penerapan Algoritma C4.5 dan Linear Regression Untuk Memprediksi Dini Gangguan Autisme.

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

UNIVERSITAS
MERCU BUANA

Jakarta, 14 Juni 2023
1000
METERAI
TEMPEL
74852AKX349252972
Hafid Ghulam Sharwar


HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Hafid Ghulam Sharwar
NIM : 41519010020
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Penerapan Algoritma C4.5 dan Linear Regression
Untuk Memprediksi Dini Gangguan Autisme

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan Oleh:

Pembimbing : Muhaimin Hasanudin S.T., M.Kom ()
NIDN : 0420027508
Ketua Penguji : Eliyani, Dr. Ir ()
NIDN : 0321026901
Penguji 1 : Roy Mubarak S.T, M.Kom ()
NIDN : 0310027402

Jakarta, 20 Juli 2023

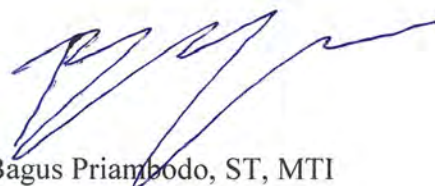
Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jekonowo, S.Si,MTI



Bagus Priambodo, ST, MTI

KATA PENGANTAR

Dengan penuh rasa syukur dan terima kasih, saya ingin mengucapkan puji kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya berhasil menyelesaikan laporan skripsi yang berjudul "**Penerapan Algoritma C4.5 dan Linear Regression Untuk Memprediksi Dini Gangguan Autisme**" Ini dilakukan sebagai salah satu syarat untuk mendapatkan gelar Sarjana Ilmu Komputer dari Fakultas Ilmu Komputer Universitas Mercu Buana Jakarta. Saya menyadari bahwa akan sangat sulit bagi saya untuk menyusun laporan skripsi ini dari masa perkuliahan hingga selesai tanpa bantuan dan bimbingan dari berbagai pihak. Akibatnya, saya ingin mengucapkan terima kasih kepada :

1. Bapak Muhaimin Hasanudin, S.T., M.KOM selaku dosen pembimbing skripsi.
2. Bapak Prof. Dr. Ir. Andi Adriansyah, M.Eng selaku Rektor Universitas Mercu Buana.
3. Bapak Dr. Bambang Jokonowo, S.Si., M.Kom selaku Dekan Fakultas Ilmu Komputer.
4. Bapak Bagus Priambodo, ST., M.T.I selaku Ketua Program Studi Teknik Informatika.
5. D. Bapak Indra Ranggadara, S.Kom., MT., MMSI. Selaku Kordinator Tugas Akhir Jurusan Teknik Informatika.Seluruh Dosen pengajar dan Staff Fakultas Ilmu Komputer Universitas Mercu Buana.
6. Orang tua saya, Ayah Yusup dan Ibunda Ita Rosita, selalu mendoakan dan mendukung saya.
7. Teman-teman Universitas Mercu Buana Bidang Studi Teknik Informatika 2019.
8. Para sahabat, Eko Prasetyo, Adam Firdhaus, Muhammad Deva, Nanda Kusuma, Fajrun Nadhif, Aryadhuna Randum, Yudha Eric, Fadly Lutfian dan Rahmat Widayat dari semester 1 hingga saat ini.
9. PEMMZ ID, selaku tempat intern saya yang telah memberikan saya support dan fasilitas, terutama pada tim Pemmz R&D.
10. Calista Nathania Gumanto, yang selalu membantu dan mendoakan saya, serta menjadi penyemangat saya.

Akhir kata, saya ingin meminta maaf atas segala kekurangan dalam laporan ini, baik dari segi isi maupun presentasi, karena terbatasnya pengetahuan dan pengalaman saya. Oleh karena itu, saya sangat mengharapkan saran dan kritik yang konstruktif untuk dijadikan pedoman di masa mendatang. Semoga skripsi ini dapat memberikan manfaat dan pengetahuan bagi semua.

Jakarta, Juni 2023

Hafid Ghulam Sharwar



HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Hafid Ghulam Sharwar
NIM : 41519010020
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Penerapan Algoritma C4.5 dan Linear Regression Untuk Memprediksi Dini Gangguan Autisme

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalty-Free Right)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS
MERCU BUANA

Jakarta, 14 Juni 2023

Yang menyatakan bahwa,



Hafid Ghulam Sharwar

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENYATAAN KARYA SENDIRI.....	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	vi
ABSTRAK.....	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN.....	xv
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Penelitian	5
BAB II.....	6
TINJAUAN PUSTAKA	6
2.1 Penelitian Terdahulu	6
2.2 Teori Pendukung	13
2.2.1 Autisme	13
2.2.2 Prediksi	14
2.2.3 Data Mining	14

2.2.4	Python	14
2.2.5	Machine Learning	15
2.2.6	Algoritma C4.5	16
2.2.7	Algoritma <i>Linear Regression</i>	16
2.2.8	Scikit-learn.....	18
2.2.9	RMSE dan MAE.....	19
2.2.10	<i>Accuracy</i>	20
BAB III		21
METODE PENELITIAN		21
3.1	Jenis Penelitian.....	21
3.2	Metode Pengumpulan Data.....	21
3.3	Tahapan Penelitian	22
3.3.1	<i>Identification</i>	23
3.3.2	<i>Exploration Data</i>	23
3.3.3	<i>Preparation Data</i>	24
3.3.4	<i>Modeling</i>	24
3.3.5	Metode <i>Linear Regression</i>	25
3.3.6	Metode C4.5.....	25
3.3.7	<i>Model Evaluation</i> dan Hasil.....	26
3.3.8	Komparasi Hasil.....	27
BAB IV		28
HASIL DAN PEMBAHASAN.....		28
4.1	Exploration Data	28
4.2	Preparation Data.....	31
4.2.1	<i>Data Cleaning</i>	34
4.2.2	<i>Data Transformation</i>	37
4.2.3	<i>Split Data</i>	38

4.3	Modeling	40
4.3.1	<i>Modeling Linear Regression</i>	40
4.3.2	<i>Modeling C4.5</i>	41
4.4	Visualisasi	42
4.4.1	Visualisasi Distribusi Column Age.....	44
4.4.2	Visualisasi Distribusi Column Gender.....	46
4.5	Analisis Hasil	48
4.5.1	Model Evaluation Linear Regression.....	48
4.5.2	Model Evaluation C4.5	55
4.6	Komparasi Hasil	61
BAB V	69
KESIMPULAN DAN SARAN	69
5.1	Kesimpulan.....	69
5.2	Saran.....	69
DAFTAR PUSTAKA	70
LAMPIRAN	73



DAFTAR TABEL

TABEL 2. 1 PENELITIAN TERKAIT	6
TABEL 4. 1 HASIL SPLIT DATA 3 KATEGORI.....	36
TABEL 4. 2 KOMPARASI DENGAN BEBERAPA SPLIT DATA.....	57



DAFTAR GAMBAR

GAMBAR 1. 1 AUTISM SCREENING ADULT DATA SET	2
GAMBAR 2. 1 RUMUS SELISIH NILAI PREDIKSI DAN NILAI OBSERVASI ..	17
GAMBAR 2. 2 RUMUS NILAI ABSOLUT	17
GAMBAR 2. 3 RUMUS NILAI MSE.....	18
GAMBAR 2. 4 RUMUS NILAI MAE.....	18
GAMBAR 2. 5 RUMUS NILAI RMSE.....	18
GAMBAR 3. 1 TAHAP PENELITIAN	22
GAMBAR 4. 1 PENJELASAN DATA	28
GAMBAR 4. 2 SOURCE CODE VISUALISASI RAW DATA	28
GAMBAR 4. 3 OUTPUT RAW DATA.....	25
GAMBAR 4. 4 SOURCE CODE VISUALISASI RAW DATA	26
GAMBAR 4. 5 VISUALISASI RAW DATA BAR PLOT.....	27
GAMBAR 4. 6 SOURCE CODE PREAPARATON DATA	27
GAMBAR 4. 7 OUTPUT RINCIAN DATA	28
GAMBAR 4. 8 SOURCE CODE DATA CLEANING.....	30
GAMBAR 4. 9 OUTPUT SEBELUM DATA CLEANING.....	30
GAMBAR 4. 10 SOURCE CODE DATA CLEANING.....	31
GAMBAR 4. 11 OUTPUT SETELAH DATA CLEANING.....	32
GAMBAR 4. 12 TRANSFORM DATA	33
GAMBAR 4. 13 VISUALISASI TRANSFORM DATA.....	34
GAMBAR 4. 14 SOURCE CODE SPLIT DATA KATEGORI 80 : 20.....	34
GAMBAR 4. 15 <i>SOURCE CODE SPLIT DATA</i> KATEGORI 80:20.....	35
GAMBAR 4. 16 SOURCE CODE SPLIT DATA KATEGORI 60:40.....	35
GAMBAR 4. 17 <i>MODELING LINEAR REGRESSION</i>	36
GAMBAR 4. 18 MODELING C4.5	37
GAMBAR 4. 19 <i>SOURCE CODE VISUALISASI</i>	38
GAMBAR 4. 20 VISUALISASI KESELURUHAN.....	39
GAMBAR 4. 21 <i>SOURCE CODE VISUALISASI DISTRIBUSI COLUMN AGE</i>	40
GAMBAR 4. 22 VISUALISASI DISTRIBUSI <i>COLUMN AGE</i>	41
GAMBAR 4. 23 <i>SOURCE CODE VISUALISASI DISTRIBUSI COLUMN GENDER</i>	42

GAMBAR 4. 24 VISUALISASI DISTRIBUSI <i>COLUMN GENDER</i>	43
GAMBAR 4. 25 <i>MODEL EVALUATION LINEAR REGRESSION</i>	44
GAMBAR 4. 26 HASIL <i>ACCURACY LINEAR REGRESSION</i>	45
GAMBAR 4. 27 <i>MODEL EVALUATION MAE DAN RMSE LINEAR REGRESSION</i>	46
GAMBAR 4. 28 HASIL <i>MAE DAN RMSE LINEAR REGRESSION</i>	46
GAMBAR 4. 29 SOURCE CODE R2 SCORE.....	47
GAMBAR 4. 30 OUTPUT R2 SCORE.....	47
GAMBAR 4. 31 <i>SOURCE CODE VISUALISASI SCATTERPLOT MODEL LINEAR</i> <i>REGRESSION</i>	47
GAMBAR 4. 32 <i>VISUALISASI SCATTERPLOT MODEL LINEAR REGRESSION</i> ..	48
GAMBAR 4. 33 <i>SOURCE CODE VISUALISASI MODEL BARPLOT LINEAR</i> <i>REGRESSION</i>	49
GAMBAR 4. 34 <i>VISUALISASI BARPLOT MODEL LINEAR REGRESSION</i>	50
GAMBAR 4. 35 <i>SOURCE CODE MODEL EVALUATION C4.5</i>	51
GAMBAR 4. 36 <i>HASIL ACCURACY MODEL EVALUATION C4.5</i>	52
GAMBAR 4. 37 <i>SOURCE CODE VISUALISASI PLOTTREE MODEL C4.5</i>	52
GAMBAR 4. 38 <i>VISUALISASI PLOTTREE MODEL C4.5</i>	48
GAMBAR 4. 39 <i>ROOTNODE</i>	49
GAMBAR 4. 40 <i>CABANG TRUE</i>	50
GAMBAR 4. 41 <i>SOURCE CODE KOMPARASI HASIL</i>	51
GAMBAR 4. 42 HASIL KOMPARASI MODEL.	52
GAMBAR 4. 43 RUMUS MODEL LINEAR REGRESSION	53

DAFTAR LAMPIRAN

LAMPIRAN 1. BIMIBINGAN	63
LAMPIRAN 2. ACC DOSEN PEMBIMBING.....	64
LAMPIRAN 3. BUKTI SUBMIT JURNAL	65
LAMPIRAN 4. HALAMAN PERNYATAAN LUARAN TUGAS AKHIR	66
LAMPIRAN 5. HALAMAN PERSETUJUAN.....	67
LAMPIRAN 6. SURAT PERNYATAAN HAKI	68
LAMPIRAN 7. NASKAH JURNAL.....	70
LAMPIRAN 8. CURRICULUM VITAE	80
LAMPIRAN 9. TURNITIN.....	82

