

**IMPLEMENTASI ALGORITMA DECISION TREE C
4.5 UNTUK MEMPREDIKSI PENDAPATAN
TAHUNAN**

Oleh

Ryan Salman Gunadharna

(207006516065)



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
2024**

**IMPLEMENTASI ALGORITMA DECISION TREE C
4.5 UNTUK MEMPREDIKSI PENDAPATAN
TAHUNAN**

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Sistem
Informasi dari Fakultas Teknologi Komunikasi dan Informatika

Oleh

Ryan Salman Gunadharna

(207006516065)



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
2024**

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

TUGAS AKHIR

IMPLEMENTASI ALGORITMA DECISION TREE C 4.5 UNTUK MEMPREDIKSI PENDAPATAN TAHUNAN



Dosen Pembimbing 1

Sari Ningsih, S.Si., M.M

NIDN. 0302066701

Dosen Pembimbing 2

Gatot Soepriyono, S.Si., M.S.M

NIDN. 0317047803

HALAMAN KEASLIAN TUGAS AKHIR

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul:

IMPLEMENTASI ALGORITMA DECISION TREE C 4.5 UNTUK MEMPREDIKSI PENDAPATAN TAHUNAN

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber Informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 26 Agustus 2024



Ryan Salman Gunadharna

NPM : 207006516065

HALAMAN PENGESAHAN
TUGAS SARJANA
IMPLEMENTASI ALGORITMA DECISION TREE C 4.5 UNTUK
MEMPREDIKSI PENDAPATAN TAHUNAN

Oleh:

Ryan Salman Gunadharna

207006516065



Dosen Pembimbing 1

Sari Ningsih, S.Si., M.M

NIDN. 0302066701

Dosen Pembimbing 2

Gatot Soepriyono, S.Si., M.S.M

NIDN. 0317047803

Penguji 1

Dr. Ucuk Darusalam, S.T., M.T

NIDN. 0327047903

Penguji 2

Ariana Azimah, S.T., M.T.I

NIDN. 0307027703

LEMBAR PERSETUJUAN TUGAS AKHIR

LEMBAR PERSETUJUAN TUGAS AKHIR

Tugas Akhir dengan judul :

IMPLEMENTASI ALGORITMA DECISION TREE C 4.5 UNTUK MEMPREDIKSI PENDAPATAN TAHUNAN

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Akhir Semester Genap 2023-2024 pada tanggal 22 bulan Agustus 2024



Dosen Pembimbing 1

Sari Ningsih, S.Si., M.M

NIDN. 0302066701

Ketua Program Studi



Dr. Audrianningsih, S. Kom., MMSI

NIDN. 0303097902

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

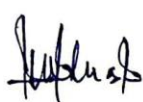
Nama : Ryan Salman Gunadharna
NPM : 207006516065
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
Program Studi : Sistem Informasi
Tanggal Sidang : 22 Agustus 2024

JUDUL DALAM BAHASA INDONESIA :

IMPLEMENTASI ALGORITMA DECISION TREE C 4.5 UNTUK
MEMPREDIKSI PENDAPATAN TAHUNAN

JUDUL DALAM BAHASA INGGRIS :

IMPLEMENTATION OF THE DECISION TREE C 4.5 ALGORITHM TO
PREDICT ANNUAL INCOME

TANDA TANGAN DAN TANGGAL		
Pembimbing 1	Ka. Prodi	Mahasiswa
TGL :	TGL : 27 08 2024	TGL :
 Sari Lingsih		 Ryan Salman Gunadharna

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Ryan Salman Gunadharna
NPM : 207006516065
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
Program Studi : Sistem Informasi
Tanggal Sidang : 22 Agustus 2024

JUDUL DALAM BAHASA INDONESIA :

IMPLEMENTASI ALGORITMA DECISION TREE C 4.5 UNTUK
MEMPREDIKSI PENDAPATAN TAHUNAN

JUDUL DALAM BAHASA INGGRIS :

IMPLEMENTATION OF THE DECISION TREE C 4.5 ALGORITHM TO
PREDICT ANNUAL INCOME

TANDA TANGAN DAN TANGGAL

Pembimbing 2	Kas Prodi	Mahasiswa
TGL :	TGL : 27-08-2024	TGL :
		 Ryan Salman Gunadharna

KATA PENGANTAR

Penulis berterima kasih kepada Allah SWT atas rahmat dan karunia-Nya yang memungkinkan penulis untuk menyelesaikan dan menyusun laporan Tugas Akhir dengan judul “IMPLEMENTASI ALGORITMA DECISION TREE C 4.5 UNUTK MEMPREDIKSI PENDAPATAN TAHUNAN”.

Penulis melakukan penelitian ini, untuk menyelesaikan tugas akhir di Universitas Nasional dalam memperoleh gelar Sarjana Sistem Informasi.

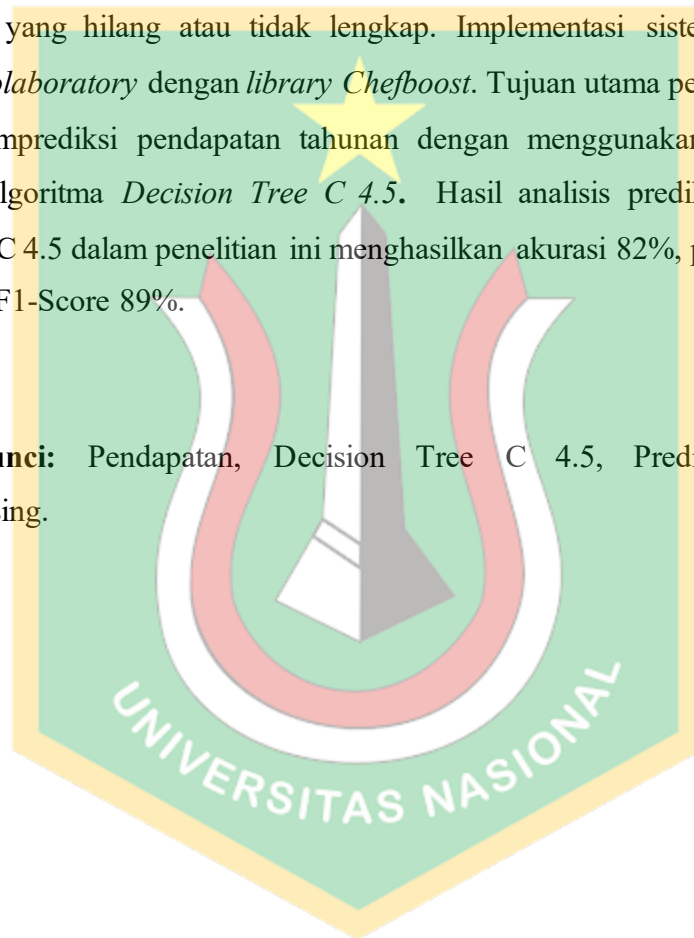
Terlepas dari bantuan yang diberikan secara langsung maupun tidak, kegiatan Tugas Akhir ini berjalan lancar. Oleh karena itu, penulis mengucapkan terima kasih kepada semua orang yang telah membantu:

1. Kedua Orang Tua saya yang selalu memberikan dukungan secara materil serta doa yang senantiasa mengantarkan penulis hingga menyelesaikan tugas akhir ini.
2. Ibu Sari Ningsih, S.Si., M.M. selaku Dosen Pembimbing 1 yang telah memeberikan bimbingan, motivasi, petunjuk, dan arahan kepada penulis dalam kegiatan penyusunan laporan tugas akhir ini.
3. Ibu Andrianingsih, S.Kom, MMSI selaku Dosen Pengampu Skripsi.
4. Bapak Gatot Soepriyono, S.Si., M.S.M. selaku Dosen Pembimbing 2 yang lebih memberikan bimbingan, motivasi, petunjuk, dan arahan kepada penulis dalam kegiatan penyusunan laporan tugas akhir ini.

ABSTRAK

Pendapatan merupakan total penghasilan dari suatu usaha atau aktivitas, baik dalam bentuk uang maupun baran. Penelitian ini memanfaatkan data sensus pendapatan di Amerika Serikat, yang dipilih karena mencakup berbagai aspek demografis, ekonomi, sosial, dan geografis yang komprehensif dan mudah diakses. Data mentah yang digunakan dalam penelitian ini melalui tahap preprocessing untuk menangani informasi yang hilang atau tidak lengkap. Implementasi sistem menggunakan *Google Colaboratory* dengan *library Chefboost*. Tujuan utama penelitian ini adalah untuk memprediksi pendapatan tahunan dengan menggunakan sistem prediksi berbasis algoritma *Decision Tree C 4.5*. Hasil analisis prediksi menggunakan algoritma C 4.5 dalam penelitian ini menghasilkan akurasi 82%, presisi 84%, recall 94%, dan F1-Score 89%.

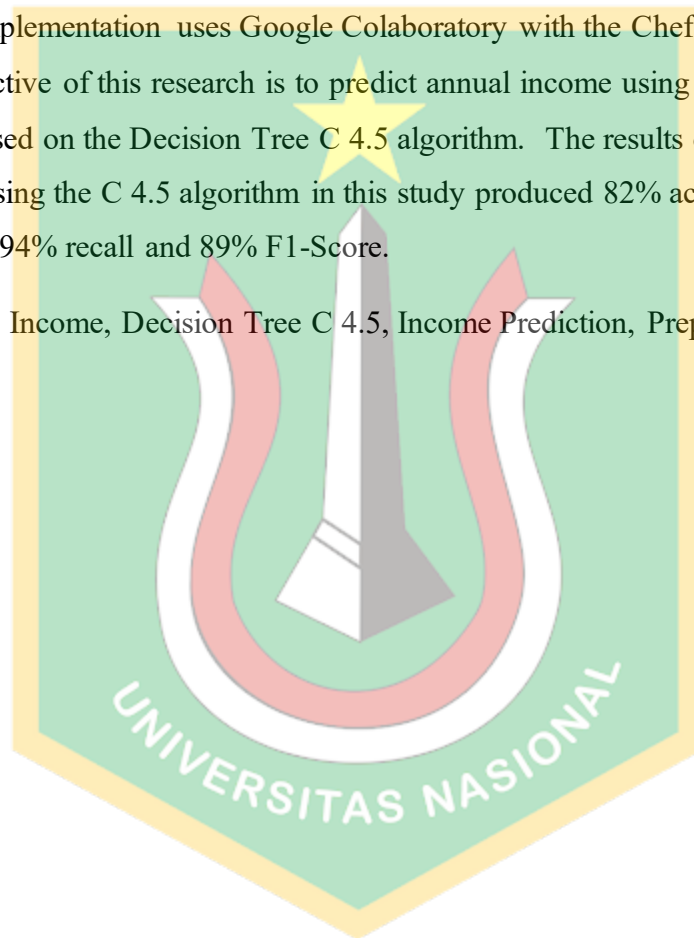
Kata Kunci: Pendapatan, Decision Tree C 4.5, Prediksi Pendapatan, Preprocessing.



ABSTRAK

Income is the total income from a business or activity, both in the form of money and goods. This research utilizes US census data, which was chosen because it covers various demographic, economic, social and geographic aspects in a comprehensive and easily accessible manner. The raw data used in this research went through a preprocessing stage to handle missing or incomplete information. System implementation uses Google Colaboratory with the Chefboost library. The main objective of this research is to predict annual income using a prediction system based on the Decision Tree C 4.5 algorithm. The results of prediction analysis using the C 4.5 algorithm in this study produced 82% accuracy, 84% precision, 94% recall and 89% F1-Score.

Keyword: Income, Decision Tree C 4.5, Income Prediction, Preprocessing



DAFTAR ISI

ABSTRAK.....	x
ABSTRAK.....	xi
DAFTAR ISI.....	xii
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xiv
BAB I PENDAHULUAN.....	Error! Bookmark not defined.
1.1 Latar Belakang.....	Error! Bookmark not defined.
1.2 Rumusan Masalah.....	Error! Bookmark not defined.
1.3 Tujuan Penelitian.....	Error! Bookmark not defined.
1.4 Manfaat Penelitian.....	Error! Bookmark not defined.
1.5 Batasan Masalah.....	Error! Bookmark not defined.
1.6 Sistematika Penulisan.....	Error! Bookmark not defined.
BAB II KAJIAN LITERATUR.....	Error! Bookmark not defined.
2.1 Penelitian Terdahulu.....	Error! Bookmark not defined.
2.2 Landasan Teori.....	Error! Bookmark not defined.
2.2.1 Machine Learning.....	Error! Bookmark not defined.
2.2.2 Data Mining.....	Error! Bookmark not defined.
2.2.3 Decicion Tree.....	Error! Bookmark not defined.
BAB III METODOLOGI PENELITIAN.....	Error! Bookmark not defined.
3.1 Waktu penelitian.....	Error! Bookmark not defined.
3.2 Fokus Penelitian.....	Error! Bookmark not defined.
3.3 Sumber Data.....	Error! Bookmark not defined.

3.4 Tahapan Penelitian	Error! Bookmark not defined.
3.4.1 Identifikasi masalah	Error! Bookmark not defined.
3.4.2 Studi Literatur.....	Error! Bookmark not defined.
3.4.3 Pengambilan Sumber Data	Error! Bookmark not defined.
3.4.4 Perancangan program	Error! Bookmark not defined.
3.4.5 Metode Pengujian.....	Error! Bookmark not defined.
3.4.6 Activity Diagram	Error! Bookmark not defined.
BAB IV PENBAHASAN DAN PENGUJIAN.....	Error! Bookmark not defined.
4.1 Analisis Kebutuhan	Error! Bookmark not defined.
4.1.1 Kebutuhan Fungsional.....	Error! Bookmark not defined.
4.2 Implementasi Algoritma Decision Tree C 4.5	Error! Bookmark not defined.
4.2.1 Perhitungan Manual Algoritma Decision Tree C 4.5	Error! Bookmark not defined.
4.2.2 Struktur Decision Tree	Error! Bookmark not defined.
4.2.3 Tahap Pemodelan Algoritma Decision Tree C 4.5..	Error! Bookmark not defined.
4.2.5 Implementasi Sistem Prediksi	Error! Bookmark not defined.
BAB V KESIMPULAN.....	Error! Bookmark not defined.
5.1 Kesimpulan	Error! Bookmark not defined.
5.2 Saran	Error! Bookmark not defined.
DAFTAR PUSTAKA	Error! Bookmark not defined.
Lampiran	Error! Bookmark not defined.

DAFTAR TABEL

Tabel 3.1 Tabel Waktu Penelitian	Error! Bookmark not defined.
Tabel 3.2 Variabel Independen	Error! Bookmark not defined.
Tabel 3.3 Variabel Dependen.....	Error! Bookmark not defined.
Tabel 4.1 Root Node	Error! Bookmark not defined.

DAFTAR GAMBAR

Gambar 3.1 Flowchart Tahapan Penelitian	Error! Bookmark not defined.
Gambar 3.2 Dataset Adult Cencus Income	Error! Bookmark not defined.
Gambar 3.3 Flowchart Program	Error! Bookmark not defined.
Gambar 3.4 Pemanggilan Dataset	Error! Bookmark not defined.
Gambar 3.5 Activity Diagram Prediksi Pendapatan Tahunan	Error! Bookmark not defined.
Gambar 4.1 Sample Data	Error! Bookmark not defined.
Gambar 4.2 Root Node.....	Error! Bookmark not defined.
Gambar 4.3 Struktur Decision Tree	Error! Bookmark not defined.
Gambar 4.4 Menghapus Duplicate	Error! Bookmark not defined.
Gambar 4.5 Cek Tanda “?”.....	Error! Bookmark not defined.
Gambar 4.6 Mengganti “?” Dengan Value Terbanyak	Error! Bookmark not defined.
Gambar 4.7 Dataset Setelah Transformasi Data	Error! Bookmark not defined.
Gambar 4.8 Library Chefboost.....	Error! Bookmark not defined.
Gambar 4.9 Tingkat Akurasi Preproccsing data.....	Error! Bookmark not defined.
Gambar 4.10 Tingkat Akurasi Transformasi Data .	Error! Bookmark not defined.
Gambar 4.11 Plot Bar Prediksi Decision Tree C 4.5	Error! Bookmark not defined.
Gambar 4.12 Trend Akurasi.....	Error! Bookmark not defined.
Gambar 4.13 Impementasi Sistem Prediksi	Error! Bookmark not defined.