

TUGAS AKHIR

**IMPLEMENTASI DATA MINING DALAM
MENENTUKAN PENERIMAAN PROGRAM
KELUARGA HARAPAN (PKH) MENGGUNAKAN
ALGORITMA C4.5
(STUDI KASUS: DESA KARANGRAU, SOKARAJA)**



DISTY DIKRIANI

19103047

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2023**

TUGAS AKHIR

**IMPLEMENTASI DATA MINING DALAM
MENENTUKAN PENERIMAAN PROGRAM
KELUARGA HARAPAN (PKH) MENGGUNAKAN
ALGORITMA C4.5
(STUDI KASUS: DESA KARANGRAU, SOKARAJA)**



DISTY DIKRIANI

19103047

Khairun Nisa Meiah,S.Pd.,M.Kom

(0605049301)

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2023**

HALAMAN PENGESAHAN

**IMPLEMENTASI DATA MINING DALAM
MENENTUKAN PENERIMAAN PROGRAM
KELUARGA HARAPAN (PKH) MENGGUNAKAN
ALGORITMA C4.5
(STUDI KASUS: DESA KARANGRAU, SOKARAJA)**

Dipersiapkan dan Disusun Oleh

Disty Dikriani

19103047

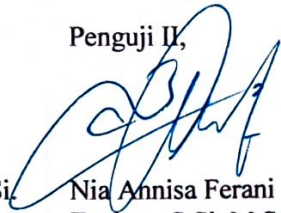
Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas Akhir
pada tanggal 16 Februari 2023

Penguji I.



Atika Ratna Dewi, S.Si.,
M.Sc.
NIDN 0615128703

Penguji II,



Nia Annisa Ferani
Tanjung, S.Si., M.Sc.
NIDN 0630049203

Penguji III,



Sena Wijayanto, S.Pd.,
M.T.
NIDN 0613109201

Pembimbing Utama,



Khairun Nisa Mejah, S.Pd.,
M.Kom
NIDN 0605049301

Dekan



Burhanuddin, S.Si., M.Kom
NIK 19820008

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama mahasiswa : Disty Dikriani

NIM : 19103047

Program Studi : S1 Sistem Informasi

Menyatakan bahwa Tugas Akhir dengan judul berikut :

IMPLEMENTASI DATA MINING DALAM MENENTUKAN PENERIMAAN PROGRAM KELUARGA HARAPAN (PKH) MENGGUNAKAN ALGORITMA C4.5

Dosen Pembimbing Utama: Khairun Nisa Meiah, S.Pd., M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Institut Teknologi Telkom Purwokerto maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan, dan penelitian Saya Sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing.
3. Dalam Karya ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab Saya, bukan tanggung jawab Institut Teknologi Telkom Purwokerto.
5. Pernyataan ini Saya buat dengan sesungguhnya, apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima Sanksi Akademik dengan pencabutan gelar yang sudah diperoleh serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Purwokerto, 05 Februari 2023

Yang Menyatakan,


(Disty Dikriani)

KATA PENGANTAR

Puji dan syukur dipanjatkan kehadirat Allah SWT, atas segala berkat dan ridhanya sehingga dapat menyelesaikan skripsi yang berjudul “Implementasi Data Mining Dalam Menentukan Penerimaan Program Keluarga Harapan (PKH) Menggunakan Algoritma C4.5”. Penulis menyadari akan banyaknya pihak yang senantiasa mendukung dan memberi bantuan selama menyelesaikan studi dan tugas akhir ini. Oleh karena itu, sudah sepantasnya penulis dengan penuh hormat mengucapkan terimakasih dan berdoa semoga Allah SWT memberikan balasan yang paling terbaik kepada:

1. Kedua orangtua beserta keluarga yang telah memberikan dukungan baik melalui materi maupun doa.
2. Dr. Arfianto Fahmi, ST., MT., IPM selaku Rektor Institut Teknologi Telkom Purwokerto.
3. Auliya Burhanudin, S.SI., M.Kom selaku Dekan Fakultas Komputasi (FIF) Institut Teknologi Telkom Purwokerto.
4. Dwi Mustika K, S.Kom., M.Kom selaku Ketua Program Studi S1 Sistem Informasi Institut Teknologi Telkom Purwokerto.
5. Khairun Nisa Meiah, S.Pd., M.Kom selaku dosen pembimbing yang telah memberikan bimbingan, arahan, nasehat serta motivasi dalam penulisan skripsi.
6. Seluruh dosen Program Studi Sistem Informasi Institut Teknologi Telkom Purwokerto.
7. Teman-teman mahasiswa Institut Teknologi Telkom Purwokerto Angkatan 2019 yang saling memberikan semangat, perhatian, dan dukungan.

Dalam penyusunan tugas akhir ini, penulis menyadari bahwa masih terdapat kekurangan. Penulis mengharapkan banyak saran dan kritik yang membangun untuk menyempurnakan tugas akhir ini agar dapat bermanfaat untuk para pembaca.

Purwokerto, 05 Februari 2023



Disty Dikriani

DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	iii
KATA PENGANTAR.....	iv
DAFTAR ISI.....	vi
DAFTAR GAMBAR.....	viii
DAFTAR TABEL	ix
DAFTAR SINGKATAN.....	x
DAFTAR ISTILAH	xi
DAFTAR LAMPIRAN	xii
ABSTRAK	xiii
<i>ABSTRACT</i>	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	5
1.3 Pertanyaan Penelitian.....	5
1.4 Batasan Masalah	5
1.5 Tujuan Penelitian.....	6
1.6 Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Penelitian Sebelumnya.....	8
2.2 Dasar Teori.....	14
2.2.1 Knowledge Discovery in Database (KDD).....	14
2.2.2 Data Mining	15
2.2.3 Dataset	17
2.2.4 Metode Decision Tree	18
2.2.5 Klasifikasi	19
2.2.6 Algoritma C4.5	20
2.2.7 Program Keluarga Harapan (PKH).....	22
2.2.8 RapidMiner.....	24
2.2.9 Confusion Matrix.....	24
2.2.10 Nilai Akurasi	25

BAB III METODOLOGI PENELITIAN	26
3.1 Subjek dan Obyek Penelitian.....	26
3.2 Alat Dan Bahan Penelitian	26
3.2.1 Alat Penelitian	26
3.2.2 Bahan Penelitian	27
3.3 Diagram Alir Penelitian.....	27
3.3.1 Studi Literatur.....	28
3.3.2 Observasi	28
3.3.3 Identifikasi Masalah	28
3.3.4 Pengumpulan Data.....	29
3.3.5 Cleaning dan Integrasi.....	30
3.3.6 Selection dan Transformation	30
3.3.7 Proses Mining	30
3.3.8 Hasil penelitian dan Kesimpulan.....	30
BAB IV HASIL DAN PEMBAHASAN	31
4.1 Analisis Perhitungan.....	31
4.2 Penggunaan Rapidminer	48
4.3 Perbandingan Nilai Akurasi	54
BAB V KESIMPULAN DAN SARAN.....	59
5.1 Kesimpulan	59
5.2 Saran	59
DAFTAR PUSTAKA.....	60
LAMPIRAN.....	64

DAFTAR GAMBAR

Gambar 2. 1 Proses yang sesuai KDD [18]	16
Gambar 2. 2 Blok diagram model klasifikasi.....	19
Gambar 2. 3 Contoh Confusion Matrix [30].....	25
Gambar 3. 1 diagram alir penelitian.....	27
Gambar 4. 1 Pohon Keputusan.....	47
Gambar 4. 2 Tampilan awal halaman Progress.....	48
Gambar 4. 3 Langkah pertama import data.....	48
Gambar 4. 4 Halaman untuk memilih lokasi data.....	49
Gambar 4. 5 Halaman memilih data yang akan diimpor	49
Gambar 4. 6 Peninjauan data	50
Gambar 4. 7 Menyimpan lokasi data import.....	50
Gambar 4. 8 Drop and Drag file yang digunakan	51
Gambar 4. 9 Operator pencarian set role	51
Gambar 4. 10 proses pemilihan label keputusan.....	52
Gambar 4. 11 pemilihan Operator split Validation	52
Gambar 4. 12 drag and drop split validation.....	52
Gambar 4. 13 Drag and drop Decision tree pada training.....	53
Gambar 4. 14 Drag and drop pada testing.....	53
Gambar 4. 15 tampilan luar proses validation	53
Gambar 4. 16 rule algoritma c 4.5 dari 20% data testing.....	54
Gambar 4. 17 Rule pohon keputusan dengan 50% data testing.....	55
Gambar 4. 18 Hasil nilai akurasi dengan 50% data testing.....	55
Gambar 4. 19 Pohon Keputusan dengan 30% data testing	56
Gambar 4. 20 Hasil Akurasi dengan 30% data testing.....	56
Gambar 4. 21 Pohon keputusan dengan 20% data testing	56
Gambar 4. 22 Hasil akurasi dengan 20% data testing.....	57
Gambar 4. 23 Pohon Keputusan dengan 10% data testing	57
Gambar 4. 24 Hasil akurasi dengan 10% data testing.....	57

DAFTAR TABEL

Tabel 2. 1 Penelitian Sebelumnya.....	9
Tabel 3. 1 Contoh Data DTKS	29
Tabel 4. 1 Data yang belum dibersihkan.....	32
Tabel 4. 2 Sampel data transformation penerima PKH.....	33
Tabel 4. 3 Tabel Klasifikasi	35
Tabel 4. 4 Perhitungan Node 1.....	41
Tabel 4. 5 Perhitungan untuk node 2	42
Tabel 4. 6 Perhitungan untuk node 3	43
Tabel 4. 7 Perhitungan untuk node 4-1	44
Tabel 4. 8 Perhitungan untuk node 4-2	44
Tabel 4. 9 Perhitungan untuk node 4-3	45
Tabel 4. 10 Perhitungan untuk kode 4-4	46
Tabel 4. 11 Perbandingan Nilai Akurasi	58

DAFTAR SINGKATAN

PKH	= Program Keluarga Harapan;
KDD	= <i>Knowledge Discovery In Database</i> ;
ID3	= <i>Inteative Dichotomiser</i> ;
SLT	= Subsidi Langsung Tunai;
CCT	= <i>Conditional Cash Transfers</i> ;
DTKS	= Data Terpadu Kesejahteraan Sosial;
GUI	= <i>Graphic User Interface</i> ;
XML	= <i>Extensible Markup Languange</i> ;
NIK	= Nomor Induk Kependudukan;
E-KTP	= Elektronik Kartu Tanda Penduduk;
KTP	= Kartu Tanda Penduduk;
Lansia	= Lanjut Usia

DAFTAR ISTILAH

S	= Himpunan kasus
A	= Atribut
n	= Jumlah partisi atribut a
$ S_i $	= Jumlah pada partisi ke i
$ S $	= Jumlah kasus dalam S
A	= Atribut
$\text{Gain}(S,A)$	= Info gain pada atribut A
$\text{Split data}(S,A)$	= Pembagian data pada atribut A .
p_i	= proporsi dari S_i terhadap S

DAFTAR LAMPIRAN

Lampiran 1 Surat Permohonan Pengambilan Data	64
Lampiran 2 Sampel Data Desa Karangrau.....	65