

**TUGAS AKHIR
PERBANDINGAN ALGORITMA KLASIFIKASI
NAÏVE BAYES, C4.5 DAN KNN UNTUK
MENENTUKAN PEROKOK AKTIF DAN PEROKOK
PASIF**



MUHAMMAD ISRA MUNTAHA TANJUNG

NIM. 17102040

**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2023**

TUGAS AKHIR

PERBANDINGAN ALGORITMA KLASIFIKASI

NAÏVE BAYES, C4.5 DAN KNN UNTUK

MENENTUKAN PEROKOK AKTIF DAN PEROKOK

PASIF

COMPARISON OF NAÏVE BAYES, C4.5 AND KNN

CLASSIFICATION ALGORITHMS TO DETERMINE

ACTIVE AND PASSIVE SMOKERS

Disusun Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Komputer



MUHAMMAD ISRA MUNTAHA TANJUNG

NIM. 17102040

PROGRAM STUDI S1 TEKNIK INFORMATIKA

FAKULTAS INFORMATIKA

INSTITUT TEKNOLOGI TELKOM PURWOKERTO

2023

LEMBAR PERSETUJUAN PEMBIMBING

**PERBANDINGAN ALGORITMA KLASIFIKASI
NAÏVE BAYES, C4.5 DAN KNN UNTUK
MENENTUKAN PEROKOK AKTIF DAN PEROKOK
PASIF**

***COMPARISON OF NAÏVE BAYES, C4.5 AND KNN
CLASSIFICATION ALGORITHMS TO DETERMINE
ACTIVE AND PASSIVE SMOKERS***

Dipersiapkan dan Disusun oleh:

MUHAMMAD ISRA MUNTAHA TANJUNG
NIM. 17102040

Fakultas Informatika
Institut Teknologi Telkom Purwokerto
Pada Tanggal: 6 Februari 2023

Pembimbing Utama,



Amalia Beladinna Arifa, S.Pd., M.Cs.

NIDN 0606019201

LEMBAR PENETAPAN PENGUJI
PERBANDINGAN ALGORITMA KLASIFIKASI
NAÏVE BAYES, C4.5 DAN KNN UNTUK
MENENTUKAN PEROKOK AKTIF DAN PEROKOK
PASIF

COMPARISON OF NAÏVE BAYES, C4.5 AND KNN
CLASSIFICATION ALGORITHMS TO DETERMINE
ACTIVE AND PASSIVE SMOKERS

Dipersiapkan dan Disusun oleh:

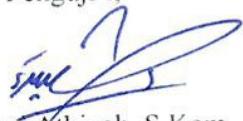
MUHAMMAD ISRA MUNTAHA TANJUNG

NIM. 17102040

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas

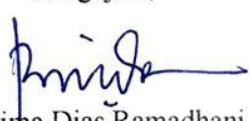
Akhir Pada Hari Kamis, Tanggal 16 Februari 2023

Penguji I,



Ummi Athiyah, S.Kom.,
M.Kom.
NIDN. 0621129001

Penguji II,



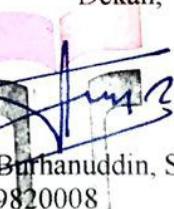
Rima Dias Ramadhani,
S.Kom., M.Kom.
NIDN. 0602039301
Pembimbing Utama,

Penguji III,



Aditya Dwi Putro W,
S.Kom., M.Kom.
NIDN. 0624119303

Dekan,



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

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Isra Muntaha Tanjung
NIM : 17102040
Program Studi : S1 Informatika

Menyatakan bahwa Tugas Akhir dengan judul berikut:

**PERBANDINGAN ALGORITMA KLASIFIKASI NAÏVE BAYES, C4.5
DAN KNN UNTUK MENENTUKAN PEROKOK AKTIF DAN PEROKOK
PASIF**

Dosen Pembimbing Utama : Amalia Beladinna Arifa, S.Pd., M.Cs.

1. Seluruh tulisan dalam Tugas Akhir 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. Seluruh tulisan dalam Tugas Akhir ini merupakan gagasan, rumusan, dan penelitian Penulis Sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing dan menjadi tanggung jawab Penulis sepenuhnya.
3. Seluruh tulisan dalam Tugas Akhir 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.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggungjawab Penulis, bukan tanggungjawab 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, 6 Februari 2023

Yang Menyatakan,



(Muhammad Isra Muntaha Tanjung)

KATA PENGANTAR

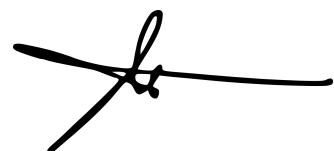
Puji Syukur penulis panjatkan kepada Allah Swt yang telah memberikan kelancaran sehingga penulis mampu menyelesaikan Tugas Akhir dengan judul “Perbandingan Algoritma Klasifikasi Naïve Bayes, C4.5 Dan K-NN Untuk Menentukan Perokok Aktif Dan Perokok Pasif”.

Dalam proses penyusunan tugas akhir ini, tentunya banyak pihak yang telah memberikan bantuan baik moril maupun materil. Oleh karena itu penulis tak lupa mengucapkan terima kasih kepada :

1. Bapak Dr. Arfianto Fahmi, S.T., M.T. IPM selaku Rektor dari Institut Teknologi Telkom Purwokerto.
2. Bapak Auliya Burhanuddin, S.Si., M.Kom selaku Dekan Fakultas Informatika Institut Teknologi Telkom Purwokerto.
3. Ibu Amalia Beladinna Arifa, S.Pd., M.Cs., selaku Ketua Program Studi S1 Teknik Informatika.
4. Ibu Amalia Beladinna Arifa, S.Pd., M.Cs., selaku Dosen Pembimbing yang telah memberikan bimbingan dan pengarahan dalam penyusunan tugas akhir.
5. Kedua Orang tua yang telah memberikan semangat, dukungan dan do'a yang tak pernah putus untuk penulis serta motivasi untuk menyelesaikan tugas akhir ini.
6. Teman-teman Institut Teknologi Telkom Purwokerto yang memberikan dukungan dan semangat terhadap penelitian ini.

Penulis menyadari bahwa penyusunan tugas akhir ini masih jauh dari sempurna dengan segala kekurangannya. Untuk itu penulis berharap adanya kritik dan saran dari semua pihak demi kesempurnaan tugas akhir ini.

Purwokerto, 6 Februari 2023



Muhammad Isra Muntaha Tanjung

TUGAS AKHIR.....	i
LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENETAPAN PENGUJI.....	iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	viii
DAFTAR GAMBAR	ix
ABSTRAK	x
ABSTRACT	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	4
1.3 Pertanyaan Penelitian.....	4
1.4 Tujuan Penelitian	5
1.5 Batasan Masalah	5
1.6 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI	6
2.1 Tinjauan Pustaka.....	6
2.2 Dasar Teori	12
2.2.1 Data <i>Mining</i>	12
2.2.2 <i>Naïve Bayes</i>	13
2.2.3 C4.5	15
2.2.4 K-NN (<i>K-Nearest Neighbor</i>)	16
2.2.5 Confussion <i>Matrix</i>	17
BAB III METODOLOGI PENELITIAN.....	19
3.1 Subjek dan Objek Penelitian.....	19
3.2 Alat dan Bahan Penelitian.....	19
3.3 Diagram Alir Penelitian	19

3.3.1	Studi Literatur	20
3.3.2	Perumusan Masalah dan Tujuan Penelitian	20
3.3.3	Pengumpulan Data	21
3.3.4	Perancangan komparasi <i>Naïve Bayes</i> , C4.5 dan K-NN	24
3.4	Pengujian Metode	28
3.5	Kesimpulan	29
BAB IV	HASIL DAN PEMBAHASAN	30
4.1	Hasil Aplikasi RapidMiner	30
4.1.1	Hasil Pengujian Algoritma C4.5	30
4.1.2	Hasil Pengujian Algoritma <i>K-Nearest Neighbour</i>	34
4.1.3	Hasil Pengujian Algoritma <i>Naïve Bayes</i>	36
4.2	Pembahasan	37
BAB V	KESIMPULAN DAN SARAN.....	39
5.1	Kesimpulan	39
5.2	Saran	39
DAFTAR PUSTAKA	40

DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	10
Tabel 2.2 Confussion Matrix.....	18
Tabel 3.1 Atribut Dataset	22
Tabel 4.1 Aturan atau Rule Hasil Pohon Keputusan RapidMiner	31
Tabel 4.2 Hasil Pengujian Perokok Aktif	37

DAFTAR GAMBAR

Gambar 2.1 Proses Data Mining [20].	12
Gambar 3.1 Flowchart Metode Penelitian	20
Gambar 3.2 Dataset <i>Body Signal Of Smoking</i>	21
Gambar 3.3 Diagram Alur Perancangan Model.....	24
Gambar 3.4 Proses <i>Missing value</i>	25
Gambar 3.5 Proses <i>Split Data</i>	25
Gambar 3.6 Flowchart Klasifikasi Naïve Bayes	26
Gambar 3.7 Flowchart Klasifikasi C4.5 (Decision Tree)	27
Gambar 3.8 Flowchart Klasifikasi K-Nearest Neighbor.....	28
Gambar 4.1 Desain Proses algoritma C4.5	30
Gambar 4.2 Hasil Pohon Keputusan RapidMiner.....	31
Gambar 4.3 Dataset Algoritma C4.5.....	33
Gambar 4.4 Hasil nilai accuracy algoritma C4.5	33
Gambar 4.5 Desain Proses algoritma <i>K-Nearest Neighbor</i>	34
Gambar 4.6 Dataset Algoritma <i>K-Nearest Neighbor</i>	34
Gambar 4.7 Hasil nilai accuracy algoritma <i>K-Nearest Neighbor</i>	35
Gambar 4.8 Desain Proses algoritma <i>Naïve Bayes</i>	36
Gambar 4.9 Dataset Algoritma <i>Naïve Bayes</i>	36
Gambar 4.10 Hasil nilai accuracy algoritma <i>Naïve Bayes</i>	37