



**Performa Algoritma *C4.5* dan *Naïve Bayes* untuk Klasifikasi
Penerima Bantuan Langsung Tunai Dana Desa**

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana

Oleh :

Fariska Desi Rakhmawati

B2A221026

**PROGRAM STUDI S1 STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS MUHAMMADIYAH SEMARANG
2023**


PERSETUJUAN PEMBIMBING

Skripsi dengan judul “Performa Algoritma *C4.5* dan *Naive Bayes* untuk Klasifikasi Penerima Bantuan Langsung Tunai Dana Desa” yang disusun oleh :

Nama : Fariska Desi Rakhmawati
NIM : B2A221026
Program Studi : S1-Statistika

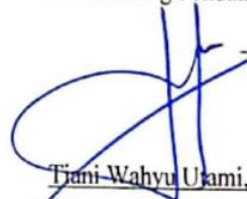
Telah disetujui oleh dosen pembimbing pada tanggal 26 Mei 2023.

Pembimbing Utama



Dr. Rochdi Warsono, M.Si
NIK.28.6.1026.119

Pembimbing Pendamping



Tiani Wahyu Utami, S.Si, M.Si
NIK. 28.6.1026.341

Mengetahui,

Ketua Program Studi



Indah Manfaati Nur, M.Si

NIK. 28.6.1026.221

PENGESAHAN KELULUSAN

Skripsi dengan Judul “ Performa Algoritma *C4.5* dan *Naive Bayes* untuk Klasifikasi
Penerima Bantuan Langsung Tunai Dana Desa” yang disusun oleh:

Nama : Fariska Desi Rakhmawati

NIM : B2A221026

Program Studi : S1 Statistika

Telah dipertahankan dalam Sidang Panitia Ujian Skripsi Program Sarjana Statistika
Universitas Muhammadiyah Semarang pada tanggal: 26 Mei 2023

Panitia Ujian
Ketua Tim Penguji



Indah Manfaati Nur, M.Si

NIK. 28.6.1026.221

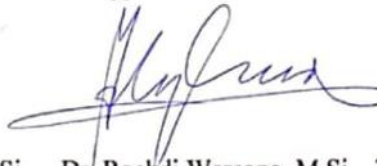
Anggota Tim Penguji I



M. Al Haris, S.Si, M.Si

NIK.28.6.1026.442

Anggota Tim Penguji II



Dr. Rochdi Warsono, M.Si

NIK.28.6.1026.119

Anggota Tim Penguji III

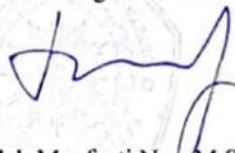


Tianj Wahyu Utami, S.Si, M.Si

NIK. 28.6.1026.341

[
M.Si

Mengetahui,
Ketua Program Studi



Indah Manfaati Nur, M.Si

NIK. 28.6.1026.221

MOTTO

“Hai orang-orang beriman, Jadikanlah sabar dan sholatmu
Sebagai penolongmu, sesungguhnya
Allah SWT beserta orang-orang yang sabar”
(Al-Baqarah :153)

“Allah mengangkat orang-orang beriman di antara kalian dan orang-orang yang
diberi ilmu beberapa derajat”
(Q.S Al-Mujadalah: 11)

“Keberhasilan adalah sebuah proses. Niatmu adalah awal keberhasilan. Peluh
keringatmu adalah pendedapnya. Tetesan air matamu adalah pewarnanya. Doamu
dan doa orang-orang disekitarmu adalah bara api yang mematangkannya.
Kegagalan di setiap langkahmu adalah pengawetnya. Akan dari itu, bersabarlah,
Karena Allah SWT selalu menyertai orang-orang yang penuh
kesabaran dalam proses menuju keberhasilan”
(Muhammad Kamil)

PERNYATAAN KEASLIAN

Dengan ini saya menyatakan bahwa:

1. Karya tulis saya, skripsi ini, adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana, Magister, dan/atau Doktor), baik di *Universitar Muhammadiyah Semarang* maupun di perguruan tinggi lain.
2. Karya tulis ini adalah murni gagasan, rumusan, dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali arahan Tim Pembimbing dan masukan Tim Penguji.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademika berupa pencabutan gelar yang telah diperoleh karena karya ini, serta sanksi lainnya sesuai dengan norma yang berlaku di perguruan tinggi ini.

Semarang, 25 Mei 2023

Yang membuat pernyataan,



Fariska Desi Rakhmawati
NIM. B2A221026

PERSEMBAHAN SKRIPSI

Alhamdulillahirobil'alamin

Segala puji bagi mu Tuhan ku Allah SWT yang Maha pengasih dan Maha penyayang atas segala karunia yang telah Kau berikan, atas kelancaran, pertolongan, dan kekuatan beserta rahmat-Mu sehingga saya dapat menyelesaikan skripsi saya sesuai dengan rencana, karena sesungguhnya saya sebagai manusia hanya bisa merencanakan tetapi pada akhirnya Engkaulah yang berhak memutuskan. Karya sederhana ini sebagai bentuk tanggung jawab pengabdian, ungkapan terima kasih, hormat, dan kasih sayang saya kepada :

1. Bapak dan Ibu, selaku orang tua yang telah memberikan kasih sayang, dukungan serta doa yang tiada henti untuk anaknya.
2. Suamiku pendukung nomor satu, yang selalu mendengarkan segala keluhan kesahku, I love you.
3. Anak ku tercinta Jesna Aisy Aurensia, yang selalu menemani hari hariku, penyemangat hidupku.
4. Mbak, Mas, Adik, Serta Ponakan-Ponakan tercinta
5. Dan untuk siapapun, dimanapun, yang telah menyempatkan dan meluangkan waktu untuk membaca karya ilmiahku.

KATA PENGANTAR

Assalamu'alaikum Wr.Wb

Segala puji dan syukur penulis haturkan kehadiran Allah Subhanahu Wata'ala yang telah memberikan rahmat, taufiq, serta hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “Performa Algoritma C4.5 dan Naive Bayes Untuk Klasifikasi Penerima Bantuan Langsung Tunai Dana Desa.”. Skripsi ini disusun sebagai salah satu syarat meraih gelar Sarjana Statistika pada Program Studi Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang. Shalawat serta Salam semoga senantiasa tercurahkan kepada junjungan kita Nabi Muhammad Shallallahu 'Alaihi Wasallam, semoga kita semua mendapatkan syafa'at-Nya di yaumul akhir nanti, Amin.

Penulis menyadari sepenuhnya, bahwa dalam penyelesaian skripsi ini tidak lepas dari bantuan dan bimbingan berbagai pihak. Pada kesempatan ini penulis ingin menyampaikan terima kasih dan penghargaan yang setinggi - tingginya kepada pihak-pihak yang telah memberikan bantuan baik secara langsung maupun secara tidak langsung, sehingga penulis dapat menyelesaikan skripsi ini. Penulis menyampaikan terima kasih dan penghargaan setinggi-tingginya kepada:

1. Bapak Prof. Dr. Masrukhi, M.Pd, selaku Rektor Universitas Muhammadiyah Semarang.
2. Ibu Dr. Eny Winaryati, M.Pd, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang.
3. Ibu Indah Manfaati Nur, S.Si, M.Si, selaku Ketua Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang.
4. Dr. Rochdi Warsono, M.Si selaku Dosen Pembimbing I yang dengan sabar membimbing dari awal sampai akhir pembuatan skripsi ini.
5. Tiani Wahyu Utami, S.Si, M.Si, selaku Dosen Pembimbing II yang selalu siap sedia untuk membimbing dan meluangkan banyak waktunya.

6. Bapak dan Ibu Dosen Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang, yang telah memberikan bekal ilmu yang bermanfaat kepada penulis sehingga akhirnya penulis mampu menyelesaikan skripsi.
7. Kepada Rekan – Rekan Mahasiswa Program Studi Sarjana Statistika yang sudah memberikan motivasi kepada penulis dalam menyelesaikan skripsi ini.
8. Kepada semua pihak yang tidak dapat penulis sebutkan satu per satu, atas do'a dan motivasi dalam penyelesaian skripsi ini.

Penulis menyadari akan segala keterbatasan dan kekurangan dalam penulisan skripsi ini. Oleh karena itu, kritik dan saran yang bersifat membangun dari semua pihak akan penulis terima dengan senang hati. Semoga hasil penelitian ini dapat memberikan manfaat dan kontribusi bagi pengembangan pembelajaran statistika di masa yang akan datang.

Wassalamu'alaikum Wr.Wb.

Semarang, 25 Mei 2023


Fariska Desi Rakhmawati

NIM. B2A221026

DAFTAR ISI

PERSETUJUAN PEMBIMBING.....	ii
PENGESAHAN KELULUSAN	iii
PERNYATAAN KEASLIAN.....	v
MOTTO	iv
PERSEMBAHAN SKRIPSI.....	ii
ABSTRAK.....	vii
<i>ABSTRACT</i>	viii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	6
1.5 Batasan Penelitian	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Data Mining.....	8
2.2 <i>Preprocessing</i> Data.....	9
2.3 Konsep dan Model Klasifikasi	11
2.4 Klasifikasi <i>Decision Tree</i>	12
2.5 Algoritma <i>C4.5</i>	13
2.6 <i>Naïve bayes</i>	14
2.7 <i>Confusion matrix</i>	16
2.8 Kurva <i>ROC</i>	18
2.9 RapidMiner.....	19
2.10 Kemiskinan.....	20

2.11	Bantuan Langsung Tunai.....	23
2.12	Dana Desa.....	25
2.13	Desa Asemdayong.....	27
BAB III METODOLOGI PENELITIAN.....		30
3.1	Sumber data.....	30
3.2	Variabel Penelitian.....	30
3.3	Struktur Data.....	32
3.4	Langkah Penelitian.....	32
3.5	Alur Penelitian.....	34
BAB IV HASIL DAN PEMBAHASAN.....		35
4.1	<i>Preprocessing</i> Data.....	35
4.1.1	<i>Cleaning</i> data.....	35
4.1.2	Transformasi Data.....	35
4.2	Hasil eksperimen dan Pengujian Model.....	38
4.2.1	Hasil Eksperimen dan Pengujian Model Naïve Bayes.....	39
4.2.2	Hasil Eksperimen dan Pengujian Model C4.5.....	43
4.2.3	Perbandingan Model <i>Naïve Bayes</i> dan C4.5.....	48
BAB IV KESIMPULAN DAN SARAN.....		53
5.1	Kesimpulan.....	53
5.2	Saran.....	54
DAFTAR PUSTAKA.....		55
LAMPIRAN.....		59

DAFTAR TABEL

Tabel 2.1 Confusion Matrix	17
Tabel 3. 1 Variabel Independen.....	30
Tabel 3. 2 Struktur Data	32
Tabel 4. 1 Confussion matrix algoritma Naïve Bayes (data testing 30% : data training 70%)	39
Tabel 4. 2 Confussion matrix algoritma Naïve Bayes (data testing 25% : data training 75%)	40
Tabel 4. 3 Confussion matrix Algoritma Naïve Bayes (data testing 20% : data training 80%)	42
Tabel 4. 4 Confussion matrix algoritma C.45 (data testing 30% : data training 70%)	44
Tabel 4. 5 Confussion Matrix Algoritma C.45 (data testing 25% : data training 75%)	45
Tabel 4. 6 Confussion Matrix Algoritma C.45 (data testing 20% : data training 80%)	47
Tabel 4. 7 Hasil komparasi Algoritma C.45 dan Algoritma Naïve Bayes berdasarkan data training dan data testing	48
Tabel 4. 8 Rata-rata hasil Komparasi Algoritma Berdasarkan Data Testing dan Data Training	50

DAFTAR GAMBAR

Gambar 1.1 Jumlah dan Persentase Penduduk Miskin Ekstrem	1
Gambar 2. 1 Alur Keja Model Klasifikasi.....	12
Gambar 2.2 Kurva ROC	19
Gambar 3. 1 Alur Penelitian	34
Gambar 4.1 Area Under Curve (AUC) Algoritma Naïve Bayes (data testing 30% : data training 70%).....	40
Gambar 4.2 Area Under Curve (AUC) Algoritma Naïve Bayes (data testing 25% : data training 75%).....	41
Gambar 4. 3 Area Under Curve (AUC) Algoritma Naïve Bayes (data testing 20% : data training 80%).....	43
Gambar 4. 4 Area Under Curve (AUC) algoritma C4.5 (data testing 30% : data training 70%)	44
Gambar 4. 5 Area Under Curve (AUC) Algoritma C4.5 (data testing 25% : data training 75%)	46
Gambar 4. 6 Area Under Curve (AUC) Algoritma C4.5 (data testing 20% : data training 80%)	47
Gambar 4. 7 Perbandingan Uji Performa algoritma Naïve Bayes dan algoritma C4.5	51