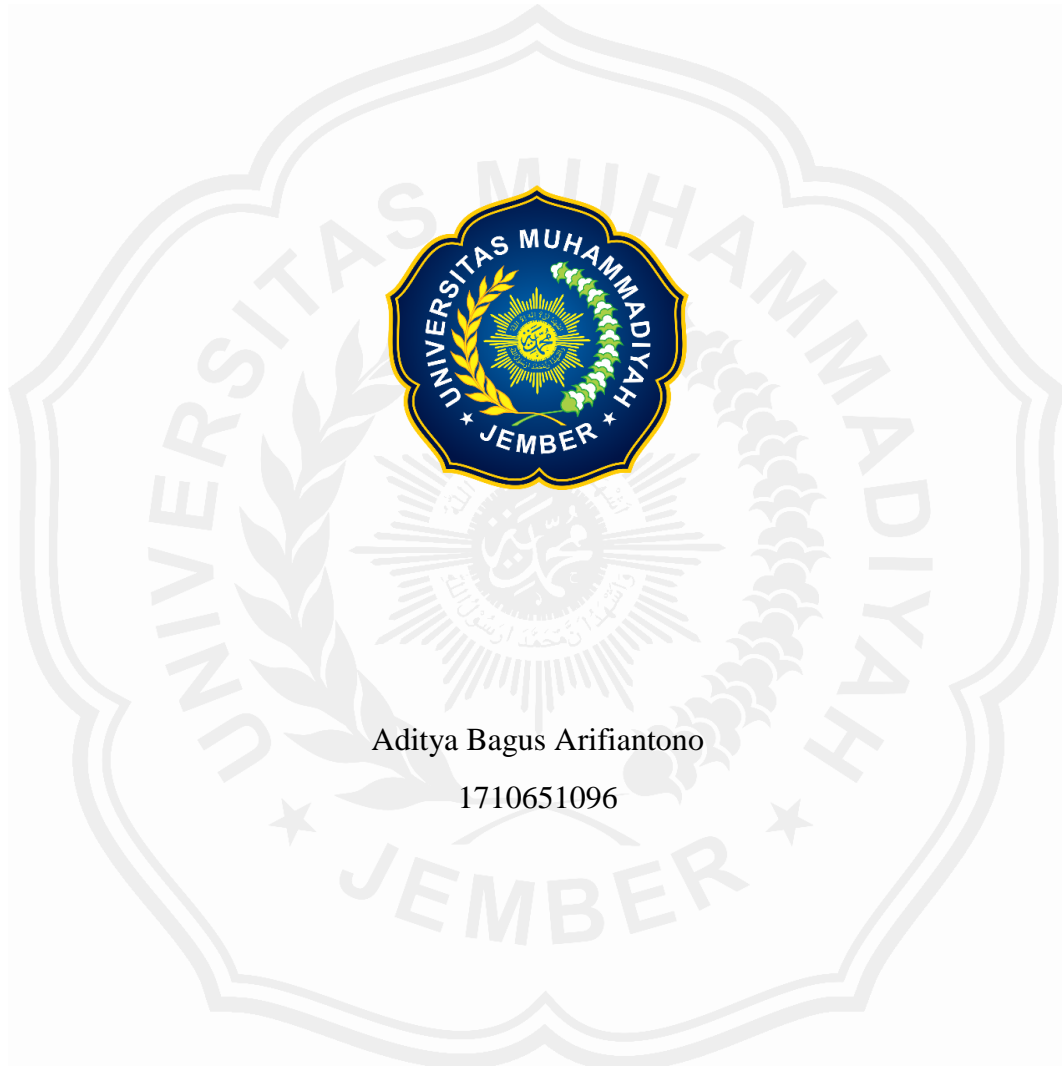


TUGAS AKHIR

PEMODELAN *MACHINE LEARNING* PADA DATA PASIEN DIAGNOSIS DEMAM *TIFOID* MENGGUNAKAN METODE *K-NEAREST NEIGHBOR*



Aditya Bagus Arifiantono

1710651096

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
2024**

TUGAS AKHIR

**PEMODELAN *MACHINE LEARNING* PADA DATA PASIEN
DIAGNOSIS DEMAM *TIFOID* MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR***

Disusun untuk Melengkapi Tugas dan Memenuhi Syarat Kelulusan
Program Strata 1 Jurusan Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember



Aditya Bagus Arifiantono

1710651096

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
2024**

LEMBAR PERSETUJUAN

**PEMODELAN *MACHINE LEARNING* PADA DATA PASIEN DIAGNOSIS
DEMAM *TIFOID* MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR***

Diajukan Oleh:

Aditya Bagus Arifiantono

1710651096

Telah disetujui bahwa Laporan Tugas Akhir ini untuk diajukan pada sidang
Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar

Sarjana Komputer(S.Kom)

di

Universitas Muhammadiyah Jember

Disetujui oleh:

Dosen Pembimbing I

Dosen Pembimbing II



Hardian Oktavianto, S.Si., M.Kom
NIDN. 0722108105



Guruh Wjiayah, S.T., M.Kom
NIDN. 20110059

LEMBAR PENGESAHAN

**PEMODELAN *MACHINE LEARNING* PADA DATA PASIEN DIAGNOSIS
DEMAM *TIFOID* MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR***

Oleh :

Aditya Bagus Arifianto

1710651096

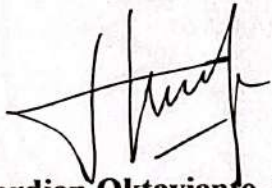
Telah mempertanggung jawabkan Laporan Tugas Akhir nya pada sidang Tugas Akhir tanggal 29 Juni 2024 sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

Disetujui Oleh :

Pembimbing I



Hardian Oktavianto, S.Si., M.Kom

NIDN: 0722108105

Pembimbing II



Guruh Wijayah, S.T., M.Kom

NIDN:20110059

Penguji I



Nur Qodariyah Fitriyah, S.T., M.Kom

NIDN:0727097501

Penguji II



Dudi Irawan, S.T., M.Kom

NIDN:0730037703

Mengesahkan,
Dekan Fakultas Teknik



Dr. Ir. Mukhar, S.T., M.T., IPM

NIDN:0010067301

Mengetahui,
Ketua Program Studi Teknik Informatika



Rosita Yanuarti, S.Kom., M.Cs

NIDN:0629018601

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama : Aditya Bagus Arifiantono

NIM : 1710651096

Institusi : S1 Teknik Informatika, Fakultas Teknik.
Universitas Muhammadiyah Jember

Menyatakan dengan sesungguhnya bahwa Tugas Akhir berjudul "**Pemodelan *Machine Learning* Pada Data Pasien Diagnosis Demam *Tifoid* Menggunakan *K-Nearest Neighbor***" adalah benar merupakan karya sendiri. Hal-hal yang bukan karya saya, dalam skripsi tersebut diberi tanda citasi dan ditunjukkan dalam daftar pustaka.

Apabila di kemudian hari terbukti pernyataan saya tidak benar dan ditemukan pelanggaran atas karya Tugas Akhir ini, saya bersedia menerima sanksi akademik berupa pencabutan Tugas Akhir dan gelar yang saya peroleh dari Tugas Akhir tersebut.

Jember, 29 Juni 2024

Yang menyatakan,



Aditya Bagus Arifiantono
NIM.1710651096

PERSEMBAHAN

Alhamdulillah segala puji bagi Allah SWT yang senantiasa memberikan nikmat yang begitu luar biasa kepada penulis, Serta berkat rahmat-Nya yang begitu melimpah penulis diberikan kesehatan dan kelancaran dalam menyelesaikan studi di kampus Universitas Muhammadiyah Jember

Atas segala upaya, bimbingan dan arahan dari semua pihak, penulis mengucapkan banyak terimakasih. Dalam kesempatan ini penulis mengucapkan terimakasih yang sebesar-besarnya kepada :

1. Allah SWT segala urusan ridho, rahmat dan hidayah-Nya penulis dapat menyelesaikan segala urusan pada saat menyusun laporan Tugas Akhir dan diberikan kesempatan mendapatkan gelar Sarjana Komputer.
2. Kedua orang tua penulis Ibu Wiwin Sriwinarti dan Bapak Soekartono serta saudara laki-laki Tonofan Bayu Winata, terimakasih telah mendukung dan mensupport penulis untuk bisa menyelesaikan penulisan Tugas Akhir ini.
3. Bapak Dr. Ir. Muhtar, S.T.,M.T.,IPM, selaku Dekan Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
4. Ibu Rosita Yanuarti, S.Kom, M.Cs, selaku Ketua Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember.
5. Hardian Oktavianto,S.Si., M.Kom, selaku dosen pembimbing 1 dan Bapak Guruh Wijaya, S.T, M.Kom, selaku dosen pembimbing 2 Tugas Akhir yang telah meluangkan waktu untuk memberi bimbingan dan arahan dalam menyelesaikan Tugas Akhir ini.
6. Ibu Nur Qodariyah Fitriyah, S.T., M.Kom, selaku dosen penguji 1 dan Bapak Dudi Irawan, S.T., M.Kom, selaku dosen penguji 2 yang telah memberikan kritik, saran, serta masukkan.
7. Dan tidak lupa penulis ucapkan banyak terima kasih kepada teman-teman, sahabat, dan kerabat khususnya di Universitas Muhammadiyah Jember yang sudah berjuang bersama hingga saat ini.

KATA PENGANTAR

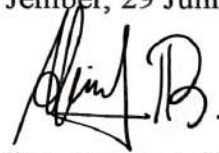
Dengan mengucapkan Alhamdulillah segala puji dan syukur penulis panjatkan atas kehadiran Allah SWT, karena berkat rahmat dan hidayah-Nya penulis dapat menyelesaikan tugas akhir dengan judul "**Pemodelan *Machine Learning* Pada Data Pasien Diagnosis Demam *Tifoid* Menggunakan *K-Nearest Neighbor***".

Penelitian tugas akhir ini mendasarkan pada isu kesehatan, Tugas akhir ini merupakan karya ilmiah yang disusun dalam upaya untuk menyelesaikan pendidikan sarjana (S1) pada Fakultas Teknik Prodi Teknik Informatika Universitas Muhammadiyah Jember.

Penulis menyadari dalam penyusunan Tugas Akhir ini tidak selesai tanpa adanya dukungan dari bapak, ibu pembimbing dan penguji akan sangat sulit untuk menyelesaikan tugas akhir ini. Oleh karena itu pada kesempatan kali ini penulis menyampaikan terima kasih pada bapak, ibu pembimbing serta penguji. Penulis menyadari bahwa tugas akhir ini jauh dari kata sempurna.

Akhir kata, penulis mengharapkan tugas akhir ini dapat memberikan manfaat bagi penulis khususnya dan bagi pembaca pada umumnya.

Jember, 29 Juni 2024



Aditya Bagus Arifiantono

NIM.1710651096

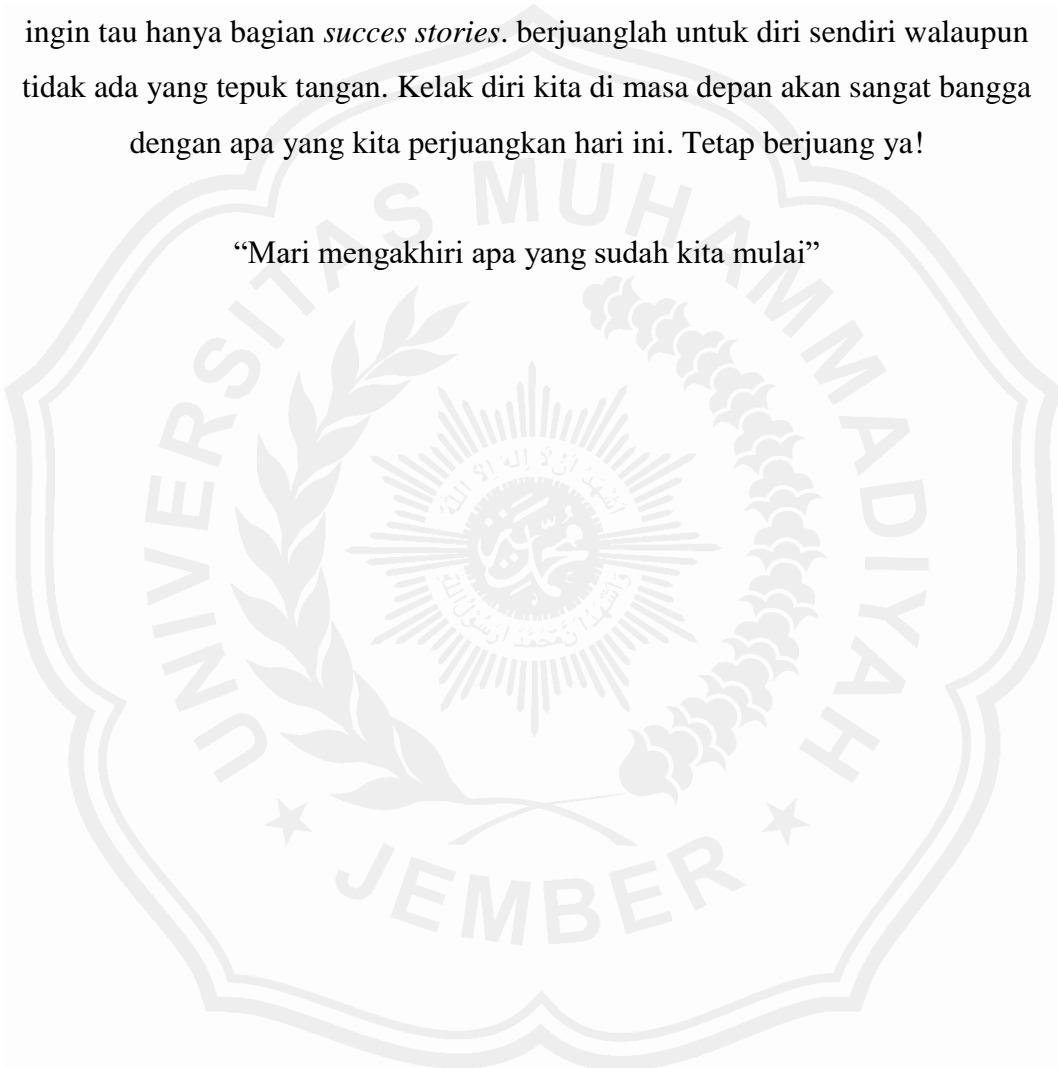
MOTTO

“Sesungguhnya Allah tidak akan merubah nasib suatu kaum hingga mereka mengubah diri mereka sendiri”

(QS.Ar-Ra'd : 11)

“Orang lain tidak akan bisa paham *struggle* dan masa sulitnya kita, yang mereka ingin tau hanya bagian *succes stories*. berjuanglah untuk diri sendiri walaupun tidak ada yang tepuk tangan. Kelak diri kita di masa depan akan sangat bangga dengan apa yang kita perjuangkan hari ini. Tetap berjuang ya!

“Mari mengakhiri apa yang sudah kita mulai”



DAFTAR ISI

COVER	i
LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN	iv
PERSEMBAHAN	v
KATA PENGANTAR	vi
MOTTO	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
ABSTRAK	xiii
<i>ABSTRACT</i>	xiv
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan.....	3
1.4. Manfaat.....	3
1.5. Batasan Masalah.....	3
BAB 2 TINJAUAN PUSTAKA	5
2.1. Penelitian Terdahulu.....	5
2.2. Demam <i>Tifoid</i>	7
2.3. Klasifikasi.....	8
2.4. <i>Label Encoding</i>	8
2.5. <i>TF-IDF (Term Frequency - Inverse Document Frequency)</i>	9
2.6. <i>K-Nearest Neighbor</i>	9
2.7. <i>K-Fold Cross Validation</i>	10
2.8. <i>Confusion Matrix</i>	11
2.9. Pengembangan Penelitian.....	12
BAB 3 METODOLOGI PENELITIAN.....	13
3.1. Tahapan Penelitian	13
3.1.1. Pengumpulan Data	14
3.1.2. <i>Preprocessing data</i>	14

3.1.3.	Partisi Data	15
3.1.4.	Pemodelan	15
3.1.5.	Pengujian dan Pengukuran	16
3.2.	Studi Kasus penggunaan metode	17
3.2.1.	Studi kasus implementasi <i>TF-IDF</i>	17
3.2.2.	Studi kasus implementasi <i>Label encoder</i>	18
3.2.3.	Studi kasus Implementasi <i>K-Nearest Neighbor</i>	19
3.2.4.	Studi kasus implementasi <i>Confusion Matrix</i>	20
BAB 4 HASIL DAN PEMBAHASAN		23
4.1.	Gambaran Data	23
4.1.1.	Sumber Data	23
4.1.2.	Jumlah Data	25
4.1.3.	Jenis Atribut	25
4.1.4.	Jenis Diagnosis	26
4.2.	<i>Preprocessing</i>	26
4.2.1.	<i>Missing value</i>	26
4.2.2.	Transformasi Data	28
4.2.3.	<i>TF-IDF</i>	29
4.3.	Partisi Data	30
4.4.	Pemodelan	31
4.4.1.	Pemodelan Menggunakan <i>K-NN</i> dan <i>K-Fold Cross Validation</i>	31
4.4.2.	Pemodelan Menggunakan <i>K-NN</i> dan <i>K-Fold Cross Validation</i> dengan Implementasi <i>SMOTE</i> pada Data Latih	33
4.4.3.	Pemodelan Menggunakan <i>K-NN</i> dan <i>K-Fold Cross Validation</i> pada Atribut Dataset dengan Tipe Data Teks	35
4.4.4.	Pemodelan Menggunakan <i>K-NN</i> dan <i>K-Fold Cross Validation</i> pada Atribut Data Teks dengan Implementasi <i>SMOTE</i> pada Data Latih	37
4.5.	Pengujian	40
4.5.1.	Pengujian menggunakan model data awal tanpa <i>SMOTE</i>	40
4.5.2.	Pengujian menggunakan model dengan implementasi <i>SMOTE</i> pada data latih	41
4.5.3.	Pengujian menggunakan model tipe data teks	43
4.5.4.	Pengujian menggunakan model tipe data teks dengan implementasi <i>SMOTE</i> pada data latih	44
4.6.	Hasil	46

4.6.1. Hasil Analisis Pengujian Model.....	46
4.6.2. Hasil Analisis Keluhan Menggunakan N-Gram	47
BAB 5 KESIMPULAN DAN SARAN	49
5.1. Kesimpulan.....	49
5.2. Saran.....	50
DAFTAR PUSTAKA	51
LAMPIRAN	



DAFTAR GAMBAR

Gambar 3. 1 Tahapan Penelitian	13
Gambar 3. 2 Tahapan <i>Preprocessing</i>	14
Gambar 3. 3 Proses Partisi data.....	15
Gambar 3. 4 Proses Pemodelan.....	15
Gambar 3. 5 Proses Pengujian dan Pengukuran.....	16
Gambar 4. 1 Kondisi Data Keseluruhan	25
Gambar 4. 2 Kondisi data sebelum partisi	30
Gambar 4. 3 Kondisi data latih (kiri) dan data uji (kanan) setelah dipartisi	31
Gambar 4. 4 Hasil pengujian metode <i>KNN</i> pada data awal tanpa implementasi <i>SMOTE</i>	41
Gambar 4. 5 Hasil pengujian metode <i>KNN</i> pada data awal setelah implementasi <i>SMOTE</i>	42
Gambar 4. 6 Hasil pengujian metode <i>KNN</i> pada fitur teks tanpa implementasi <i>SMOTE</i>	43
Gambar 4. 7 Hasil pengujian metode <i>KNN</i> pada fitur teks setelah implementasi <i>SMOTE</i>	44
Gambar 4. 8 Perbandingan hasil pemodelan tiap skenario	46



DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	5
Tabel 2. 2 Model Confusion Matrix.....	11
Tabel 3. 1 Data Studi kasus <i>TF-IDF</i>	17
Tabel 3. 2 Hasil hitung TF dan DF pada Studi kasus <i>TF-IDF</i>	17
Tabel 3. 3 Hasil hitung IDF pada studi kasus <i>TF-IDF</i>	18
Tabel 3. 4 Proses pada studi kasus <i>Label encoder</i>	18
Tabel 3. 5 Data latih pada studi kasus implementasi <i>K-Nearest Neighbor</i>	19
Tabel 3. 6 Data uji pada studi kasus implementasi <i>K-Nearest Neighbor</i>	19
Tabel 3. 7 Proses Hitung jarak <i>Euclidean Distance</i> pada studi kasus implementasi <i>KNN</i>	20
Tabel 3. 8 Hasil klasifikasi pada data uji studi kasus implementasi <i>KNN</i>	20
Tabel 3. 9 Perbandingan nilai aktual dan hasil klasifikasi pada studi kasus <i>Confusion Matrix</i>	21
Tabel 3. 10 Hasil pengelompokan kategori pada hasil klasifikasi studi kasus <i>Confusion Matrix</i>	21
Tabel 4. 1 Data pasien hasil pemeriksaan Puskesmas.....	23
Tabel 4. 2 Kondisi <i>Missing value</i> pada Data.....	26
Tabel 4. 3 Hasil hitung nilai rata-rata suhu berdasarkan Hasil IMT	27
Tabel 4. 4 Kondisi <i>Missing value</i> setelah <i>preprocessing</i>	28
Tabel 4. 5 Hasil penerapan <i>Label encoder</i> pada transformasi data.....	28
Tabel 4. 6 Hasil pemodelan metode <i>KNN</i> pada data latih tanpa implementasi <i>SMOTE</i>	32
Tabel 4. 7 Hasil pemodelan metode <i>KNN</i> pada data latih setelah implementasi <i>SMOTE</i>	33
Tabel 4. 8 Hasil pemodelan metode <i>KNN</i> pada data latih pada fitur teks.....	36
Tabel 4. 9 Hasil pemodelan metode <i>KNN</i> pada data latih dengan fitur teks dan diimplementasi <i>SMOTE</i>	38
Tabel 4. 10 Hasil hitung frekuensi kata menggunakan N-Gram.....	47