

TUGAS AKHIR

**PERBANDINGAN ALGORITMA *K-NEAREST NEIGHBOR (KNN)* DAN
GAUSSIAN NAIVE BAYES (GNB) DALAM KLASIFIKASI *BREAST
CANCER COIMBRA***

Disusun untuk Melengkapi dan Memenuhi Syarat Kelulusan Program Strata 1

Program Studi Teknik Informatika Fakultas Teknik

Universitas Muhammadiyah Jember



JOHAN TARUNA WIJAYA

1610651068

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2020

TUGAS AKHIR

**PERBANDINGAN ALGORITMA *K-NEAREST NEIGHBOR (KNN)* DAN
GAUSSIAN NAIVE BAYES (GNB) DALAM KLASIFIKASI *BREAST
CANCER COIMBRA***

Disusun untuk Melengkapi dan Memenuhi Syarat Kelulusan Program Strata 1

Program Studi Teknik Informatika Fakultas Teknik

Universitas Muhammadiyah Jember



JOHAN TARUNA WIJAYA

1610651068

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2020

HALAMAN PERSETUJUAN

**PERBANDINGAN ALGORITMA *K-NEAREST NEIGHBOR (KNN)* DAN
GAUSSIAN NAIVE BAYES (GNB) DALAM KLASIFIKASI *BREAST
CANCER COIMBRA***

Oleh:

Johan Taruna Wijaya

1610651068

Telah disetujui bahwa Laporan Tugas Akhir ini untuk diajukan pada Sidang Tugas Akhir sebagai salah satu syarat kelulusan dan mendapat gelar Sarjana Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

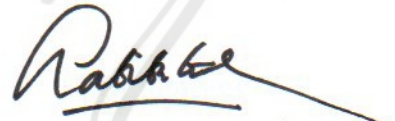
Disetujui oleh:

Pembimbing I



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

Pembimbing II



Habibatul Azizah Al Faruq. M. Pd.
NIDN. 0718128901

HALAMAN PENGESAHAN

**PERBANDINGAN ALGORITMA *K-NEAREST NEIGHBOR (KNN)* DAN
GAUSSIAN NAIVE BAYES (GNB) DALAM KLASIFIKASI *BREAST
CANCER COIMBRA***

Oleh:

Johan Taruna Wijaya

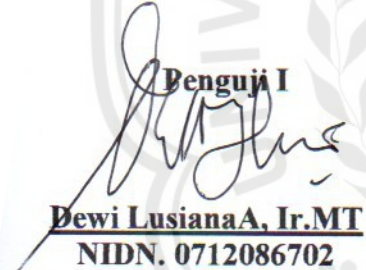
1610651068

Telah mempertanggung jawabkan Laporan Tugas Akhirnya pada Sidang Tugas Akhir tanggal
(30 April 2021) sebagai salah satu syarat kelulusan dan mendapatkan gelar
Sarjana Komputer (S.Kom)

Di
Universitas Muhammadiyah Jember

Disetujui oleh:

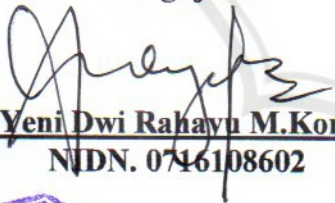
Penguji I


Dewi Lusiana A. Ir. MT
NIDN. 0712086702

Pembimbing I


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

Penguji II


Yeni Dwi Rahayu M.Kom
NIDN. 0746108602

Pembimbing II


Habibatul Azizah Al Faruq. M. Pd.
NIDN. 0718128901

**Mengesahkan,
Dekan Fakultas Teknik**


Dr. Nanang Saiful Rizal. S.T., M.T.
NIDN. 0705047806

**Mengetahui,
Ketua Program Studi Teknik Informatika**


Triawan Adi Cahyanto. M.Kom.
NIDN. 0702098804

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

NAMA : Johan Taruna wijaya
NIM : 1610651068
Program Studi : Teknik Informatika
Perguruan Tinggi : Universitas Muhammadiyah Jember

Menyatakan bahwa Tugas Akhir yang berjudul **“PERBANDINGAN ALGORITMA *K-NEAREST NEIGHBOR (KNN)* DAN *GAUSSIAN NAIVE BAYES (GNB)* DALAM KLASIFIKASI *BREAST CANCER COIMBRA*”** bukan merupakan Tugas Akhir orang lain baik sebagian maupun keseluruhan kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya dan apabila pernyataan ini tidak benar penulis bersedia mendapatkan sanksi akademik.

Jember, 30 April 2021

Yang menyatakan,



Johan Taruna Wijaya
NIM. 1610651068

KATA PENGANTAR

Bismillahirrahmanirrahim

Alhamdulillah, segala puji bagi Allah SWT atas rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan laporan tugas akhir ini. Atas segala upaya, bimbingan, dan arahan dari semua pihak, penulis mengucapkan terimakasih sebesar-besarnya kepada :

1. Allah SWT yang telah melimpahkan rahmat dan hidayahnya, sehingga saya dapat menyelesaikan tugas akhir ini.
2. Bapak Dr. Nanang Saiful Rizal, S.T., M.T selaku dekan Fakultas Teknik Universitas Muhammadiyah Jember.
3. Bapak Triawan Adi Cahyanto, M.Kom selaku Kepala Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember.
4. Bapak Hardian Oktavianto, S.Si., M.Kom selaku dosen pembimbing I tugas akhir yang telah memberikan bimbingan serta pengarahan sehingga tugas akhir ini dapat terselesaikan.
5. Ibu Habibatul Azizah Al Faruq, M.Pd selaku dosen pembimbing II laporan tugas akhir yang telah memberikan bimbingan serta pengarahan sehingga tugas akhir ini dapat terselesaikan.
6. Ibu Ir, Dewi Lusiana, MT selaku dosen penguji I yang telah memberikan kritik dan saran yang membangun dalam penelitian ini.
7. Ibu Yeni Dwi Rahayu, M.Kom selaku dosen penguji II yang telah memberikan kritik dan saran yang membangun dalam penelitian ini.
8. Keluarga saya terutama ayah saya H, Jayus Imam dan ibu saya Hj Susetyo Handayani, yang telah banyak mendukung dan memberikan doanya dalam penyusunan laporan tugas akhir ini.
9. Teman-Teman “Jarkom Family & Klan Data Mining”, terutama Hamdan, Farid, dan Feni terimakasih atas waktu, dan dukungannya.
10. Teman-teman “Himajo dan Katrine club”, yang sudah menghibur saat-saat patah semangat & depresi karena revisian.

MOTTO

“Jawaban dari Sebuah Keberhasilan Adalah Terus Belajar dan Tak Kenal Menyerah.”

(John Rohn)

“Tidak Masalah jika kamu berjalan dengan lambat, asalkan kamu tidak pernah berhenti berusaha”

(Confucius)

“Pedang Terbaik yang Dimiliki ialah Sebuah Kesabaran Tanpa Batas.”

(Nelson Mandela)

“Jangan Menunda Pekerjaan Sampai Besok, Jika Hari Ini Dapat Diselesaikan ,Maka Tuntaskanlah”

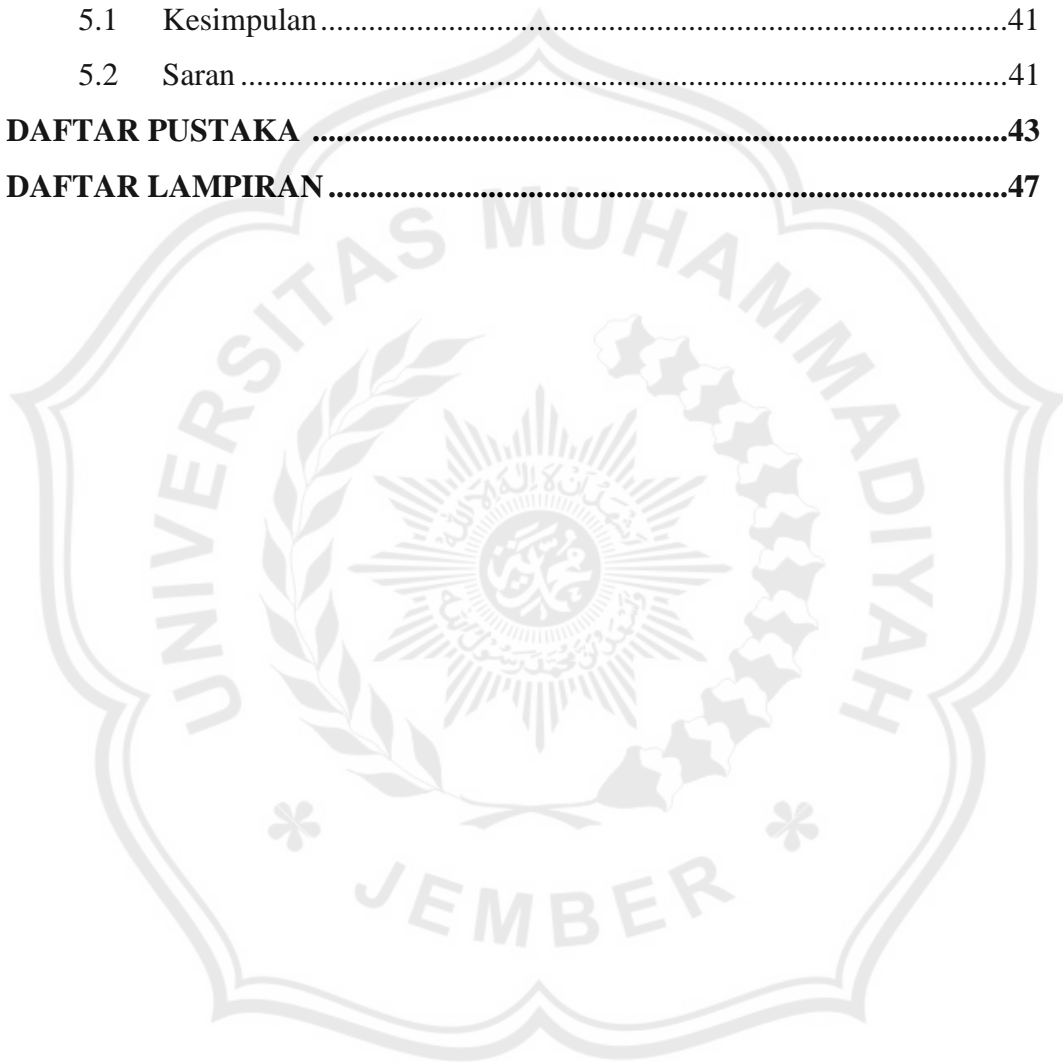
(Jhon Wick)

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
HALAMAN PERNYATAAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
MOTTO	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR GRAFIK	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Penelitian	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA	4
2.1 Kanker Payudara	4
2.2 <i>Machine Learning</i>	4
2.3 Klasifikasi	5
2.4 <i>K-Nearest Neighbor</i>	6
2.5 <i>Naive Bayes</i>	7
2.5.1 <i>Gaussian Naïve Bayes</i>	7
2.6 <i>Jupyter Notebook</i>	8
2.7 <i>k-Fold Cross-Validation</i>	8

2.8	<i>Confusion Matrix</i>	9
2.9	Akurasi, Presisi, dan <i>Recall</i>	9
BAB III METODE PENELITIAN		11
3.1	Pengumpulan Data	12
3.2	Penerapan Algoritma	12
3.2.1	<i>K-NEAREST NEIGHBORS</i> (K-NN).....	12
3.2.1.1	Data & Nilai K	13
3.2.1.2	Jarak Euclidean	13
3.2.1.3	Hasil Klasifikasi	15
3.2.2	<i>GAUSSIAN NAIVE BAYES</i> (GNB)	17
3.2.2.1	Data	18
3.2.2.2	Nilai Probabilitas	18
3.2.2.3	Nilai Probabilitas Data Kontinu	18
3.2.2.4	Nilai Hasil Data Kontinu	20
3.2.2.5	Klasifikasi	20
3.3	Validasi dan Evaluasi.....	21
BAB IV PEMBAHASAN.....		22
4.1	Data.....	22
4.2	Proses Klasifikasi	23
4.2.1	Pengujian Menggunakan <i>Kfold 2</i> Algoritma KNN.....	23
4.2.2	Pengujian Menggunakan <i>Kfold 2</i> Algoritma GNB	26
4.2.3	Pengujian Menggunakan <i>Kfold 3</i> Algoritma KNN.....	27
4.2.4	Pengujian Menggunakan <i>Kfold 3</i> Algoritma GNB	29
4.2.5	Pengujian Menggunakan <i>Kfold 4</i> Algoritma KNN.....	30
4.2.6	Pengujian Menggunakan <i>Kfold 4</i> Algoritma GNB	32
4.2.7	Pengujian Menggunakan <i>Kfold 5</i> Algoritma KNN.....	33
4.2.8	Pengujian Menggunakan <i>Kfold 5</i> Algoritma GNB	35
4.3	Perbandingan Kinerja Algoritma	37
4.3.1	Perbandingan hasil Algoritma Pada <i>Kfold 2</i>	37

4.3.2	Perbandingan hasil Algoritma Pada <i>Kfold</i> 3	38
4.3.3	Perbandingan hasil Algoritma Pada <i>Kfold</i> 4	39
4.3.4	Perbandingan hasil Algoritma Pada <i>Kfold</i> 5	40
BAB V KESIMPULAN DAN SARAN		41
5.1	Kesimpulan	41
5.2	Saran	41
DAFTAR PUSTAKA		43
DAFTAR LAMPIRAN		47



DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	10
Gambar 3.2 <i>Flowchart</i> Algoritma <i>KNN</i>	13
Gambar 3.3 <i>Flowchart</i> Algoritma <i>Gaussian Naïve bayes</i>	17
Gambar 3.4 Hasil Data Kontinu Untuk <i>Class 1</i>	20
Gambar 3.5 Hasil Data Kontinu Untuk <i>Class 2</i>	20
Gambar 4.1 Proses Memasukan Data.....	23
Gambar 4.2 Hasil pencarian K Tertinggi pada <i>Kfold 2</i>	24
Gambar 4.3 <i>Kfold 2</i> Algoritma <i>KNN</i>	24
Gambar 4.4 <i>Kfold 2</i> Algoritma <i>GNB</i>	26
Gambar 4.5 Hasil pencarian K Tertinggi pada <i>Kfold 3</i>	27
Gambar 4.6 <i>Kfold 3</i> Algoritma <i>KNN</i>	28
Gambar 4.7 <i>Kfold 3</i> Algoritma <i>GNB</i>	29
Gambar 4.8 Hasil pencarian K Tertinggi pada <i>Kfold 4</i>	31
Gambar 4.9 <i>Kfold 4</i> Algoritma <i>KNN</i>	31
Gambar 4.10 <i>Kfold 4</i> Algoritma <i>GNB</i>	32
Gambar 4.11 Hasil pencarian K Tertinggi pada <i>Kfold 5</i>	34
Gambar 4.12 <i>Kfold 5</i> Algoritma <i>KNN</i>	34
Gambar 4.13 <i>Kfold 5</i> Algoritma <i>GNB</i>	35
Gambar 4.14 Hasil Pengujian Pada <i>Kfold 2</i>	37
Gambar 4.14 Hasil Pengujian Pada <i>Kfold 3</i>	38
Gambar 4.14 Hasil Pengujian Pada <i>Kfold 4</i>	39
Gambar 4.14 Hasil Pengujian Pada <i>Kfold 5</i>	40

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	9
Tabel 2.2 Rumus Akurasi, Presisi dan Recall	10
Tabel 3.1 Contoh Data Yang Akan Diuji	12
Tabel 3.2 Nilai <i>Euclidean</i>	14
Tabel 3.3 <i>Rank</i> Kedekatan Jarak <i>Euclidean</i>	14
Tabel 3.4 Pembagian Data Training dan Data Testing di Tiap Percobaan	15
Tabel 3.5 Hasil Klasifikasi KNN	16
Tabel 3.6 Hasil Probabilitas Tiap Kelas.....	18
Tabel 3.7 Nilai Probabilitas Data Kontinu <i>Class 1</i>	19
Tabel 3.8 Nilai Probabilitas Data Kontinu <i>Class 2</i>	19
Tabel 3.9 Perbandingan <i>Class 1</i> dan <i>Class 2</i>	21
Tabel 4.1 <i>Confusion Matrix</i> Pada Algoritma <i>KNN Kfold 2</i>	25
Tabel 4.2 <i>Confusion Matrix</i> pada algoritma <i>GNB Kfold 2</i>	26
Tabel 4.3 <i>Confusion Matrix</i> Pada Algoritma <i>KNN Kfold 3</i>	28
Tabel 4.4 <i>Confusion Matrix</i> pada algoritma <i>GNB Kfold 3</i>	30
Tabel 4.5 <i>Confusion Matrix</i> Pada Algoritma <i>KNN Kfold 4</i>	32
Tabel 4.6 <i>Confusion Matrix</i> pada algoritma <i>GNB Kfold 4</i>	33
Tabel 4.7 <i>Confusion Matrix</i> Pada Algoritma <i>KNN Kfold 5</i>	35
Tabel 4.8 <i>Confusion Matrix</i> pada algoritma <i>GNB Kfold 5</i>	36