

TUGAS AKHIR

**PERBANDINGAN AKURASI ALGORITMA *MULTINOMIAL NAIVE BAYES* (MNB)
DAN *K-NEAREST NEIGHBOUR* (KNN) DALAM KLASIFIKASI
ABSTRAK TUGAS AKHIR
(STUDI KASUS: FAKULTAS TEKNIK UNIVERSITAS MUHAMMADIYAH JEMBER)**



YUKA REKSA NUGROHO

1610651052

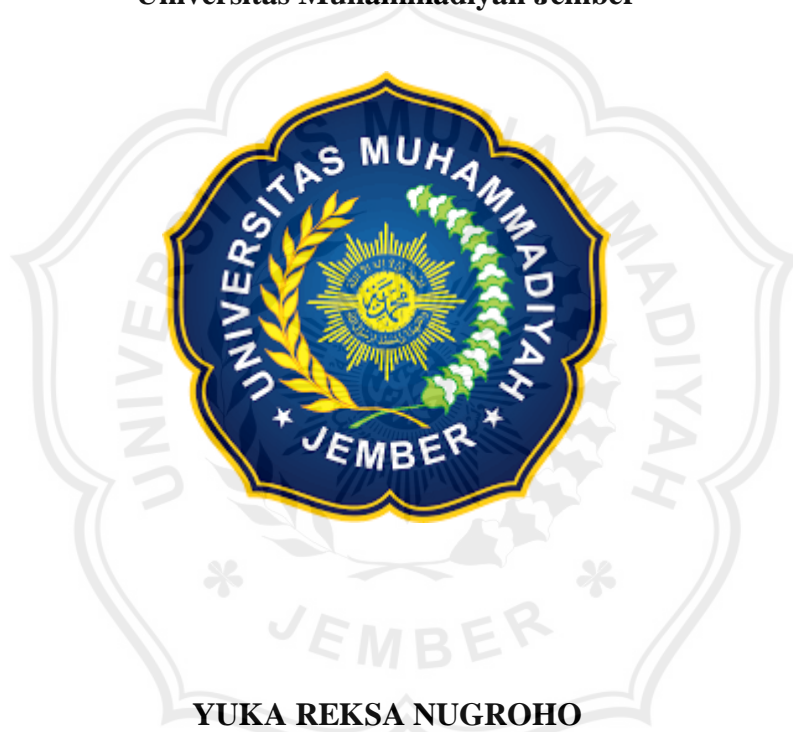
**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER**

2020

TUGAS AKHIR

**PERBANDINGAN AKURASI ALGORITMA *MULTINOMIAL NAIVE BAYES* (MNB)
DAN *K-NEAREST NEIGHBOUR* (KNN) DALAM KLASIFIKASI
ABSTRAK TUGAS AKHIR
(STUDI KASUS: FAKULTAS TEKNIK UNIVERSITAS MUHAMMADIYAH JEMBER)**

**Disusun untuk Melengkapi dan Memenuhi Syarat Kelulusan Progam Strata 1
Progam Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember**



YUKA REKSA NUGROHO

1610651052

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

2020

HALAMAN PERSETUJUAN

**PERBANDINGAN AKURASI ALGORITMA *MULTINOMIAL NAIVE BAYES (MNB)*
DAN *K-NEAREST NEIGHBOUR (KNN)* DALAM KLASIFIKASI
ABSTRAK TUGAS AKHIR**

Oleh:

Yuka Reksa Nugroho

1610651052

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

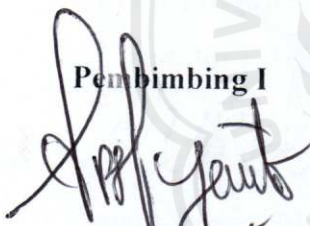
Sarjana Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

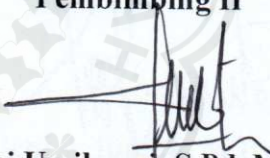
Disetujui oleh:

Pembimbing I


Deni Arifianto, M.Kom.

NIDN. 0718068103

Pembimbing II


Reni Umilasari, S.Pd, M.Si.

NIDN. 0728079101

HALAMAN PENGESAHAN

**PERBANDINGAN AKURASI ALGORITMA *MULTINOMIAL NAIVE BAYES (MNB)* DAN
K-NEAREST NEIGHBOUR (KNN) DALAM KLASIFIKASI
ABSTRAK TUGAS AKHIR**

Oleh:

Yuka Reksa Nugroho

1610651052

Telah mempertanggung jawabkan Laporan Tugas Akhirnya pada Sidang Tugas Akhir tanggal
17 Juli 2020 sebagai salah satu syarat kelulusan dan mendapatkan gelar

Sarjana Komputer (S.Kom)

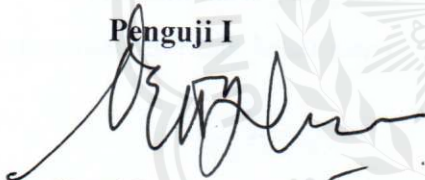
di

Universitas Muhammadiyah Jember

Disetujui oleh:

Dosen Penguji:

Penguji I



Ir. Dewi Lusiana, M.T.

NIDN. 0712086702

Penguji II

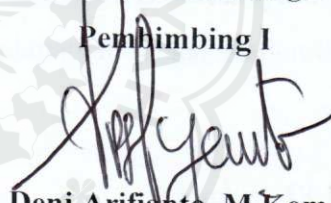


Rosita Yanuarti, S.Kom., M. Cs.

NIDN. 0629018601

Dosen Pembimbing:

Pembimbing I



Deni Arifianto, M.Kom.

NIDN. 0718068103

Pembimbing II



Reni Umilasari, S.Pd, M.Si.

NIDN. 0728079101

Mengesahkan,

Dekan Fakultas Teknik



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 dibawah ini:

NAMA : YUKA REKSA NUGROHO
NIM : 1610651052
INSTITUSI : Strata-1 Program Studi Teknik Informatika Fakultas Teknik,
Universitas Muhammadiyah Jember

Menyatakan bahwa Tugas Akhir yang berjudul “**Perbandingan Akurasi Algoritma *Multinomial Naive Bayes (MNB)* Dan *K-Nearest Neighbour (KNN)* Dalam Klasifikasi Abstrak Tugas Akhir**” 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, 18 Juli 2020

Yang Menyatakan,



Yuka Reksa Nugroho

NIM. 1610651052

KATA PENGANTAR

Alhamdulillah segala puji bagi Allah yang senantiasa memberikan nikmat yang sangat luar biasa dalam hidup ini kepada penulis berupa nikmat untuk selalu bersujud dan meminta ampunan kepada-Nya. Serta berkat rahmat-Nya penulis diberikan kemudahan dalam menyelesaikan studi di kampus Universitas Muhammadiyah Jember.

Atas segala upaya, bimbingan dan arahan dari semua pihak, penulis mengucapkan banyak terima kasih. Dalam kesempatan ini penulis mengucapkan terima kasih sebanyak-banyaknya kepada :

1. Allah SWT atas segala rahmat, nikmat dan hidayah-Nya, terima kasih atas kehendak-Mu telah memberikan kesempatan ini untuk mendapatkan gelar Sarjana Komputer.
2. Bapak Nanang Saiful Rizal, S.T, M.T. selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
3. Bapak Triawan Adi Cahyanto, M.Kom., selaku Ketua Program Studi Teknik Informatika Universitas Muhammadiyah Jember.
4. Bapak Deni Arifianto, M.Kom., selaku dosen pembimbing I dan Ibu Reni Umilasari, S.Pd, M.Si., selaku dosen pembimbing II yang telah meluangkan waktu untuk memberi arahan dan membimbing penulis dalam menyelesaikan tugas akhir ini.
5. Ibu Ir. Dewi Lusiana, M.T., selaku dosen penguji I dan Ibu Rosita Yanuarti, S.Kom., M. Cs., selaku dosen penguji II yang telah memberikan saran dan masukan dalam penyelesaian tugas akhir ini.
6. Kepada Bapak dan Ibu Dosen Program Studi Teknik Informatika Universitas Muhammadiyah Jember yang telah memberikan ilmunya kepada penulis.
7. Kepada Bapak dan Ibu Staf dan Laboran Program Studi Teknik Informatika Universitas Muhammadiyah Jember yang telah membantu penulis dalam melaksanakan tugas akhir.
8. Kepada orang tua penulis Bapak Yusak Joko Nugroho dan Ibu

Kasmawati Majid yang selalu mendoakan, memberikan motivasi, dukungan dan memberikan kasih sayang sehingga penulis dapat menyelesaikan tugas akhir ini.

9. Kepada Pengumpul Token, Feni, Farid, dan Hamdhan yang telah banyak membantu dan memberikan semangat, dukungan, bantuan, nasihat, saran, serta doa ke penulis untuk menyelesaikan tugas akhir ini.
10. Kepada keluarga Jarkom, yang telah membantu untuk memberikan semangat, dukungan, bantuan, nasihat, saran, dan doa ke penulis untuk menyelesaikan tugas akhir ini.
11. Serta kepada teman-teman Program Studi Teknik Informatika Angkatan 2016 yang telah memberikan doa, dukungan, dan bantuannya.

Dan untuk semuanya, semoga Allah menunjukkan kepada kita pada jalan yang telah diridhoi oleh-Nya, diberikan ilmu yang barokah, serta diampuni dosa-dosa kita, dan senantiasa tercurah limpahkan rahmat, taufiq dan hidayah-Nya kepada kita semua.

Aamiin... Yaa Robbal Alamiin...

Jember, 18 Juli 2020

Penulis

MOTTO

“Ketika kau sedang mengalami kesusahan dan bertanya-tanya kemana Allah, cukup ingat bahwa seorang guru selalu diam saat ujian berjalan.”

– Nourman Ali Khan

“Hidup itu adalah seni menggambar tanpa penghapus.”

– John W. Gardner

“Hidup dapat dipahami dengan berpikir ke belakang. Tapi ia juga harus dijalani dengan berpikir ke depan.”

– Soren Kiekegaard

“Bekerja keras dan bersikap baiklah. Hal luar biasa akan terjadi.”

– Conan O'Brien

“If you're too open-minded, your brains will fall out.”

– Lawrence Ferlinghetti

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
HALAMAN PERNYATAAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
MOTTO	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiv
DAFTAR GRAFIK	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA	5
2.1 Pengertian Abstrak	5
2.2 <i>Text Mining</i>	5
2.3 <i>Pre-Processing Text</i>	6
2.4 Pembobotan TF-IDF	7
2.4.1 <i>Term Frequency</i>	8
2.4.2 <i>Inverse Document Frequency (IDF)</i>	8
2.4.3 <i>Term Weighting TF-IDF</i>	9
2.5 <i>Naïve Bayes Classifier</i>	9
2.6 <i>Multinomial Naïve Bayes</i>	10
2.7 <i>K-Nearest Neighbour</i>	11

2.8	<i>K-Folf Cross Validation</i>	12
2.9	<i>Python</i>	13
2.10	<i>Natural language Toolkit (NLTK)</i>	13
2.11	<i>Scikit-Learen</i>	13
2.12	<i>Library Sastrawi</i>	14
2.13	<i>Jupyter Notebook</i>	15
2.14	<i>Confusion Matrix</i>	15
BAB III METODE PENELITIAN		17
3.1	Rancangan Penelitian.....	17
3.2	Pengumpulan Data	18
3.3	Pelabelan data.....	18
3.4	<i>Pre-Processing</i>	18
3.5	Pembobotan	21
3.6	Klasifikasi	34
3.6.1	Pelatihan Metode <i>Multinomial Naive Bayes</i>	36
3.6.2	Pengujian Metode <i>Multinomial Naive Bayes</i>	37
3.6.3	Pelatihan Metode <i>K-Nearest Neighbour</i>	41
3.6.4	Pengujian Metode <i>K-Nearest Neighbour</i>	41
3.7	Evaluasi.....	42
BAB IV PEMBAHASAN		43
4.1	Data.....	43
4.2	<i>Pre-processing data</i>	44
4.3	Implementasi Pembobotan.....	45
4.4	Hasil Klasifikasi.....	46
4.4.1	Hasil Klasifikasi <i>Multinomial Naive Bayes</i> dan <i>K-Nearest Neighbour Fold K = 2</i>	50
4.4.2	Hasil Klasifikasi <i>Multinomial Naive Bayes</i> dan <i>K-Nearest Neighbour Fold K = 4</i>	52
4.4.3	Hasil Klasifikasi <i>Multinomial Naive Bayes</i> dan <i>K-Nearest Neighbour Fold K = 5</i>	54

4.4.4 Hasil Klasifikasi <i>Multinomial Naive Bayes</i> dan <i>K-Nearest Neighbour Fold K = 10</i>	56
BAB V KESIMPULAN DAN SARAN	61
5.1 Kesimpulan	61
5.2 Saran	62
DAFTAR PUSTAKA	63
LAMPIRAN	66

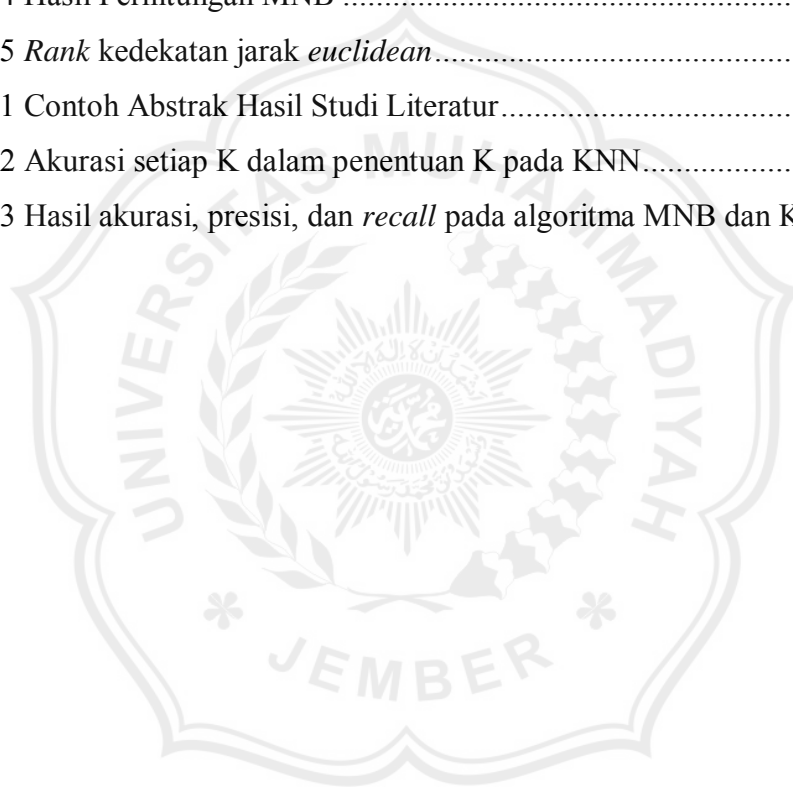


DAFTAR GAMBAR

Gambar 2.1 Contoh <i>K-Fold Cross Validation</i>	12
Gambar 3.1 Tahapan Rancangan Penelitian	17
Gambar 3.2 Alur <i>Text Pre-processing</i>	18
Gambar 3.3 Alur proses klasifikasi.....	34
Gambar 3.4 <i>Flowchart</i> alur klasifikasi MNB dan KNN.....	35
Gambar 4.1 <i>Pre-processing</i> pada <i>Jupyter Notebook</i>	44
Gambar 4.2 <i>Source code</i> pelabelan data	45
Gambar 4.3 Proses dan hasil dari pembobotan pada <i>Jupyter Notebook</i>	46
Gambar 4.4 Klasifikasi algoritma <i>Multinomial Naïve Bayes</i>	47
Gambar 4.5 Klasifikasi algoritma <i>K-Nearest Neighbour</i>	49
Gambar 4.6 <i>Confusion Matrix</i> pada algoritma MNB, <i>fold K = 2</i>	50
Gambar 4.7 <i>Confusion Matrix</i> pada algoritma KNN, <i>fold K = 2</i>	51
Gambar 4.8 <i>Confusion Matrix</i> pada algoritma MNB, <i>fold K = 4</i>	52
Gambar 4.9 <i>Confusion Matrix</i> pada algoritma KNN, <i>fold K = 4</i>	53
Gambar 4.10 <i>Confusion Matrix</i> pada algoritma MNB, <i>fold K = 5</i>	55
Gambar 4.11 <i>Confusion Matrix</i> pada algoritma KNN, <i>fold K = 5</i>	55
Gambar 4.12 <i>Confusion Matrix</i> pada algoritma MNB, <i>fold K = 10</i>	57
Gambar 4.13 <i>Confusion Matrix</i> pada algoritma KNN, <i>fold K = 10</i>	58

DAFTAR TABEL

Tabel 2.1 Contoh <i>Pre-processing Text</i>	7
Tabel 2.2 <i>Term-document Matrix</i>	8
Tabel 2.3 Representasi hasil proses klasifikasi <i>Confusion Matrix</i>	16
Tabel 3.1 Perhitungan <i>Term Frequency</i>	21
Tabel 3.2 Perhitungan Pembobotan IDF	26
Tabel 3.3 Hasil Perhitungan <i>TF-IDF</i>	30
Tabel 3.4 Hasil Perhitungan MNB	39
Tabel 3.5 <i>Rank</i> kedekatan jarak <i>euclidean</i>	42
Tabel 4.1 Contoh Abstrak Hasil Studi Literatur	43
Tabel 4.2 Akurasi setiap K dalam penentuan K pada KNN	48
Tabel 4.3 Hasil akurasi, presisi, dan <i>recall</i> pada algoritma MNB dan KNN	59



DAFTAR GRAFIK

Grafik 4.1 Penentuan K terbaik dari KNN	47
Grafik 4.2 Hasil Perhitungan Akurasi Pada <i>Fold</i> K=2	51
Grafik 4.3 Hasil Perhitungan Akurasi Pada <i>Fold</i> K=4	54
Grafik 4.4 Hasil Perhitungan Akurasi Pada <i>Fold</i> K=5	56
Grafik 4.5 Hasil Perhitungan Akurasi Pada <i>Fold</i> K=10	58

