

**IMPLEMENTASI ALGORITMA NAIVE BAYES CLASSIFIER UNTUK
KLASIFIKASI BIDANG DOSEN BERDASARKAN JUDUL PENELITIAN
MAHASISWA**

SKRIPSI



**Disusun oleh
Muhammad Farhan Khoirulamri Darmawan
18.11.2273**

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2022**

**IMPLEMENTASI ALGORITMA NAIVE BAYES CLASSIFIER UNTUK
KLASIFIKASI BIDANG DOSEN BERDASARKAN JUDUL PENELITIAN
MAHASISWA**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



Disusun oleh
Muhammad Farhan Khoirulamri Darmawan
18.11.2273

PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2022

PERSETUJUAN

SKRIPSI

IMPLEMENTASI ALGORITMA NAIVE BAYES CLASSIFIER UNTUK KLASIFIKASI BIDANG DOSEN BERDASARKAN JUDUL PENELITIAN MAHASISWA

yang dipersiapkan dan disusun oleh
Muhammad Farhan Khoirulamri Darmawan
18.11.2273

Telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 20 Maret 2022

Dosen Pembimbing,

Windha Mega Pradnya D.M.Kom
NIK. 190302185

PENGESAHAN

SKRIPSI

IMPLEMENTASI ALGORITMA NAIVE BAYES CLASSIFIER UNTUK KLASIFIKASI BIDANG DOSEN BERDASARKAN JUDUL PENELITIAN MAHASISWA

yang dipersiapkan dan disusun oleh

Muhammad Farhan Khoirulamri Darmawan

18.11.2273

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Februari 2022

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Erni Seniwati, S.Kom, M.Cs

NIK. 190302231

Nuraini, M.Kom

NIK. 190302066

Windha Mega P.D, M.Kom

NIK. 190302185

Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer

Tanggal 2 Maret 2022

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, M.Kom

NIK. 190302096

PERNYATAAN

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah dipublikasi oleh orang lain untuk memperoleh gelar akademis di suatu instansi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara sah dan/atau secara sah dan/atau disebut dalam daftar pustaka.

Seperti tersebut yang tertera dengan sah dan/atau yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 30 September 2021



Muhammad Fauzan Ghobrialanif Darmasari

NIM. 18.11.2773

Scanned by TapScanner

MOTTO

"Sesungguhnya sesudah kesulitan itu ada kemudahan"

(Al-Insyirah :6)

"...Allah tidak membebani seseorang itu melainkan sesuai dengan kesanggupannya.."

(Al-Baqarah :286)

"Janganlah kamu bersikap lemah dan janganlah pula kamu bersedih hati, padahal kamulah orang-orang yang paling tinggi derajatnya jika kamu beriman."

(Ali-Imran :139)

"Work hard in silence, let success be your noise."

(Frank)

"The best preparation for tomorrow is doing your best today."

(H. Jackson Brown Jr.)

"Aja dadi uwong sing rumangsa bisa lan rumangsa pinter. Nanging dadiya uwong sing bisa lan pinter rumangsa."

(Peribahasa Bahasa Jawa)

"Mohon, mangesthi, mangastuti, marem."

(Peribahasa Bahasa Jawa)

PERSEMBAHAN

Alhamdulillah kupanjatkan puji syukur atas rahmat Allah SWT dan bantuan atau bimbingan dari berbagai pihak sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Oleh karena itu Skripsi ini penulis persembahkan kepada :

1. Terima kasih kepada Allah SWT atas rahmat dan hidayahnya. Serta kepada Rasulullah Muhammad SAW yang telah membawa manusia dari zaman kegelapan menuju zaman kecerahan.
2. Bapak dan ibu tersayang, terima kasih atas segala dukungan kalian. Skripsi ini dipersembahkan sebagai wujud rasa terima kasih atas pengorbanan dan jerih payah kalian.
3. Terima Kasih kepada Ibu Windha Mega Pradnya selaku dosen pembimbing.
4. Terima Kasih kepada Andi Setyawan, Muhammad Pandi Fery, Nurrofiqi, Agita Putri, dan Lasmita yang sudah membantu dan mendukung selama proses pengerjaan skripsi.
5. Terima kasih kepada Keluarga besar HMIF AMIKOM yang selalu memberikan dukungan.
6. Terima Kasih kepada teman seperjuangan kelas 18-IF-07 yang telah sama sama berjuang.
7. Teman-Teman yang belum saya sebutkan. Terima kasih atas doa dan semangat yang kalian berikan.
8. Someone special in my heart... terima kasih untuk semuanya. Aku harap kamu selalu sukses baik di dunia dan akhirat.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Implementasi Algoritma Naïve Bayes Classifier Untuk Klasifikasi Bidang Dosen Berdasarkan Judul Penelitian Mahasiswa”.

Selama penulisan skripsi ini, penulis banyak menerima bantuan dan dukungan sehingga dapat menyelesaikan skripsi ini. Oleh karena itu, penulis mengucapkan terima kasih sebesar-besarnya kepada:

1. Bapak Prof. Dr. M. Suyanto selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom, M.Kom selaku Dekan Fakultas Ilmu Komputer.
3. Ibu Windha Mega Pradnya D, M.Kom selaku Ketua Program Studi Informatika dan selaku Dosen Pembimbing yang selalu memberi arahan sehingga skripsi ini selesai dengan baik.
4. Kepada orang tua dan segenap keluarga yang telah memberikan doa dan dukungan demi tercapainya cita-cita penulis.
5. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah membantu sehingga skripsi ini dapat selesai.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, oleh karena itu, semua kritik dan saran yang bersifat membangun akan penulis terima dengan senang hati. Penulis berharap, semoga skripsi ini dapat bermanfaat kedepan-nya.

Yogyakarta, 4 February 2022

Muhammad Farhan Khoirulamri Darmawan

18.11.2273

DAFTAR ISI

COVER.....	0
JUDUL.....	i
PERSETUJUAN.....	ii
PENGESAHAN.....	iii
PERNYATAAN.....	iv
MOTTO.....	v
PERSEMBAHAN.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
INTISARI.....	xii
ABSTRACT.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	1
1.3 Batasan Masalah.....	2
1.4 Maksud dan Tujuan Penelitian.....	2
1.5 Manfaat Penelitian.....	2
1.6 Metodologi Penelitian.....	2
1.6.1 Metode Pengumpulan Data.....	3
1.6.2 Metode Analisis.....	3
1.6.3 Metode Perancangan.....	4
1.6.4 Metode Testing.....	4
1.7 Sistematika Penulisan.....	4
BAB II LANDASAN TEORI.....	6
2.1 Kajian Pustaka.....	6
2.2 Daftar Teori I.....	10
2.2.1 Data Mining.....	10
2.2.2 Text Mining.....	10
2.2.3 Text Preprocessing.....	10
2.2.4 Klasifikasi.....	11
2.2.5 Pembobotan Kata (<i>Term Weighting</i>).....	12
2.2.6 Algoritma Naïve Bayes Classifier.....	14

2.2.7	Confusion Matrix.....	16
2.2.8	Metode Black Box	18
2.2.9	Python.....	20
BAB III ANALISIS DAN PERANCANGAN		21
3.1	Tinjauan Umum.....	21
3.2	Alat dan Bahan Penelitian.....	21
3.2.1	Perangkat Keras (<i>Hardware</i>)	21
3.2.2	Perangkat Lunak (<i>Software</i>)	22
3.3	Alur Penelitian.....	23
3.3.1	Data Privat Universitas Amikom.....	23
3.3.2	Preprocessing	28
3.3.3	Data Split	32
3.3.4	Feature Extraction	33
3.3.5	Naïve Bayes Classifier	43
3.3.6	Confusion Matrix.....	44
3.4	Perancangan Antar Muka (<i>Interface</i>).....	44
BAB IV IMPLEMENTASI DAN PEMBAHASAN		46
4.1	Deskripsi Implementasi.....	46
4.2	Implementasi Memasukkan Data	46
4.3	Implementasi Pelabelan Data	46
4.4	Implementasi Perancangan Fungsi dan Sistem.....	48
4.4.1	Implementasi Fungsi Preprocessing	48
4.4.2	Implementasi Data Split.....	50
4.4.3	Implementasi TF-IDF	50
4.4.4	Implementasi Algoritma Naïve Bayes Classifier.....	51
4.4.5	Implementasi Confusion Matrix	52
4.5	Implementasi Interface.....	53
4.5.1	Halaman Utama	53
4.5.2	Proses Input Judul	54
4.5.3	Hasil Klasifikasi Bidang Dosen.....	55
4.5.4	Hasil List Dosen Yang Sesuai Dengan Bidang.....	56
BAB V KESIMPULAN.....		58
5.1	Kesimpulan.....	58
5.2	Saran.....	58
DAFTAR PUSTAKA		59
LAMPIRAN.....		60

DAFTAR TABEL

Table 2. 1 Daftar Penelitian Kajian Pustaka	8
Table 3. 1 Spesifikasi Hardware	22
Table 3. 2 Perangkat Lunak (Software)	22
Table 3. 3 Dataset Utuh	25
Table 3. 4 Dataset Utuh dengan 8 Kolom	26
Table 3. 5 Contoh Case Folding	28
Table 3. 6 Tahapan Menghilangkan Link dan Angka	30
Table 3. 7 Tahapan Menghilangkan Tanda Baca	30
Table 3. 8 Tahapan Filtering/Stopword Removal	31
Table 3. 9 Term Kata	33
Table 3. 10 Perhitungan IDF	36
Table 3. 11 Perhitungan TF-IDF	40
Table 3. 12 Data Uji	43
Table 4. 1 Perbandingan Data Testing	53

DAFTAR GAMBAR

Gambar 2. 1 Persamaan TF-IDF	13
Gambar 2. 2 Persamaan Teorema Bayes	14
Gambar 3. 1 Alur Penelitian	23
Gambar 3. 2 Labeling Data	28
Gambar 3. 3 Data Split	33
Gambar 3. 4 Interface Halaman Utama	45
Gambar 4. 1 Menambahkan Kolom 'Bidang Dosen'	46
Gambar 4. 2 Script Pelabelan Data 'Bidang Dosen'	47
Gambar 4. 3 Hasil Pelabelan Data 'Bidang Dosen'	47
Gambar 4. 4 Script Ubah 'Bidang Dosen' Menjadi Case Folding	48
Gambar 4. 5 Hasil Ubah 'Bidang Dosen' Menjadi Case Folding	48
Gambar 4. 6 Script Open File .xlsx	48
Gambar 4. 7 Script Case Folding	49
Gambar 4. 8 Script untuk menghilangkan link atau angka	49
Gambar 4. 9 Script untuk menghilangkan tanda baca	49
Gambar 4. 10 Script Stopword	50
Gambar 4. 11 Script Data Split	50
Gambar 4. 12 Script Mengubah Teks Menjadi Angka	51
Gambar 4. 13 Script Untuk Mengetahui Bobot Setiap Kata	51
Gambar 4. 14 Script Akurasi NBC	52
Gambar 4. 15 Script Confusion Matrix	52
Gambar 4. 16 Implementasi Confusion Matrix	53
Gambar 4. 17 Implementasi Halaman Utama	54
Gambar 4. 18 Kolom Input Judul	55
Gambar 4. 19 Input Judul Random	55
Gambar 4. 20 Hasil Rekomendasi Klasifikasi Bidang Dosen	55
Gambar 4. 21 Hasil List Dosen Yang Sesuai dengan Bidang	56
Gambar 4. 22 Hasil Rekomendasi dan List Dosen Data Random	57

INTISARI

Pada perguruan tinggi, khususnya untuk S1 syarat kelulusan adalah menyelesaikan skripsi. Untuk dapat menyelesaikan skripsi, mahasiswa akan didampingi oleh dosen pembimbing. Dosen pembimbing sendiri merupakan dosen yang ditunjuk oleh pimpinan jurusan untuk membimbing mahasiswa sesuai dengan bidang atau keahlian yang dimiliki dan tema skripsi yang dipilih. Masalahnya, masih ada mahasiswa yang sering mengeluhkan dosen pembimbing tidak sesuai dengan bidangnya sehingga mengakibatkan sulitnya dalam proses bimbingan.

Dalam penelitian kali ini, data yang digunakan diperoleh dari Universitas AMIKOM Yogyakarta Program Studi Informatika tahun 2014 hingga 2016 yang terdiri dari 1598 data. Kemudian, dataset tersebut akan diolah melalui proses Preprocessing, pembobotan kata dan diklasifikasikan menggunakan algoritma naïve bayes classifier apakah dataset tersebut sudah sesuai dengan bidang dosen terdiri dari Data Science, Multimedia, Networking, Software Engineer, dan Game. Setelah itu dataset dibagi menjadi data training dan data test sebanyak 9:1 (90% data training dan 10% data test).

Berdasarkan hasil penelitian yang telah dilakukan, hasil akurasi yang diperoleh dari klasifikasi menggunakan algoritma naïve bayes classifier sebesar 88.29%.

Kata Kunci: *Algoritma naïve bayes classifier, bidang dosen, mahasiswa.*

ABSTRACT

In college, especially for S-1 graduation requirements are to complete a thesis. To be able to complete the thesis, students will be accompanied by a guidance lecturer. The guidance lecturer himself is a lecturer appointed by the head of the department to guide students in accordance with the field or expertise owned and the theme of the thesis chosen. The problem is, there are still students who often complain that the guidance lecturer is not in accordance with the field, resulting in difficulty in the guidance process.

In this study, the data used was obtained from AMIKOM Yogyakarta University Informatics Study Program from 2014 to 2016 consisting of 1598 data. Then, the dataset will be processed through the process of preprocessing, word weighting and classified using naïve bayes classifier algorithm whether the dataset is in accordance with the lecturer field consisting of Data Science, Multimedia, Networking, Software Engineer, and Game. After that the dataset is divided into training data and test data as much as 9:1 (90% training data and 10% test data),

Based on the results of research that has been done, the accuracy results obtained from classification using the naïve algorithm bayes classifier amounted to 88.29%.

Keywords: *Algorithm naïve bayes classifier, field of lecturers, students.*

