



**Klasifikasi Data Mahasiswa untuk Merekomendasikan Kelompok Studi
Mahasiswa Menggunakan Algoritma Naive Bayes**

SKRIPSI

ALFIAN PRATAMA

NIM. 1910511050

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI INFORMATIKA

2023



**Klasifikasi Data Mahasiswa untuk Merekomendasikan Kelompok Studi
Mahasiswa Menggunakan Algoritma Naive Bayes**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana
Komputer**

ALFIAN PRATAMA

NIM. 1910511050

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI INFORMATIKA

2023

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua sumber referensi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Alfian Pratama
NIM : 1910511050
Tanggal : 17 Juli 2023

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan berlaku.

Jakarta, 17 Juli 2023

Yang Menyatakan,


10000
METRAL
TEMPEL
E4AD56XXB12797453
(Alfian Pratama)

PERNYATAAN PERSETUJUAN PUBLIKSASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Alfian Pratama
NIM : 1910511050
Fakultas : Ilmu Komputer
Program Studi : SI Informatika

Demu pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya berjudul :

Klasifikasi Data Mahasiswa untuk Merekomendasikan Kelompok Studi Mahasiswa Menggunakan Algoritma Naive Bayes

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 17 Juli 2023

Yang Menyatakan,



(Alfian Pratama)

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama : Alfian Pratama
NIM : 1910511050
Program Studi : S1 Informatika
Judul Tugas Akhir : Klasifikasi Data Mahasiswa untuk Merekomendasikan
Kelompok Studi Mahasiswa Menggunakan Algoritma Naive Bayes

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Henki Bayu Seta, S.Kom., MTI.

Penguji I



Jayanta, S.Kom., M.Si.

Penguji II



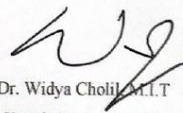
Nurhafifah Matondang, S.Kom., M.M., M.TI

Pembimbing



Dr. Ermatita, M.Kom.

Dekan



Dr. Widya Cholil, M.T.

Kepala Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 13 Juli 2023



ABSTRAK

Mahasiswa memiliki peran penting dalam berjalannya suatu pendidikan di tingkat universitas. Sebagai seorang mahasiswa, mengembangkan soft skill dan hard skill yang dimiliki merupakan hal yang sangat penting. Dua hal tersebut menjadi poin utama dalam pengembangan diri dan juga untuk merekomendasikan mahasiswa terkait dengan kelompok studi mahasiswa yang terdapat di Fakultas Ilmu Komputer UPN Veteran Jakarta. Permasalahan tersebut membuat sebuah pemikiran untuk dapat menciptakan suatu kelompok yang bisa mewadahi mahasiswa dalam mengembangkan diri di dalam perkuliahan. Namun karena hal tersebut, timbul masalah baru bagi mahasiswa yang tidak tahu arah untuk melangkah dalam mendalami suatu bidang tertentu. Oleh karena itu, untuk memudahkan mahasiswa yang memiliki keraguan dalam memilih kelompok studi mahasiswa, di dalam penelitian ini dilakukan klasifikasi menggunakan data mahasiswa untuk dapat menentukan pilihan kelompok studi mahasiswa. Dalam pengaplikasiannya, dataset didapatkan melalui survey menggunakan platform *google form* sebanyak 136 data mahasiswa Fakultas Ilmu Komputer Angkatan 2021-2022. Data yang diperoleh melewati beberapa proses yaitu praproses data, pembersihan data, dan transformasi data. Data yang diolah dibangun menggunakan algoritma klasifikasi naïve bayes dengan jenis multinomial naïve bayes dikarenakan hampir seluruh data memiliki jenis data kategorikal. Berdasarkan hasil penelitian, akurasi yang didapatkan oleh model dan setelah dilakukan perhitungan akurasi menggunakan confusion matrix terhadap model, diperoleh akurasi sebesar 85,7 %, presisi sebesar 100 %, dan nilai recall sebesar 77,7%.

Kata Kunci : Klasifikasi, Naive Bayes, Kelompok Studi Mahasiswa

ABSTRACT

Students have an important role in the running of an education at the university level. As a student, developing soft skills and hard skills is very important. These two things are the main points in self-development and also to recommend students related to student study groups at the UPN Veteran Jakarta Faculty of Computer Science. These problems make a thought to be able to create a group that can accommodate students in developing themselves in lectures. However, because of this, new problems arise for students who do not know which way to go in exploring a particular field. Therefore, to make it easier for students who have doubts in choosing a student study group, in this study a classification was carried out using student data to be able to determine the choice of student study group. In its application, the dataset was obtained through a survey using the Google form platform as many as 136 data from students of the Faculty of Computer Science Batch 2021-2022. The data obtained went through several processes, namely data pre-processing, data cleaning, and data transformation. The processed data is built using the naïve Bayes classification algorithm with the naïve Bayes multinomial type because almost all data has categorical data types. Based on the results of the study, the accuracy obtained by the model and after calculating the accuracy using the confusion matrix on the model, obtained an accuracy of 85.7%, a precision of 100%, and a recall value of 77.7%.

Keywords : Classify, Naive Bayes, Student Study Group

KATA PENGANTAR

Dengan memanjatkan puji dan syukur atas kehadiran Allah Swt., yang telah memberikan rahmat dan karunia-Nya sampai pada saat proposal skripsi ini telah dibuat. Berkat rahmat dan karunia-Nya, penulis dapat menyelesaikan proposal skripsi ini dengan judul “Klasifikasi Data Mahasiswa untuk Menentukan Pilihan Kelompok Studi Mahasiswa Menggunakan Algoritma Naive Bayes” dengan baik.

Adapun proses penyelesaian penulisan proposal skripsi ini tidak lepas dari bantuan berbagai pihak kepada penulis. Untuk itu, penulis mengucapkan terima kasih kepada:

1. Allah Swt.
2. Kedua orang tua yang selalu menjadi penyemangat dan memberikan dukungan kepada penulis sehingga penulis dapat menyelesaikan tugas akhir ini dengan baik dan tepat waktu
3. Ibu Nurhafifah Matondang, S.Kom., M.T.I. selaku dosen pembimbing proposal skripsi yang telah memberikan bimbingan dan saran selama berjalannya proposal skripsi.
4. Ibu Dr. Widya Cholil, S.Kom., M.I.T. selaku Ketua Program Studi S1 Informatika
5. Kepada Mutiara Dwi Jayanti selaku teman, sahabat, dan kekasih saya yang telah bersama dan mendukung saya sejak pertama kali skripsi ini dibuat sampai dengan sekarang skripsi ini selesai.
6. Seluruh pihak yang terlibat dalam kelancaran penulisan tugas akhir ini dan yang belum disebutkan di atas, penulis ucapkan terima kasih.

Akhir kata, penulis menyadari bahwa proposal skripsi ini masih jauh dari kata sempurna, hal tersebut disebabkan keterbatasan penulis dari segi pengetahuan dan kemampuan. Oleh karena itu, penulis mengharapkan saran dan kritik dari pembaca yang bersifat membangun agar penyusunan proposal skripsi ini dapat disempurnakan.

Jakarta, 30 Oktober 2022

Penulis

DAFTAR ISI

PERNYATAAN ORISINALITAS	i
PERNYATAAN PERSETUJUAN PUBLIKSASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....	ii
LEMBAR PENGESAHAN	iii
ABSTRAK	iv
ABSTRACT.....	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR.....	xii
DAFTAR SIMBOL	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I.....	1
PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	7
1.3. Tujuan Penelitian.....	8
1.4. Manfaat Penelitian.....	8
1.5. Ruang Lingkup.....	8
1.6. Luaran yang Diharapkan	8
1.7. Sistematika Penulisan.....	8
BAB II	10
TINJAUAN PUSTAKA.....	10
2.1. Mahasiswa	10
2.2. Pendidikan	10
2.3. Kelompok Studi Mahasiswa.....	11
2.4. Data Mining	13
2.4.1. Pengertian Data Mining	13
2.4.2. Teknik Data Mining	14
2.5. Machine Learning.....	16
2.6. Klasifikasi.....	18
2.7. Algoritma Naive Bayes	18

2.8. Studi Literatur.....	20
METODOLOGI PENELITIAN.....	25
3.1. Kerangka Pikir.....	25
3.1.1. Menentukan Topik.....	26
3.1.2. Identifikasi Masalah.....	26
3.1.3. Studi Literatur.....	27
3.1.4. Pengumpulan Data.....	27
3.1.5. Pengolahan Data.....	28
3.1.6. Pembagian Data.....	28
3.1.7. Data Latih.....	29
3.1.8. Training.....	29
3.1.9. Model Naive Bayes.....	29
3.1.10. Data Uji.....	30
3.1.11. Testing.....	30
3.1.12. Hasil Akurasi.....	30
3.1.13. Evaluasi.....	31
3.2. Perangkat Penelitian.....	31
BAB IV.....	32
HASIL DAN PEMBAHASAN.....	32
4.1. Pengumpulan Data.....	32
4.2. Data.....	32
4.3. Pra-proses Data.....	34
4.4. Data Cleaning.....	38
4.5. Transformasi Data.....	38
4.6. Pembagian Data.....	50
4.7. Pemodelan Naive Bayes.....	51
BAB V.....	76
PENUTUP.....	76
5.1. Kesimpulan.....	76
5.2. Saran.....	76
DAFTAR PUSTAKA.....	77
RIWAYAT HIDUP.....	80
LAMPIRAN.....	81

DAFTAR TABEL

Tabel 4. 1 Variabel beserta tipe data dan nilai atribut.....	32
Tabel 4. 2 Detail Data	35
Tabel 4. 3 Tabel probabilitas atribut Pemilihan Fakultas dan Program Studi berdasarkan keinginan diri sendiri	38
Tabel 4. 4 Tabel probabilitas atribut Program Studi yang dipilih berdasarkan dengan karir yang diinginkan.....	39
Tabel 4. 5 Tabel probabilitas atribut Program Studi yang dipilih sudah sesuai dengan bayangan yang dimiliki sebelumnya.....	39
Tabel 4. 6 Tabel probabilitas atribut Mata kuliah yang ada sesuai dengan ilmu yang ingin dipelajari pada Program Studi yang dipilih	39
Tabel 4. 7 Tabel probabilitas atribut Menyukai mata kuliah yang dipelajari pada Program Studi yang dipilih.....	40
Tabel 4. 8 Tabel probabilitas atribut Ingin mempelajari ilmu diluar dari mata kuliah yang ada	40
Tabel 4. 9 Tabel probabilitas atribut Ilmu yang ingin dipelajari diluar perkuliahan masih berhubungan dengan Program Studi yang dipilih.....	40
Tabel 4. 10 Tabel probabilitas atribut Ilmu dari beberapa bidang studi yang ingin dipelajari 41	
Tabel 4. 11 Tabel probabilitas atribut Memiliki waktu lebih untuk belajar diluar dari jam perkuliahan.....	41
Tabel 4. 12 Tabel probabilitas atribut Mengetahui informasi terkait dengan Kelompok Studi Mahasiswa.....	41
Tabel 4. 13 Tabel probabilitas atribut Ingin mengikuti organisasi yang ada di fakultas	42
Tabel 4. 14 Tabel probabilitas atribut Memilih untuk bergabung ke dalam organisasi yang ada	42
Tabel 4. 15 Tabel probabilitas atribut Tertarik untuk belajar mengenai programming berbasis android	42
Tabel 4. 16 Tabel probabilitas atribut Tertarik untuk belajar mengenai graphic design, video editing, dan semacamnya	42
Tabel 4. 17 Tabel probabilitas atribut Tertarik untuk belajar mengenai keamanan data dalam jaringan dan sebagainya	43
Tabel 4. 18 Tabel probabilitas atribut Tertarik untuk belajar mengenai mekanisme pembuatan robot dan semacamnya.....	43
Tabel 4. 19 Tabel probabilitas atribut Tertarik menjalani karir diluar dari ilmu yang anda pelajari di perkuliahan.....	43
Tabel 4. 20 Tabel probabilitas atribut Tertarik memiliki skill diluar dari perkuliahan yang anda pelajari.....	44
Tabel 4. 21 Tabel probabilitas atribut Memilih beberapa skill yang ada.....	44
Tabel 4. 22 Tabel probabilitas atribut Mampu dalam menyelesaikan mata kuliah yang dipelajari	44
Tabel 4. 23 Tabel probabilitas atribut Mampu dalam mengimplementasikan mata kuliah yang dipelajari	45



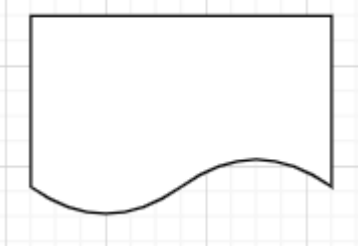
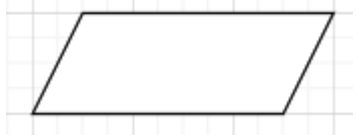
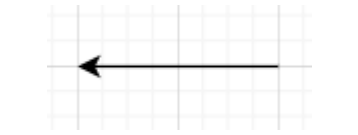
Tabel 4. 24 Tabel probabilitas atribut Merasa sudah puas dengan mata kuliah yang dipelajari	45
Tabel 4. 25 Tabel probabilitas atribut Mempelajari hal selain di perkuliahan dapat mengganggu perkuliahan	45
Tabel 4. 26 Tabel probabilitas atribut Kemampuan yang dimiliki tidak sesuai dengan perkuliahan yang dijalani	46
Tabel 4. 27 Tabel probabilitas atribut Memiliki kemampuan dalam melakukan pemrograman berbasis android	46
Tabel 4. 28 Tabel probabilitas atribut Memiliki kemampuan dalam melakukan software editing dan sejenisnya	46
Tabel 4. 29 Tabel probabilitas atribut Memiliki kemampuan dalam membuat sebuah robot yang berbasis komputer	47
Tabel 4. 30 Tabel probabilitas atribut Memiliki kemampuan dalam melakukan peretasan terhadap suatu jaringan dan sejenisnya.....	47
Tabel 4. 31 Tabel probabilitas atribut Karir yang dipilih	47
Tabel 4. 32 Tabel probabilitas atribut Merasa mampu dalam mempelajari ilmu yang ada di perkuliahan dan juga yang ada di dalam Kelompok Studi Mahasiswa	48
Tabel 4. 33 Tabel probabilitas atribut Mumpuni untuk bergabung ke dalam Kelompok Studi Mahasiswa.....	48
Tabel 4. 34 Data sebelum di transformasi	49
Tabel 4. 35 Data setelah di transformasi.....	50
Tabel 4. 36 Data untuk training	52
Tabel 4. 37 Tabel perhitungan probabilitas atribut Pemilihan Fakultas dan Program Studi berdasarkan keinginan diri sendiri	53
Tabel 4. 38 Tabel perhitungan probabilitas atribut Program Studi yang dipilih berdasarkan dengan karir yang diinginkan	54
Tabel 4. 39 Tabel perhitungan probabilitas atribut Program Studi yang dipilih sudah sesuai dengan bayangan yang dimiliki sebelumnya	54
Tabel 4. 40 Tabel perhitungan probabilitas atribut Mata kuliah yang ada sesuai dengan ilmu yang ingin dipelajari pada Program Studi yang dipilih	54
Tabel 4. 41 Tabel perhitungan probabilitas atribut Menyukai mata kuliah yang dipelajari pada Program Studi yang dipilih	54
Tabel 4. 42 Tabel perhitungan probabilitas atribut Ingin mempelajari ilmu diluar dari mata kuliah yang ada	55
Tabel 4. 43 Tabel perhitungan probabilitas atribut Ilmu yang ingin dipelajari diluar perkuliahan masih berhubungan dengan Program Studi yang dipilih	55
Tabel 4. 44 Tabel perhitungan probabilitas atribut Ilmu dari beberapa bidang studi yang ingin dipelajari	55
Tabel 4. 45 Tabel perhitungan probabilitas atribut Memiliki waktu lebih untuk belajar diluar dari jam perkuliahan	56
Tabel 4. 46 Tabel perhitungan probabilitas atribut Mengetahui informasi terkait dengan Kelompok Studi Mahasiswa	56
Tabel 4. 47 Tabel perhitungan probabilitas atribut Ingin mengikuti organisasi yang ada di fakultas.....	56

Tabel 4. 48 Tabel perhitungan probabilitas atribut Memilih untuk bergabung ke dalam organisasi yang ada.....	56
Tabel 4. 49 Tabel perhitungan probabilitas atribut Tertarik untuk belajar mengenai programming berbasis android	57
Tabel 4. 50 Tabel perhitungan probabilitas atribut Tertarik untuk belajar mengenai graphic design, video editing, dan semacamnya.....	57
Tabel 4. 51 Tabel perhitungan probabilitas atribut Tertarik untuk belajar mengenai keamanan data dalam jaringan dan sebagainya	57
Tabel 4. 52 Tabel perhitungan probabilitas atribut Tertarik untuk belajar mengenai mekanisme pembuatan robot dan semacamnya.....	57
Tabel 4. 53 Tabel perhitungan probabilitas atribut Tertarik menjalani karir diluar dari ilmu yang anda pelajari di perkuliahan	58
Tabel 4. 54 Tabel perhitungan probabilitas atribut Tertarik memiliki skill diluar dari perkuliahan yang anda pelajari	58
Tabel 4. 55 Tabel perhitungan probabilitas atribut Memilih beberapa skill yang ada	58
Tabel 4. 56 Tabel perhitungan probabilitas atribut Mampu dalam menyelesaikan mata kuliah yang dipelajari.....	59
Tabel 4. 57 Tabel perhitungan probabilitas atribut Mampu dalam mengimplementasikan mata kuliah yang dipelajari.....	59
Tabel 4. 58 Tabel perhitungan probabilitas atribut Merasa sudah puas dengan mata kuliah yang dipelajari.....	59
Tabel 4. 59 Tabel perhitungan probabilitas atribut Mempelajari hal selain di perkuliahan dapat mengganggu perkuliahan	60
Tabel 4. 60 Tabel perhitungan probabilitas atribut Kemampuan yang dimiliki tidak sesuai dengan perkuliahan yang dijalani	60
Tabel 4. 61 Tabel perhitungan probabilitas atribut Memiliki kemampuan dalam melakukan pemrograman berbasis android.....	60
Tabel 4. 62 Tabel perhitungan probabilitas atribut Memiliki kemampuan dalam melakukan software editing dan sejenisnya	60
Tabel 4. 63 Tabel perhitungan probabilitas atribut Memiliki kemampuan dalam membuat sebuah robot yang berbasis komputer.....	61
Tabel 4. 64 Tabel perhitungan probabilitas atribut Memiliki kemampuan dalam melakukan peretasan terhadap suatu jaringan dan sejenisnya.....	61
Tabel 4. 65 Tabel perhitungan probabilitas atribut Karir yang dipilih	61
Tabel 4. 66 Tabel perhitungan probabilitas atribut Merasa mampu dalam mempelajari ilmu yang ada di perkuliahan dan juga yang ada di dalam Kelompok Studi Mahasiswa	62
Tabel 4. 67 Tabel perhitungan probabilitas atribut Mumpuni untuk bergabung ke dalam Kelompok Studi Mahasiswa	62

DAFTAR GAMBAR

Gambar 2. 1 Data mining sebagai dari proses knowledge discovery	15
Gambar 2. 2 <i>Pemrograman Tradisional dan Machine Learning memiliki perbedaan yang signifikan. Berikut adalah perbandingan antara keduanya (Brownlee, 2015).</i>	17
Gambar 3. 1 <i>Kerangka Pikir</i>	25

DAFTAR SIMBOL

No.	Gambar	Nama	Keterangan
1		Terminal (Start dan End)	Menampilkan permulaan dan kegiatan berakhir
2		Process	Menampilkan sebuah proses yang dilakukan
3		Activity	Menampilkan kegiatan yang dilakukan terhadap suatu proses
4		Data	Menampilkan keluaran data ataupun inputan data
5		Flow	Menjelaskan arah proses berjalan

DAFTAR LAMPIRAN

LAMPIRAN 1 : Hasil Turnitin	81
LAMPIRAN 2 : Hasil Rangkuman Kuisisioner	82
LAMPIRAN 3 : Data sebelum di transformasi.....	90
LAMPIRAN 4 : Data setelah di Transformasi	93
LAMPIRAN 5 : Source Code	95