

**IMPLEMENTASI KOMBINASI ALGORITMA *NAÏVE BAYES* DAN
ALGORITMA *ROUGH SET* UNTUK MEMPREDIKSI INDEKS PRESTASI
MAHASISWA**

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar
Sarjana Komputer

DISUSUN OLEH:

MUHAMMAD FEBRI MAULANA

1811102441059



**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR**

2023

**Implementasi Kombinasi Algoritma *Naïve Bayes* dan Algoritma
Rough Set untuk Memprediksi Indeks Prestasi Mahasiswa**

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar
Sarjana Komputer

Disusun Oleh:

Muhammad Febri Maulana

1811102441059



PROGRAM STUDI S1 TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR

2023

HALAMAN PENGESAHAN

IMPLEMENTASI KOMBINASI ALGORITMA *NAÏVE BAYES* DAN ALGORITMA *ROUGH SET* UNTUK MEMPREDIKSI INDEKS PRESTASI MAHASISWA

DISUSUN OLEH:

MUHAMMAD FEBRI MAULANA

1811102441059

Telah melaksanakan ujian skripsi dan dinyatakan lulus,

Pada tanggal 6 Januari 2023

Dosen Pembimbing



Rofle Hasudungan, S.Kom., M.Sc
NIDN. 1107048601

Penguji



Wawan Joko Pranoto, S.Kom., M.TI
NIDN. 1102057701

Dekan



Prof. Dr. Sarjito, MT., Ph.D
NIDN. 0610116204

Ketua Program Studi



Asslia Johar Latipah, S.Kom., M.Cs
NIDN. 1124098902

PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Febri Maulana
NIM : 1811102441059
Konsentrasi : Sistem Informasi

Menyatakan bahwa Skripsi dengan judul berikut:

Implementasi Kombinasi Algoritma *Naïve Bayes* dan Algoritma *Rough Set* untuk Memprediksi Indeks Prestasi Mahasiswa

Dosen Pembimbing : Rofie Hasudungan, S.Kom., M.Sc

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar Sarjana Komputer, baik di Universitas Muhammadiyah Kalimantan Timur (UMKT) maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas Muhammadiyah Kalimantan Timur (UMKT)
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi Muhammadiyah.

Samarinda, 6 Januari 2023



(Muhammad Febri Maulana)

PRAKATA

Segala puji dan syukur atas kehadiran Allah SWT karena Rahmat dan Hidayah-Nya peneliti dapat menyelesaikan skripsi yang berjudul “Implementasi Kombinasi Algoritma Naïve Bayes dan Algoritma Rough Set untuk Memprediksi Indeks Prestasi Mahasiswa”. Peneliti menyadari bahwa dalam penyelesaiannya tidak dapat terwujud tanpa dukungan dan bantuan dari berbagai pihak. Oleh sebab itu peneliti mengucapkan rasa terima kasih kepada:

1. Kedua orang tua Ibu Sumiati dan Bapak Muhammad Jubri yang senantiasa mendoakan dan memotivasi peneliti selama ini.
 2. Bapak Rofile Hasudungan, S.Kom., M.Sc selaku Dosen pembimbing yang telah memberikan banyak kesempatan bimbingan dan masukkan hingga skripsi selesai kepada peneliti.
 3. Ibu Asslia Johar Latipah, M.Cs., selaku Ketua program studi teknik Informatika Universitas Muhammadiyah Kalimantan Timur.
 4. Bapak Prof. Dr. H. Bambang Setiaji, selaku rektor Universitas muhammadiyah kalimantan timur.
 5. Bapak Prof. Ir. Sarjito, M.T.,Ph.D., selaku dekan fakultas sains dan teknologi Universitas muhammadiyah kalimantan timur.
 6. Bapak dan Ibu dosen Program Studi Teknik Informatika yang telah memberikan ilmu pengetahuan selama penulis menempuh pendidikan di Universitas muhammadiyah kalimantan timur.
 7. Teman – teman satu bimbingan peneliti Aulia rachman, Doni pratama, Hary subekti, Ilham Qasthalani, M. Fahri Al Sya’ba, Muhammad Iqbal Wahyu Perdana, dan Taufiqurrahman yang membantu dan memberikan masukkan
- Peneliti menyadari bahwa dalam penulisan skripsi masih berada jauh dalam kata sempurna, karena keterbatasan ilmu yang dimiliki oleh penulis. Oleh karenanya peneliti mengharapkan saran dan kritik bersifat membangun dari semua pihak.

ABSTRAK

Indeks prestasi mahasiswa yang berada dibawah standar dapat menyebabkan berbagai masalah yang menimbulkan efek berantai menyebabkan turunnya performa rendah di semester-semester yang akan datang. Untuk menjaga prestasi mahasiswa tetap baik maka, diperlukan adanya sebuah pendekatan untuk menggali informasi penting mengenai faktor yang melatar belakangi mahasiswa. Salah satu cara untuk menggali informasi adalah melakukan prediksi menerapkan analisis data menggunakan data mining. Pada analisis data penelitian ini menggunakan data berjumlah 306 data mahasiswa yang terdiri dari 240 mahasiswa memiliki indeks prestasi dan 66 mahasiswa belum memiliki indeks prestasi. Menggunakan metode algoritma rough set sebagai metode pemilihan atribut untuk klasifikasi menghasilkan 4 atribut pilihan dari 14. Untuk mengimplementasi dalam klasifikasi akan membagi data menjadi 70% data latih dan 30% data uji pada model menggunakan semua atribut dan model menggunakan pemilihan atribut. Dalam implementasi menggunakan dua model yaitu model menggunakan semua atribut menghasilkan akurasi sebesar 86,76%. Sedangkan akurasi model menggunakan pemilihan atribut menghasilkan meningkat 89,71%. Sehingga dapat dilakukan prediksi pada data mahasiswa yang belum memiliki indeks prestasi mahasiswa menghasilkan 47 indeks prestasi tinggi, 4 indeks prestasi cukup, dan 7 indeks prestasi rendah.

Kata kunci : naive bayes, rough set, indeks prestasi

ABSTRACT

Grade Point Average (GPA) that are below standard can cause various problems that cause a chain effect that causes a decrease in low GPA in the coming semesters. To maintain good student learning next GPA, it is necessary to have an approach to extract important information about the factors behind students. One way to dig up information is to make predictions by applying data analysis using data mining. In data analysis, this research used data that totaled 306 student data consisting of 240 students who had a GPA and 66 students who did not yet have a GPA. The use of the rough set algorithm method as an attribute selection method for classification produces 4 attribute choices out of 14 attributes. To implement the classification, divide the data into 70% data training and 30% data testing on the model using all attributes and the model using attribute selection. In its implementation, it uses two models, namely a model that uses all the attributes to produce an accuracy of 83,58%. Meanwhile, the accuracy of the model using attribute selection resulted in an increase of 88,06%. So that predictions can be made on student data that does not yet have a GPA which results in a high GPA 47 students, a moderate GPA 4 students, and a low GPA 7 students.

Keyword : *naive bayes, rough set, GPA*

DAFTAR ISI

ABSTRAK.....	vi
<i>ABSTRACT</i>	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
Daftar Lampiran	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Batasan Penelitian	3
BAB II TINJAUAN PUSTAKA.....	4
2.1 Prediksi indeks prestasi mahasiswa	4
2.2 Algoritma Naïve Bayes.....	6
2.3 Algoritma Rough Set.....	10
2.4 Maximum Dependency of Attributes.....	13
2.5 Pemrosesan data.....	15
2.6 Evaluasi.....	16
BAB III METODE PENELITIAN	18
3.1 Desain Penelitian.....	18
3.2 Koleksi Data	19
3.3 Pemroses Data	20
3.4 Pemilihan atribut dengan Rough set.....	21
3.5 Klasifikasi dengan Naïve Bayes.....	22
3.6 Evaluasi dan Komparasi.....	22
3.7 Prediksi indeks prestasi mahasiswa	23
BAB IV PEMBAHASAN DAN HASIL.....	24
4.1 Data penelitian	24
4.2 Pemrosesan data.....	25

4.3	Integrasi data.....	25
4.3.1	Pembersihan data.....	27
4.3.2	Transformasi data.....	27
4.4	Pemilihan atribut dengan algoritma Rough set	28
4.5	Klasifikasi dengan algoritma Naïve bayes	29
4.5.1	Model klasifikasi dengan semua atribut	31
4.5.2	Model klasifikasi dengan pemilihan atribut	33
4.6	Evaluasi dan Komparasi.....	34
4.7	Prediksi indeks prestasi mahasiswa	35
BAB V PENUTUP		38
5.1	Kesimpulan	38
5.2	Saran.....	39
DAFTAR PUSTAKA.....		40
Lampiran		45

DAFTAR TABEL

Tabel 2.1 Penelitian terkait indeks prestasi mahasiswa	4
Tabel 2.2 Tabel penelitian terkait metode naïve bayes.....	7
Tabel 2.3 Tabel indeks prestasi mahasiswa menggunakan metode naive bayes...	9
Tabel 2.4 Tabel penelitian terkait Rough set	12
Tabel 2.5 Confusion matrix	16
Tabel 3.1 Tabel atribut yang digunakan dalam penelitian.....	19
Tabel 3.2 Tabel kategori indeks prestasi mahasiswa	21
Tabel 3.3 Tabel kategori jumlah anggota keluarga.....	21
Tabel 4.1 Contoh data yang diperoleh dari BAA.....	25
Tabel 4.2 Contoh data mahasiswa informatika tahun angkatan 2020 dan 2021 .	26
Tabel 4.3 Contoh data mahasiswa informatika tahun angkatan 2022	26
Tabel 4.4 Hasil Data Cleaning.....	27
Tabel 4.5 Contoh data mahasiswa angkatan 2020 dan 2021 yang telah di transformasi	27
Tabel 4.6 Contoh data mahasiswa angkatan 2022 yang telah di transformasi....	28
Tabel 4.7 Hasil dari perhitungan ketergantungan atribut MDA	29
Tabel 4.8 Perhitungan atribut Status tempat tinggal mahasiswa.....	31
Tabel 4.9 Nilai probabilitas atribut Status tempat tinggal mahasiswa	31
Tabel 4.10 Hasil prediksi klasifikasi dengan menggunakan semua atribut	32
Tabel 4.11 Hasil prediksi klasifikasi dengan menggunakan pemilihan atribut	33
Tabel 4.12 Confusion matrix model pertama	34
Tabel 4.13 Confusion matrix model kedua	34
Tabel 4.14 Tabel prediksi indeks prestasi mahasiswa 2022	36

DAFTAR GAMBAR

Gambar 2.1 Skema penyelesaian algoritma rough set	10
Gambar 2.2 Skema penyelesaian MDA.....	14
Gambar 3.1 Gambar alur penelitian	18
Gambar 4.1 Jumlah data tekumpul dari angkatan 2020 sampai 2022	24
Gambar 4.2 Akurasi model klasifikasi	35
Gambar 4.3 Hasil prediksi angkatan 2022	37

DAFTAR LAMPIRAN

- Lampiran 1.** Kuesioner Google Form yang dibagikan untuk mengambil data.
- Lampiran 2.** Data penelitian mahasiswa angkatan 2020 dan 2021
- Lampiran 3.** Data penelitian mahasiswa angkatan 2022
- Lampiran 4.** Data berasal dari BAA
- Lampiran 5.** Source code implementasi rough set library rst-tools
- Lampiran 6.** Implementasi klasifikasi naïve bayes menggunakan RapidMiner
- Lampiran 7.** Hasil Prediksi indeks prestasi mahasiswa teknik informatika angkatan 2022
- Lampiran 8.** Lembar revisi skripsi
- Lampiran 9.** Lembar Bimbingan Skripsi
- Lampiran 10.** Kuesioner Google Form yang dibagikan untuk mengambil data
- Lampiran 11.** Surat Keterangan Tidak Uji Validasi
- Lampiran 12.** Uji Plagiarisme