

***ANALISIS OPTIMASI FORWARD SELECTION PADA
KLASIFIKASI NILAI MAHASISWA DENGAN ALGORITMA
NAÏVE BAYES CLASSIFER***

STUDI KASUS : UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar
Sarjana Komputer

DISUSUN OLEH :

RAHMAT RAMADHANI

1911102441022



**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
2023**

Analisis Optimasi Forward Selection pada Klasifikasi Nilai

Mahasiswa dengan Algoritma Naïve Bayes Classifier

Studi Kasus : Universitas Muhammadiyah Kalimantan Timur

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar

Sarjana Komputer

Disusun Oleh :

Rahmat Ramadhani

1911102441022



PROGRAM STUDI S1 TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR

2023

HALAMAN PENGESAHAN

***Analisis Optimasi Forward Selection Pada Klasifikasi Nilai
Mahasiswa dengan Algoritma Naïve Bayes Classifier***
Studi Kasus: Universitas Muhammadiyah Kalimantan Timur

DISUSUN OLEH:
RAHMAT RAMADHANI
1911102441022

Telah melaksanakan ujian skripsi dan dinyatakan lulus,
Samarinda, 07 Juli 2023

Dosen Pembimbing



Wawan Joko Pranoto, S.Kom, M.TI

NIDN. 1102057701


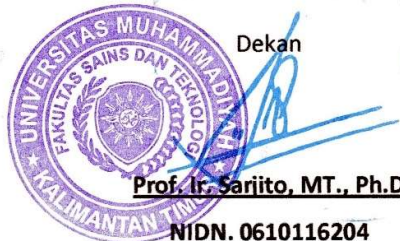
Dosen Penguji



Faldi, S.Kom, M.TI

NIDN. 1121079101

Dekan

Prof. Ir. Sariito, MT., Ph.D
NIDN. 0610116204

Ketua Program Studi




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

SURAT PERNYATAAN KEASLIAN SKRIPSI

Assalamualaikum Warahmatullahi Wabarakatuh

Saya yang bertanda tangan dibawah ini :

Nama : RAHMAT RAMADHANI
NIM : 1911102440122
Program Studi : S1 TEKNIK INFORMATIKA
Konsentrasi : Rekayasa Sistem

Menyatakan bahwa Skripsi dengan judul berikut :

"Analisis Optimasi Forward Selection Pada Klasifikasi Nilai Mahasiswa dengan Algoritma Naïve Bayes Classifier"

Dosen Pembimbing : Wawan Joko Pranoto, S.Kom, M.Ti

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 SANKI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi Muhammadiyah



KATA PENGANTAR



Alhamdulillah puji syukur kepada Allah SWT berkat Rahmat dan Karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Analisis Optimasi Forward Selection pada Klasifikasi Nilai Mahasiswa dengan Algoritma Naïve Bayes Classifier Studi Kasus : Universitas Muhammadiyah Kalimantan Timur”. Adapun tujuan dari penulisan skripsi ini adalah untuk memenuhi salah satu persyaratan penyelesaian studi di jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur.

Penulis menyadari bahwa skripsi ini tidak akan berjalan dengan sebagaimana mestinya tanpa adanya dukungan dari berbagai pihak, oleh karena itu izinkanlah penulis menyampaikan ucapan terima kasih dan penghargaan yang sebesar-besarnya kepada :

1. Ibunda saya yakni Ibu Dahlia yang saya cintai dan yang telah memberikan kasih sayang, tenaga, serta hidupnya untuk memperjuangan dan selalu mendoakan saya hingga pada titik ini.
2. Adik saya yang sangat saya banggakan Nur Fadhillah yang tidak pernah lupa memberikan waktunya untuk menyemangati saya dalam kondisi apapun dan Ryan Rizky Liansyah, S.T orang yang menjadi Kakak, Senior dan juga sahabat saya yang menemani dalam suka maupun duka.
3. Bapak Wawan Joko Parnoto, S. Kom., M,Kom selaku dosen pembimbing yang telah meluangkan waktunya, membimbing dan memberikan saran serta motivasi kepada penulis.
4. Bapak Faldi, S.Kom, M.Ti selaku dosen penguji yang telah meluangkan waktunya, membimbing dan memberikan saran serta motivasi kepada penulis.
5. Ibu Asslia Johar Latipah, M.Cs selaku ketua Program Studi S1 Teknik Informatika

6. Bapak Prof. Ir. Sarjito, M.T.,Ph.D., selaku Dekan Fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur.
7. Bapak Prof. Dr. H. Bambang Setiaji, selaku Rektor Universitas Muhammadiyah Kalimantan Timur.
8. Bapak Taghfirul Yoga Siswa, S.Kom. M.Ti selaku salah satu dosen yang membidangi Data Mining.
9. Seluruh dosen S1 Teknik Informatika Universitas Muhammadiyah Kalimantan Timur yang telah memberikan ilmunya kepada penulis.
10. Kepada rekan kerja saya yaitu Ustadz Kahar, S.Pd, Ustadz Alauddin Dzulfahmi, S.Pd, M.Ap, Ustadz Hadiyatna Hamid, Ustadz Safaruddin Hanif, Ustadz Khoirul Anam , Ustadz Sigit Biantoro, S.Pd, Ustadz Muhammad Azizi & Ustadz Wildan Hamdani yang selalu mengingatkan, memotivasi serta mensupport saya dalam segala
11. Kepada rekan-rekan seperjuangan bimbingan yaitu Putra Saymen, Khamirul Hakim, Muhammad Noor Fiqri, Dinamita Romadoni, Rahmawati, Muhammad Fath Thoriq, Suci Mawaddah, I Nyoman Purne Wijaya dan Muhammad Zulfikar Aulia Nugraha yang telah banyak membantu dan berjuang bersama-sama serta memberikan saran serta semangat sehingga dapat menyelesaikan proposal ini.
12. Kepada teman-teman S1 Teknik Informatika angkatan 2019.
13. Semua pihak yang tidak dapat penulis sebutkan satu persatu, terima kasih atas bantuan, nasehat dan dorongan dalam proses penyusunan skripsi ini. Penulis menyadari skripsi ini masih jauh dari sempurna, oleh karena itu dengan segala hormat dan kerendahan hati penulis berharap saran dan kritik yang membangun untuk hasil yang dituangkan dalam skripsi ini bisa bermanfaat bagi pihak-pihak yang memerlukan.

Samarinda, 08 Maret 2023



Rahmat Ramadhani

ABSTRAK

Nilai akademik mahasiswa Universitas Muhammadiyah Kalimantan Timur tahun ajaran 2020/2021 terjadi penurunan signifikan pada Mata Kuliah Bahasa Arab. Hal ini dipandang perlu untuk dilakukan analisis data dengan proses Data Mining yakni klasifikasi sebagai evaluasi untuk menentukan apa yang menjadi parameter penurunan dari mata kuliah tersebut. Tujuan penelitian ini akan mengimplementasikan Algoritma Naïve Bayes Classifier dengan Seleksi Fitur Forward Selection untuk mendapatkan nilai akurasi dan atribut yang paling berpengaruh dalam Mata Kuliah Bahasa Arab. Jumlah data yang digunakan adalah 943 record dengan menerapkan pembagian data menggunakan K – Fold Cross Validation. Hasil pengujian evaluasi rata rata fold yakni sebesar 86% tanpa menggunakan Seleksi Fitur dan Akurasi meningkat 7% yakni menjadi 93% setelah menggunakan Optimasi Forward Selection.

Kata kunci: *Naïve Bayes; Forward Selection; Klasifikasi, Cross Validation, Akurasi*

ABSTRACT

The academic scores of Muhammadiyah University students in East Kalimantan for the 2020/2021 academic year saw a significant decrease in the Arabic Language Course. It is deemed necessary to carry out data analysis with the Data Mining process, namely classification as an evaluation to determine what is the parameter for the decline of the course. The purpose of this research is to implement the Naïve Bayes Classifier Algorithm with Forward Selection Feature Selection to get the most influential accuracy and attribute values in the Arabic Language Course. The amount of data used is 943 records by applying data division using K - Fold Cross Validation. The results of the evaluation evaluation test show that the average fold is 86% without using Feature Selection and the accuracy increases by 7% to 93% after using Forward Selection Optimization.

Keywords: *Naïve Bayes; Forward Selection; Classification, Cross Validation, Accuracy*

DAFTAR ISI

HALAMAN PENGESAHAN	i
SURAT PERNYATAAN KEASLIAN SKRIPSI	ii
KATA PENGANTAR	iii
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
BAB 1	1
PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan Penelitian	3
1.4. Batasan Masalah	3
BAB 2	4
TINJAUAN PUSTAKA	4
2.1 Kinerja Mahasiswa	4
2.2 Data Mining	4
2.2.1. Metode Klasifikasi	5
2.2.2. Algoritma Naïve Bayes	6
2.2.3. Data Preprocessing	7
2.2.4. Metode Cross-Validation	8
2.2.5. Confusion Matrix	9
2.2.6. CRISP-DM (Cross-Industry Standard Process for Data Mining)	11
2.2.7. Optimasi Forward Selection	12
2.3 Python.	13
2.4 Scikit – Learn	13

2.5	Penelitian Terdahulu	14
BAB 3	19
METODOLOGI PENELITIAN	19
3.1.	Objek Penelitian	19
3.2.	Teknik Pengumpulan Data.....	19
3.3.	Teknik Analisis Data.....	19
3.4.	Peralatan yang digunakan	25
3.5.	Jadwal Penelitian.....	25
BAB 4	26
HASIL & PEMBAHASAN	26
4.1.	Pemahaman Data	26
4.2.	Pengumpulan Data	26
4.3.	Data Preprocessing	28
4.3.1.	Data Selection	28
4.3.2.	Data Integration	29
4.3.3.	Data Transformation	29
4.3.4.	K – Fold Cross Validation.....	30
4.4.	Modeling dan Evaluation.....	34
4.4.1.	Perhitungan Algoritma Naïve Bayes.....	34
4.4.2.	Skema Seleksi Fitur Forward Selection.....	45
4.4.3.	Implementasi Naïve Bayes pada Python	50
4.4.4.	Implementasi Naïve Bayes dengan Forward Selection pada Python.....	52
4.5.	Pembahasan	53
BAB 5	54
PENUTUP	54
5.1.	Kesimpulan	54
5.2.	Saran.....	54
DAFTAR PUSTAKA	55
LAMPIRAN	59

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	14
Tabel 3. 1 Atribut Data MKDU.....	20
Tabel 3. 2 Atribut Data BAA	21
Tabel 3. 3 Atribut Seleksi dan Integrasi Data	22
Tabel 3. 4 Jadwal Penelitian	25
Tabel 4. 1 Data Atribut BAA	26
Tabel 4. 2 Data Atribut MKDU.....	27
Tabel 4. 3 Data BAA yang telah Diseleksi.....	28
Tabel 4. 4 Data MKDU yang telah diseleksi.....	28
Tabel 4. 5 Data MKDU & BAA yang di Integrasi	29
Tabel 4. 6 Tabel Transformasi Data.....	29
Tabel 4. 7 Tabel Transformasi Data 2.....	30
Tabel 4. 8 Tabel Data yang telah di integrasi dan di transformasi.....	30
Tabel 4. 9 Skema 10 Fold Cross Validation	31
Tabel 4. 10 Skema Teknik Pembagian Data	32
Tabel 4. 11 Tabel Data Testing Fold - 1	32
Tabel 4. 12 Tabel Data Training Fold - 1.....	32
Tabel 4. 13 Tabel Data Testing Fold 2	32
Tabel 4. 14 Tabel Data Training Fold - 2.....	33
Tabel 4. 15 Tabel Data Training Fold - 2 (Lanjutan)	33
Tabel 4. 16 Tabel Data Testing Fold - 3	33
Tabel 4. 17 Tabel Data Training Fold - 3.....	34
Tabel 4. 18 Tabel Data Training Fold - 3 (Lanjutan)	34
Tabel 4. 19 Tabel Confussion Matrix Fold - 1.....	40
Tabel 4. 20 Tabel Confussion Matrix Fold - 2.....	40
Tabel 4. 21 Tabel Confussion Matrix Fold - 3.....	41
Tabel 4. 22 Tabel Confussion Matrix Fold - 4.....	41
Tabel 4. 23 Tabel Confussion Matrix Fold - 5.....	42
Tabel 4. 24 Tabel Confussion Matrix Fold - 6.....	42

Tabel 4. 25 Tabel Confussion Matrix Fold - 7.....	43
Tabel 4. 26 Tabel Confussion Matrix Fold - 8.....	43
Tabel 4. 27 Tabel Confussion Matrix Fold - 9.....	44
Tabel 4. 28 Tabel Confussion Matrix Fold - 10.....	44
Tabel 4. 29 Tabel Akurasi Masing - Masing Fold.....	45
Tabel 4. 30 Tabel Skema Penerapan Forward Selection Iterasi 1.....	46
Tabel 4. 31 Tabel Skema Penerapan Forward Selection Iterasi 1 (Lanjutan)	46
Tabel 4. 32 Tabel Penerapan Forward Selection Iterasi 2	47
Tabel 4. 33 Tabel Penerapan Optimasi Forward Selection Iterasi 2 (Lanjutan)	47
Tabel 4. 34 Tabel Penerapan Forward Selection Iterasi 3	48
Tabel 4. 35 Tabel Penerapan Forward Selection Iterasi 3 (Lanjutan).....	48
Tabel 4. 36 Tabel Penerapan Forward Selection Iterasi 3 (Lanjutan).....	48
Tabel 4. 37 Tabel Penerapan Forward Selection Iterasi 4	49
Tabel 4. 38 Tabel Penerapan Forward Selection Iterasi 4 (Lanjutan).....	49

DAFTAR GAMBAR

Gambar 2. 1 Confusion Matrix.....	9
Gambar 2. 2 CRISP-DM.....	12
Gambar 3. 1 Alur Analisis Data.....	19
Gambar 3. 2 Contoh Diagram Alur Forward Selection	24
Gambar 4. 1 Dataset MKDU & BAA.....	50
Gambar 4. 2 Jumlah Record Kelas BAIK dan BURUK.....	50
Gambar 4. 3 Implementasi Naive Bayes Pada Python.....	51
Gambar 4. 4 Hasil Confussion Matrix Naive Bayes beserta Akurasi	51
Gambar 4. 5 Confussion Matrix K - Fold Cross Validation pada Naive Bayes.....	52
Gambar 4. 6 Implementasi Naive Bayes & Forward Selection Pada Python.....	53
Gambar 4. 7 Hasil Akurasi & Fitur Fitur terbaik dari Forward Selection.....	53

DAFTAR LAMPIRAN

Lampiran 1. Riwayat Hidup	60
Lampiran 2. Data MKDU & BAA Lengkap	61
Lampiran 3. Source Code Algortima & Seleksi Fitur.....	106
Lampiran 4. Surat Izin Penelitian.....	108
Lampiran 5. Surat Keterangan tidak Uji Validitas.....	110
Lampiran 6. Lembar Konsultasi	111
Lampiran 7. Bukti Uji Plagiasi	113