

**ANALISIS KLASIFIKASI PERFORMA MAHASISWA DENGAN
NAÏVE BAYES**

SKRIPSI

Oleh:

SAPTA HADI GUNAWAN

NIM. 192112127

VERNANDA JULIANA GAFAR PUTRI

NIM. 182111060

DWI SYAH PUTRA LUBIS

NIM. 182111566



**PROGRAM STUDI S-1 SISTEM INFORMASI
FAKULTAS INFORMATIKA
UNIVERSITAS MIKROSKIL
MEDAN
2023**

ANALYSIS OF STUDENT PERFORMANCE CLASSIFICATION USING NAÏVE BAYES METHOD

FINAL PROJECT

By:

**SAPTA HADI GUNAWAN
ID NUMBER. 192112127
VERNANDA JULIANA GAFAR PUTRI
ID NUMBER. 182111060
DWI SYAH PUTRA LUBIS
ID NUMBER. 182111566**



**MAJOR OF S-1 INFORMATION SYSTEM
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN
2023**

LEMBARAN PENGESAHAN
ANALISIS KLASIFIKASI PERFORMA MAHASISWA DENGAN
NAÏVE BAYES

PROYEK


Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Sarjana
Program Studi S-1 Sistem Informasi

Oleh:

SAPTA HADI GUNAWAN
NIM. 192112127
VERNANDA JULIANA GAFAR PUTRI
NIM. 182111060
DWI SYAH PUTRA LUBIS
NIM. 182111566

Disetujui Oleh:

Dosen Pembimbing I,


Sudarto, S.Kom., M.Kom.

Dosen Pembimbing II,


Joosten, S.Kom., M.Eng.

Medan, 13 Juli 2023
Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-1 Sistem Informasi,


UNIVERSITAS
MIKROSKIL
Yuni Marlina Saragih, S.Kom., M.Kom

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Sistem Informasi Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Sapta Hadi Gunawan
NIM : 192112127

Saya telah melaksanakan penelitian dan penulisan Tugas Akhir dengan judul dan tempat penelitian sebagai berikut:

Judul Tugas Akhir : Analisis Klasifikasi Performa Mahasiswa Dengan Naïve Bayes
Tempat Penelitian : -
Alamat Tempat Penelitian : -
No. Telp. Tempat Penelitian : -

Sehubungan dengan Tugas Akhir tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Tugas Akhir tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Tugas Akhir saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tugas Akhir saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tugas Akhir saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 13 Juli 2023

Saya yang membuat pernyataan,



Sapta Hadi Gunawan

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Sistem Informasi Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Vernanda Juliana Gafar Putri

NIM : 182111060

Saya telah melaksanakan penelitian dan penulisan Tugas Akhir dengan judul dan tempat penelitian sebagai berikut:

Judul Tugas Akhir : Analisis Klasifikasi Performa Mahasiswa Dengan Naïve Bayes

Tempat Penelitian : -

Alamat Tempat Penelitian : -

No. Telp. Tempat Penelitian : -

Shubungan dengan Tugas Akhir tersebut, dengan ini saya menyatakan dengan sebenarnya bahwa penelitian dan penulisan Tugas Akhir tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Tugas Akhir saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tugas Akhir saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tugas Akhir saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 13 Juli 2023

Saya yang membuat pernyataan,



Vernanda Juliana Gafar Putri

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Sistem Informasi Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Dwi Syah Putra Lubis
NIM : 182111566

Saya telah melaksanakan penelitian dan penulisan Tugas Akhir dengan judul dan tempat penelitian sebagai berikut:

Judul Tugas Akhir : Analisis Klasifikasi Performa Mahasiswa Dengan Naïve Bayes
Tempat Penelitian : -
Alamat Tempat Penelitian : -
No. Telp. Tempat Penelitian : -

Sehubungan dengan Tugas Akhir tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Tugas Akhir tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Tugas Akhir saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tugas Akhir saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tugas Akhir saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

13 Juli 2023

Saya yang membuat pernyataan,



Dwi Syah Putra Lubis

ANALISIS KLASIFIKASI PERFORMA MAHASISWA DENGAN NAÏVE BAYES

Abstrak

Prestasi pendidikan saat ini sangat ditentukan dengan performa belajar generasi muda. Untuk meningkatkan performa belajar mahasiswa ada beberapa faktor yang menjadi pengaruhnya. Hal terpenting yang perlu diperhatikan saat proses pembelajaran berlangsung adalah peningkatan performa mahasiswa. Naïve bayes adalah sebuah algoritma *supervised learning* berdasarkan teorema Bayes yang digunakan untuk memecahkan masalah klasifikasi dengan mengikuti pendekatan probabilistik. Hasil capaian tersebut tentu saja berpengaruh pada ketepatan mahasiswa dalam menyelesaikan waktu studinya. Eksperimen ini juga menguji *accuracy*, *precision*, *recall*, dan *f1-score* menggunakan *k-fold cross validation*. Nilai *accuracy* tertinggi didapat dari hasil pengujian *10-fold* data 60% dari total atribut dengan nilai 89,31%. Nilai *precision* tertinggi didapat dari hasil pengujian *10-fold* data 60% dari total atribut dengan nilai 92,20%. Nilai *recall* tertinggi didapat dari hasil pengujian *10-fold* data 60% dan *10-fold* data 80% dari total atribut dengan nilai 92,20%. Dan nilai *f1-score* tertinggi didapat dari hasil pengujian *10-fold* data 60% dari total atribut dengan nilai 92,18%. Sedangkan pengukuran dengan menggunakan ROC curve, berdasarkan hasil AUC disetiap perbandingan yang dilakukan bahwa hasil pengujian dengan 60% dari total atribut memiliki nilai tertinggi dari pengujian lainnya.

Kata Kunci: *Algoritma Naïve Bayes, k-fold validation, klasifikasi*

Abstract

Current educational achievement is largely determined by the learning performance of the younger generation. To improve student learning performance there are several factors that influence it. The most important thing that needs to be considered when the learning process takes place is improving student performance. Naïve Bayes is a supervised learning algorithm based on Bayes' theorem that is used to solve classification problems by following a probabilistic approach. The results of these achievements certainly affect the accuracy of students in completing their study time. This experiment also tests accuracy, precision, recall, and f1-score using k-fold cross validation. The highest accuracy value is obtained from the 10-fold data test results of 60% of the total attributes with a value of 89.31%. The highest precision value is obtained from the results of the 10-fold data test with 60% of the total attributes with a value of 92.20%. The highest recall value is obtained from the results of testing 10-fold data 60% and 10-fold data 80% of the total attributes with a value of 92.20%. And the highest f1-score value is obtained from the results of the 10-fold data test of 60% of the total attributes with a value of 92.18%. While measurements using the ROC curve, based on the results of the AUC in each comparison, the results of the test with 60% of the total attributes have the highest value from other tests.

Keyword: *Naïve Bayes algorithm, k-fold cross validation, classification*

KATA PENGANTAR

Ucapan syukur kami panjatkan kehadiran Tuhan Yang Maha Kuasa berkat rahmat dan karunia-Nya, kami dapat menyelesaikan proyek Tugas Akhir yang berjudul “**ANALISIS KLASIFIKASI PERFORMA MAHASISWA DENGAN NAÏVE BAYES**”

Penyusunan Tugas Akhir ini digunakan untuk memenuhi persyaratan dari akademik dalam menyelesaikan Program Studi S-1 Sistem Informasi Fakultas Informatika Universitas Mikroskil Medan.

Dalam penyusunan Tugas Akhir ini kami menyadari banyak rintangan dan hambatan yang terjadi namun dapat dilalui dengan adanya dukungan, nasihat, bimbingan, dan bantuan dari berbagai pihak. Dalam kesempatan ini kami ingin mengucapkan terimakasih kepada:

1. Bapak Sudarto, S.Kom., M.Kom., selaku Dosen Pembimbing I yang telah banyak melakukan bimbingan, pengarahan serta memberikan saran kepada penulis dalam menyelesaikan laporan Tugas Akhir.
2. Bapak Joosten, S.Kom., M.Eng., selaku Dosen Pembimbing II yang telah banyak bersedia meluangkan waktunya untuk memberikan bimbingan, pengarahan serta memberikan saran kepada penulis dalam menyelesaikan laporan Tugas Akhir.
3. Bapak Hardy, S.Kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil Medan.
4. Bapak Ng Poi Wong, S.Kom., M.T.I., selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
5. Ibu Yuni Marlina Saragih S.Kom., M.Kom., selaku Ketua Program Studi S-1 Sistem Informasi Fakultas Informatika Universitas Mikroskil Medan.
6. Seluruh dosen pengajar Universitas Mikroskil yang telah memberikan pengajaran serta ilmu yang berharga kepada penulis selama duduk di bangku kuliah.
7. Kepada orang tua terkasih dari Dwi Syah Putra Lubis yaitu Ibu Tiarlina Sitorus, serta adik kandung Fitri Azura Lubis yang menjadi orang-orang hebat yang tidak pernah lelah memberikan doa dan dukungan, semangat, dan kasih sayang serta motivasi dalam menyelesaikan perkuliahan dan menyelesaikan tugas akhir.
8. Kepada orang tua tercinta dari Vernanda Juliana Gafar Putri yaitu Bapak Johannes Gafar dan Ibu Yenny Oktarina Siregar, serta kakak kandung Vivien Tamara Gafar dan adik kandung Vanessa Gafar yang menjadi orang-orang hebat yang tak pernah lelah memberikan do'a dan dukungan, semangat, dan kasih sayang serta motivasi dalam menyelesaikan perkuliahan dan menyelesaikan tugas akhir

9. Kepada orang tua tercinta dari Sapta Hadi Gunawan yaitu Bapak Suyanto dan Ibu Sri Wahyuni serta yang menjadi orang-orang hebat yang tak pernah lelah memberikan do'a dan dukungan, semangat, dan kasih sayang serta motivasi dalam menyelesaikan perkuliahan dan menyelesaikan tugas akhir.
10. Kepada Suzy Zulayka, Irfan Nurdiansyah Nasution, Dini Sartika Br.Tarigan yang telah memberi motivasi spesial dan kasih sayang dalam menyelesaikan tugas akhir sampai selesai.
11. Sahabat kami selaku teman seperjuangan yang telah memberikan dukungan, bantuan serta motivasi kepada kami menempuh studi dan menyusun tugas akhir di Universitas Mikroskil Medan.
12. Teman satu bimbingan proyek 2023 serta pihak lain yang tidak dapat disebutkan penulis satu-persatu yang telah mendukung dalam penulisan tugas akhir ini.

Akhir kata kami menyadari bahwa tugas akhir ini masih jauh dari kata sempurna. Oleh karena itu kritik dan saran pembaca sangat diharapkan untuk penulisan lebih baik dimasa yang akan datang. Semoga hasil tugas akhir ini dapat memberikan manfaat bagi pengembangan ilmu komunikasi terkhususnya Sistem Informasi dan bermanfaat untuk berbagai macam kepentingan.

Medan, 13 Juli 2023

Penulis,

UNIVERSITAS
MIKROSKIL

Sapta Hadi Gunawan,
Vernanda Juliana Gafar Putri,
Dwi Syah Putra Lubis

DAFTAR ISI

ABSTRAK	iv
KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan	3
1.4. Manfaat	3
1.5. Ruang Lingkup	3
BAB II KAJIAN LITERATUR	4
2.1. Data Mining	4
2.1.1. Tahap-Tahap Data Mining	4
2.1.2. Fungsi Data Mining	8
2.1.3. Faktor-Faktor Data Mining	9
2.1.4. Komponen Data Mining	10
2.2. Klasifikasi	11
2.2.1. Algoritma Naïve Bayes	12
2.3. Evaluasi	15
2.3.1. K-Fold Cross Validation	16
2.3.2. Confusion Matrix	16
2.4. Performa Mahasiswa	18
2.5. Information Gain	18
2.6. Peneliti Terdahulu	19
BAB III METODOLOGI PENELITIAN	22
3.1. Metode Penelitian	22
3.2. Desain Penelitian	22
3.3. Pengumpulan Data	24
3.4. Data Preprocessing	25
3.5. Validasi	27
3.6. Model Yang Diusulkan	28
3.7. Evaluasi	28

BAB IV HASIL DAN PEMBAHASAN	30
4.1. Hasil dan Pembahasan	30
4.1.1. Pengujian Sebelum Seleksi Atribut	31
4.1.2. Pengujian Setelah Seleksi Information Gain	35
4.2 Evaluasi	39
5.1. Kesimpulan	44
5.2. Saran	44
DAFTAR PUSTAKA	45
DAFTAR RIWAYAT HIDUP	48



UNIVERSITAS MIKROSKIL

DAFTAR GAMBAR

Gambar 2. 1. Tahap-tahap data mining	5
Gambar 2. 2. contoh sampel dataset	5
Gambar 2. 3. Jumlah missing values per atribut dari dataset contoh	6
Gambar 2. 4. Sampel atribut sebelum dan sesudah missing values imputation	7
Gambar 2. 5. Transformasi dataset sampel	7
Gambar 2. 6. Alur Proses Algoritma Naive Bayes.....	15



UNIVERSITAS
MIKROSKIL

DAFTAR TABEL

Tabel 2. 1. Tabel Confusion Matrix	17
Tabel 2. 2. Kategori nilai AUC.....	18
Tabel 2. 3. Penelitian Terdahulu[7], [14], [24]–[26]	19



UNIVERSITAS MIKROSKIL