

**ANALISIS DAN IMPLEMENTASI DATA MINING UNTUK  
MEMPREDIKSI HASIL BELAJAR MAHASISWA  
MENGUNAKAN METODE NAÏVE BAYES PADA FAKULTAS  
TEKNIK PRODI TEKNOLOGI INFORMASI UNIVERSITAS  
MUHAMMADIYAH PALEMBANG**



**TESIS**

**MUHAMMAD FEBRIAN**

**232420022**

**ENTERPRISE IT INFRASTRUCTURE**

**PROGRAM STUDI TEKNIK INFORMATIKA – S2**

**PROGRAM PASCASARJANA**

**UNIVERSITAS BINA DARMA PALEMBANG**

**2026**

**ANALISIS DAN IMPLEMENTASI DATA MINING UNTUK  
MEMPREDIKSI HASIL BELAJAR MAHASISWA  
MENGUNAKAN METODE NAÏVE BAYES PADA FAKULTAS  
TEKNIK PRODI TEKNOLOGI INFORMASI UNIVERSITAS  
MUHAMMADIYAH PALEMBANG**



**Tesis ini diajukan sebagai salah satu syarat  
untuk memperoleh gelar**

**MAGISTER KOMPUTER**

**Muhammad Febrian**

**232420022**

**ENTERPRISE IT INFRASTRUCTURE**

**PROGRAM STUDI TEKNIK INFORMATIKA S2**

**PROGRAM PASCASARJANA**

**UNIVERSITAS BINA DARMA**

**2026**

### Halaman Pengesahan Pembimbing Tesis

Judul Tesis: ANALISIS DAN IMPLEMENTASI DATA MINING UNTUK  
MEMPREDIKSI HASIL BELAJAR MAHASISWA MENGGUNAKAN  
METODE NAIVE BAYES PADA FAKULTAS TEKNIK PRODI  
TEKNOLOGI INFORMASI UNIVERSITAS MUHAMMADIYAH  
PALEMBANG

Oleh Muhammad Febrian , NIM 232420022 Tesis ini telah disetujui dan disahkan oleh  
Tim Penguji Program Studi Teknik Informatika – S2 konsentrasi ENTERPRISE IT  
INFRASTRUCTURE, Program Pascasarjana Universitas Bina Darma pada 05 Maret 2026  
dan telah dinyatakan LULUS.

Palembang, 05 Maret 2026  
Mengetahui,  
Program Studi Teknik Informatika – S2  
Universitas Bina Darma

Ketua,



Dr. Usman Ependi, S.Kom., M.Kom.

Pembimbing,



Dr. Tata Sutabri, S.Kom., M.M.S.I

## Halaman Pengesahan Penguji Tesis

Judul Tesis: ANALISIS DAN IMPLEMENTASI DATA MINING UNTUK  
MEMPREDIKSI HASIL BELAJAR MAHASISWA MENGGUNAKAN  
METODE NAIVE BAYES PADA FAKULTAS TEKNIK PRODI  
TEKNOLOGI INFORMASI UNIVERSITAS MUHAMMADIYAH  
PALEMBANG

Oleh Muhammad Febrian , NIM 232420022 Tesis ini telah disetujui dan disahkan oleh  
Tim Penguji Program Studi Teknik Informatika – S2 konsentrasi ENTERPRISE IT  
INFRASTRUCTURE, Program Pascasarjana Universitas Bina Darma pada 05 Maret 2026  
dan telah dinyatakan LULUS.

Palembang, 05 Maret 2026  
Mengetahui,  
Program Pascasarjana  
Universitas Bina Darma  
Direktur,

  
PROGRAM PASCASARJANA

Prof. Dr. Ir. Achmad Syarifudin, M.Sc.

Penguji I,



Dr. Tata Sutabri, S.Kom., M.M.S.I.

Penguji II,



Prof. Dr. Edi Surya Negara, M.Kom.

Penguji III,



Dr. Usman Ependi, S.Kom., M.Kom.

## **SURAT PERNYATAAN**

Saya yang bertanda tangan dibawah ini:

Nama : MUHAMMAD FEBRIAN

NIM : 232410024

Dengan ini menyatakan bahwa:

1. Karya tulis Saya (Tesis) ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik baik (Magister, Sarjana, dan Ahli Madya) di Universitas Bina Darma;
2. Karya tulis ini murni gagasan, rumusan dan penelitian Saya sendiri dengan arahan tim pembimbing;
3. Dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkan ke dalam daftar pustaka;
4. Karena yakin dengan keaslian karya tulis ini, Saya menyatakan bersedia Tesis/Skripsi/Tugas Akhir, yang Saya hasilkan di unggah ke internet;
5. Surat Pernyataan ini Saya tulis dengan sungguh-sungguh dan apabila terdapat penyimpangan atau ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima sanksi dengan aturan yang berlaku di perguruan tinggi ini.

Demikian Surat Pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya.

Palembang, 05 Maret 2026

Yang Membuat Pernyataan,



MUHAMMAD FEBRIAN

NIM : 232410024

## ABSTRAK

Penelitian ini bertujuan menganalisis dan mengimplementasikan metode data mining menggunakan algoritma *Naïve Bayes* dalam memprediksi hasil belajar mahasiswa pada Fakultas Teknik Program Studi Teknologi Informasi Universitas Muhammadiyah Palembang. Latar belakang penelitian ini didasarkan pada pemanfaatan data akademik yang belum optimal dalam mendukung pengambilan keputusan strategis di lingkungan perguruan tinggi. Data yang digunakan berasal dari Sistem Informasi Akademik (SIKAD) serta hasil kuesioner mahasiswa, dengan total dataset sekitar 200 mahasiswa. Variabel yang dianalisis meliputi nilai mata kuliah, Indeks Prestasi Kumulatif (IPK), tingkat kehadiran, dan jumlah SKS. Proses penelitian dilakukan melalui tahapan pengumpulan data, preprocessing (data cleaning, transformasi, dan reduksi), pembagian data training dan testing dengan rasio 80:20, pemodelan menggunakan algoritma *Naïve Bayes*, serta evaluasi menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan F1-score. Hasil penelitian menunjukkan bahwa model *Naïve Bayes* mampu memberikan performa klasifikasi yang baik dengan tingkat akurasi sebesar 82%, *precision* 85%, *recall* 82%, dan F1-score 83%. Model lebih akurat dalam mengklasifikasikan mahasiswa kategori “Sangat Aktif” dibandingkan kategori “Kurang Aktif”. Temuan ini menunjukkan bahwa metode *Naïve Bayes* efektif digunakan untuk memprediksi hasil belajar mahasiswa dan dapat dimanfaatkan sebagai sistem pendukung keputusan dalam mengidentifikasi mahasiswa yang memerlukan perhatian akademik lebih lanjut.

Kata Kunci : *Data Mining*, *Naïve Bayes*, Hasil Belajar Mahasiswa, Prediksi Akademik, *Confusion Matrix*, Sistem Pendukung Keputusan

## **ABSTRACT**

*This study aims to analyze and implement data mining methods using the Naïve Bayes algorithm in predicting student learning outcomes at the Faculty of Engineering, Information Technology Study Program, University of Muhammadiyah Palembang. The background of this research is based on the use of academic data that is not optimal in supporting strategic decision-making in the university environment. The data used came from the Academic Information System (SIKAD) and the results of student questionnaires, with a total dataset of around 200 students. The variables analyzed included course scores, Cumulative Grade Point Average (GPA), attendance rate, and number of credits. The research process was carried out through the stages of data collection, preprocessing (data cleaning, transformation, and reduction), data sharing training and testing with an 80:20 ratio, modeling using the Naïve Bayes algorithm, and evaluation using confusion matrix, accuracy, precision, recall, and F1-score. The results showed that the Naïve Bayes model was able to provide good classification performance with an accuracy level of 82%, accuracy of 85%, recall of 82%, and an F1-score of 83%. The model is more accurate in classifying students in the "Very Active" category than in the "Less Active" category. These findings show that the Naïve Bayes method is effectively used to predict student learning outcomes and can be used as a decision support system in identifying students who need further academic attention.*

*Keywords : Data Mining, Naïve Bayes, Student Learning Outcomes, Academic Predictions, Confusion Matrix, Decision Support Systems*

## MOTTO DAN PERSEMBAHAN

### MOTTO :

1. Sukses tidak datang dari apa yang diberikan oleh orang lain, Melainkan datang dari keyakinan dan kerja keras diri kita sendiri.
2. Usaha Keras itu tak akan mengkhianati Hasil.
3. Semua impian kita dapat menjadi nyata, jika kita memiliki keberanian untuk mengejanya.
4. Jangan menunda – nunda waktu anda biar pekerjaan anda bisa lancar.
5. Gunakan waktu anda dengan baik.
6. Tetap Semangat Walaupun badai selalu datang pada waktunya.
7. Kejarlah pendidikan setinggi langit.

### PERSEMBAHAN

- Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya yang tak terhingga.
- Baginda Nabi Muhammad Shallallahu alaihi wasallam.
- Kedua Orang Tua saya yang selalu support Dan Memotivasiku.
- Saudara – Saudaraku yang selalu memotivasiku.
- Dosen Pembimbing Dr. Tata Sutabri, S.Kom., M.M.S.I. terima kasih atas bimbingannya selama ini.
- Teman-teman seangkatan 2023
- Almamater Universitas Bina Darma.
- Terima kasih kepada teman teman ku Magister Teknik Informatika, Terima kasih kepada Mamang ku Karnadi, S.Kom., M.Kom turut mensupport saya untuk menyelesaikan Laporan tesis saya.
- Pada Teman dan Rekan-rekan kerja yang mendukung dan membantu Penulis penyelesaian S2 ini.

## KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa atas berkat rahmatnya dan karunia-Nya skripsi saya dapat selesai dengan baik untuk memenuhi salah satu syarat untuk mendapatkan gelar Megister Komputer di Pascasarjana Universitas Bina Darma Palembang

Dalam penulisan tesis ini, tentunya masih jauh dari sempurna. Hal ini dikarenakan keterbatasannya pengetahuan yang dimiliki. Oleh karena itu dalam rangka melengkapi kesempurnaan dari penulis skripsi ini diharapkan adanya saran dan kririk yang diberikan bersifat membangun.

Pada kesempatan yang baik ini, tak lupa penulis menghaturkan terima kasih kepada semua pihak yang telah memberi bimbingan, pengarahan, nasehat dan pemikiran dalam penulisan skripsi ini, terutama kepada :

1. Dr. Sunda Ariana, M.Pd., M.M. selaku Rektor Universitas Bina Darma Palembang.
2. Prof. Dr. Ir. Achmad Syarifudin, M.Sc. selaku Direktur Pascasarjana Universitas Bina Darma Palembang
3. Dr. Usman Ependi, S.Kom.,M.Kom ketua Program Studi Megister Teknik Informatika.
4. Dr. Tata Sutabri, S.Kom., M.M, S.I selaku dosen pembimbing yang telah memberikan bimbingan dalam penulisan Tesis.
5. Prof Dr. Edi Surya Negara, S.Kom., M.Kom selaku penguji satu saya yang telah memberikan masukan dan saran dalam penulisan tesis.
6. Segenap Dosen Magister Teknik Informatika Universitas Bina Darma Palembang yang telah memberikan ilmu pengetahuan dan bimbingannya.
7. Orang tua tercinta, terima kasih atas doanya dan serta cinta yang telah diberikan.
8. Terima kasih kepada Mamang ku Karnadi, S.Kom., M.Kom turut mensupport saya untuk menyelesaikan Laporan tesis saya.
9. Keluarga besar Universitas Bina Darma Palembang khususnya teman-teman seperjuangan pada Prodi Magister teknik informatika atas semua

dukungan, semangat, serta kerja samanya.

10. Teman – teman Magister Teknik Informatika angkatan 2023 tercinta
11. Pada Teman dan Rekan-rekan kerja yang mendukung dan membantu Penulis menyelesaikan S2 ini.
12. Penulis menyadari penulisan tesis ini tidak luput dari berbagai kekurangan. Penulis mengharapkan saran dan kritik demi kesempurnaan dan perbaikannya sehingga akhirnya laporan skripsi ini dapat memberikan manfaat bagi bidang pendidikan dan bisa dikembangkan lebih lanjut lagi. Aamiin.

Akhir kata dari penulisan menyadari bahwa penulisan tesis ini belum mencapai kesempurnaan dalam arti yang sebenarnya. Namun penulis berharap penulisan tesis ini dapat bermanfaat bagi penulis sendiri dan para pembaca pada umumnya.

Palembang,

Muhammad Febrian

## DAFTAR ISI

|                                                         |             |
|---------------------------------------------------------|-------------|
| <b>HALAMAN DEPAN .....</b>                              | <b>i</b>    |
| <b>HALAMAN DEPAN TESIS .....</b>                        | <b>ii</b>   |
| <b>HALAMAN PENGESAHAN PEMBIMBING TESIS .....</b>        | <b>iii</b>  |
| <b>HALAMAN PENGESAHAN UJIAN TESIS....</b>               | <b>iv</b>   |
| <b>SURAT PERNYATAAN .....</b>                           | <b>v</b>    |
| <b>ABSTRAK.....</b>                                     | <b>vi</b>   |
| <b><i>ABSTRACT</i>.....</b>                             | <b>vii</b>  |
| <b>MOTTO DAN PERSEMBAHAN .....</b>                      | <b>viii</b> |
| <b>KATA PENGANTAR .....</b>                             | <b>ix</b>   |
| <b>DAFTAR ISI .....</b>                                 | <b>xi</b>   |
| <b>DAFTAR GAMBAR .....</b>                              | <b>xv</b>   |
| <b>DAFTAR TABEL .....</b>                               | <b>xvi</b>  |
| <b>BAB I PENDAHULUAN .....</b>                          | <b>1</b>    |
| 1.1 Latar Belakang.....                                 | 1           |
| 1.2 Identifikasi masalah.....                           | 3           |
| 1.3 Perumusan Masalah.....                              | 4           |
| 1.4 Tujuan penelitian .....                             | 4           |
| 1.5 Kebaruan .....                                      | 4           |
| 1.6 Manfaat penelitian .....                            | 6           |
| 1.7 Pembatasan masalah.....                             | 6           |
| 1.8 Sistematika Penulisan.....                          | 7           |
| <b>BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI .....</b> | <b>9</b>    |
| 2.1 Tinjauan Pustaka.....                               | 9           |
| 2.1.1 <i>Data Maining</i> .....                         | 9           |
| 2.2 Educational Data Maining.....                       | 11          |

|                                |                                                     |           |
|--------------------------------|-----------------------------------------------------|-----------|
| 2.3                            | Hasil Belajar Mahasiswa .....                       | 12        |
| 2.4                            | <i>Naïve Bayes Classifier</i> .....                 | 13        |
| 2.5                            | Confusion Matrix .....                              | 14        |
| 2.6                            | Penelitian Terdahulu.....                           | 16        |
| 2.7                            | Kerangka kerja yang digunakan .....                 | 21        |
| 2.7.1                          | Studi literatur.....                                | 21        |
| 2.7.2                          | Identifikasi masalah dan tujuan.....                | 21        |
| 2.7.3                          | Pengumpulan data .....                              | 21        |
| 2.7.4                          | Implementasi algoritma naïve bayes .....            | 21        |
| 2.7.5                          | Evaluasi Kinerja Sistem .....                       | 21        |
| 2.7.6                          | Analisis dan Pembahasan Hasil.....                  | 22        |
| 2.8                            | Google colabs.....                                  | 22        |
| <b>BAB III METODOLOGI.....</b> |                                                     | <b>24</b> |
| 3.1                            | Metode Penelitian.....                              | 24        |
| 3.2                            | Waktu dan tempat penelitian .....                   | 24        |
| 3.2.1                          | Tempat penelitian .....                             | 24        |
| 3.2.2                          | Waktu penelitian .....                              | 25        |
| 3.3                            | Prosedur Penelitian.....                            | 25        |
| 3.3.1                          | Pengumpulan data .....                              | 26        |
| 3.3.2                          | Preprocessing Data .....                            | 26        |
| 3.3.3                          | Pemodelan (Modeling) .....                          | 27        |
| 3.3.4                          | Evaluasi Model.....                                 | 28        |
| 3.3.5                          | Analisis dan Interpretasi Hasil .....               | 28        |
| 3.4                            | Variabel Input Berdasarkan Data Kuesioner .....     | 28        |
| 3.4.1                          | Tingkat Kehadiran Mahasiswa .....                   | 29        |
| 3.4.2                          | Distribusi Responden Berdasarkan Jenis Kelamin..... | 29        |
| 3.4.3                          | Indeks Prestasi Kumulatif (IPK) Mahasiswa.....      | 30        |

|                                                                                   |           |
|-----------------------------------------------------------------------------------|-----------|
| <b>BAB IV HASIL DAN PEMBAHASAN .....</b>                                          | <b>32</b> |
| 4.1 Hasil .....                                                                   | 32        |
| 4.2 Implementasi Naïve Bayes .....                                                | 33        |
| 4.3 Preprocessing Data .....                                                      | 34        |
| 4.3.1 Data Cleaning (Pembersihan Data) .....                                      | 34        |
| 4.3.2 Data Transformation .....                                                   | 35        |
| 4.4 Preprocessing Data .....                                                      | 38        |
| 4.4.1 Import Data .....                                                           | 39        |
| 4.4.2 Seleksi Atribut .....                                                       | 39        |
| 4.4.3 Data cleaning .....                                                         | 39        |
| 4.4.4 Transformasi Data .....                                                     | 40        |
| 4.4.5 Encoding Data .....                                                         | 41        |
| 4.4.6 Pembentukan Dataset Akhir .....                                             | 41        |
| 4.5 Evaluasi Kinerja Model Berdasarkan Hasil Kuesioner .....                      | 41        |
| 4.5.1 Tingkat Kehadiran Mahasiswa .....                                           | 42        |
| 4.5.2 Distribusi IPK Responden .....                                              | 43        |
| 4.5.3 Tingkat Motivasi Belajar Mahasiswa .....                                    | 44        |
| 4.5.4 Persepsi Mahasiswa terhadap Penggunaan Data Akademik untuk<br>Prediksi..... | 44        |
| 4.5.5 Keterkaitan Hasil Kuesioner dengan Kinerja Model .....                      | 45        |
| 4.6 Confusion Matrix .....                                                        | 46        |
| 4.7 Hasil Feature Importanc .....                                                 | 48        |
| 4.8 Performa Model.....                                                           | 49        |
| 4.9 Distribusi Kelas .....                                                        | 51        |
| 4.10 5-Fold Cross-Validation .....                                                | 53        |
| 4.11 Hasil Naïve Bayes .....                                                      | 54        |
| 4.12 Hasil Distribusi Akurasi dan Akurasi per Eksperimen .....                    | 56        |

|                            |                                     |           |
|----------------------------|-------------------------------------|-----------|
| 4.13                       | Distribusi Setiap Fitur Input ..... | 58        |
| <b>BAB IV PENUTUP.....</b> |                                     | <b>61</b> |
| 5.1                        | Kesimpulan.....                     | 61        |
| 5.2                        | Saran.....                          | 61        |
| <b>DAFTAR PUSTAKA.....</b> |                                     | <b>62</b> |
| <b>LAMPIRAN .....</b>      |                                     | <b>64</b> |



## DAFTAR GAMBAR

|                                                                                      |    |
|--------------------------------------------------------------------------------------|----|
| <b>Gambar 2.1</b> Proses data mining atau <i>Knowledge Discovery from Data</i> ..... | 11 |
| <b>Gambar 4.1</b> <i>Confusion Matrix</i> .....                                      | 46 |
| <b>Gambar 4.2</b> <i>Feature Importance</i> .....                                    | 48 |
| <b>Gambar 4.3</b> Performa Model .....                                               | 50 |
| <b>Gambar 4.4</b> Distribusi Kelas.....                                              | 52 |
| <b>Gambar 4.5</b> <i>5-Fold Cross-Validation</i> .....                               | 53 |
| <b>Gambar 4.6</b> Hasil <i>Naïve Bayes</i> .....                                     | 55 |
| <b>Gambar 4.7</b> Hasil Distribusi akurasi dan akurasi eksperimen .....              | 57 |
| <b>Gambar 4.8</b> Distribusi Setiap Fitur Input.....                                 | 59 |

## DAFTAR TABEL

|                                                                              |    |
|------------------------------------------------------------------------------|----|
| <b>Tabel 2.1</b> Confusion Matrix .....                                      | 15 |
| <b>Tabel 2.2</b> Akurasi .....                                               | 15 |
| <b>Tabel 2.3</b> Presisi .....                                               | 15 |
| <b>Tabel 2.4</b> Recall .....                                                | 16 |
| <b>Tabel 2.5</b> Penelitian terdahulu.....                                   | 16 |
| <b>Tabel 3.1</b> Waktu Penelitian .....                                      | 25 |
| <b>Tabel 3.2</b> Tingkat kehadiran mahasiswa .....                           | 29 |
| <b>Tabel 3.3</b> Distribusi Responden Berdasarkan Jenis Kelamin .....        | 30 |
| <b>Tabel 3.4</b> Distribusi IPK Mahasiswa .....                              | 30 |
| <b>Tabel 4.1</b> Ringkas distribusi data .....                               | 33 |
| <b>Tabel 4.2</b> Nilai Akademik .....                                        | 35 |
| <b>Tabel 4.3</b> Transformasi Kehadiran Mahasiswa .....                      | 36 |
| <b>Tabel 4.4</b> Encoding Skala Likert.....                                  | 37 |
| <b>Tabel 4.5</b> Encoding Tingkat Aktivitas.....                             | 37 |
| <b>Tabel 4.6</b> Transformasi Data.....                                      | 40 |
| <b>Tabel 4.7</b> Transformasi IPK.....                                       | 40 |
| <b>Tabel 4.8</b> Kategori Motivasi .....                                     | 41 |
| <b>Tabel 4.9</b> Tingkat Kehadiran Mahasiswa .....                           | 42 |
| <b>Tabel 4.10</b> Distribusi IPK Mahasiswa .....                             | 43 |
| <b>Tabel 4.11</b> Tingkat Motivasi Belajar Mahasiswa .....                   | 44 |
| <b>Tabel 4.12</b> Persepsi Mahasiswa terhadap Penggunaan Data Akademik ..... | 45 |