

**ANALISIS SENTIMEN UMPAN BALIK MAHASISWA  
TERHADAP DOSEN MENGGUNAKAN NAÏVE BAYES**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**IRENIUS PUTRA AGUNG**

**22.11.5119**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2026**

**ANALISIS SENTIMEN UMPAN BALIK MAHASISWA  
TERHADAP DOSEN MENGGUNAKAN NAÏVE BAYES**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**IRENIUS PUTRA AGUNG**

**22.11.5119**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2026**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN UMPAN BALIK MAHASISWA TERHADAP  
DOSEN MENGGUNAKAN NAÏVE BAYES**

yang disusun dan diajukan oleh

**Irenius Putra Agung**

**22.11.5119**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 8 April 2026

**Dosen Pembimbing,**

**Rizqi Sukma Kharisma, M.Kom**  
**NIK. 190302215**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN UMPAN BALIK MAHASISWA TERHADAP  
DOSEN MENGGUNAKAN NAÏVE BAYES**

yang disusun dan diajukan oleh

**Irenius Putra Agung**

**22.11.5119**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 22 April 2026

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Rumini, S.Kom., M.Kom**  
**NIK. 190302246**

**Sharazita Dyah Anggita, S.Kom., M.Kom**  
**NIK. 190302285**

**Rizqi Sukma Kharisma, S.Kom., M.Kom**  
**NIK. 190302215**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 22 April 2026

**DEKAN FAKULTAS ILMU KOMPUTER**



**Prof. Dr. Kusrini, M.Kom.**  
**NIK. 190302106**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Irenius Putra Agung  
NIM : 22.11.5119

Menyatakan bahwa Skripsi dengan judul berikut:

### ANALISIS SENTIMEN UMPAN BALIK MAHASISWA TERHADAP DOSEN MENGGUNAKAN NAÏVE BAYES

Dosen Pembimbing : Rizqi Sukma Kharisma, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta 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 AMIKOM Yogyakarta.
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.

Yogyakarta, 22 April 2026

Yang Menyatakan,



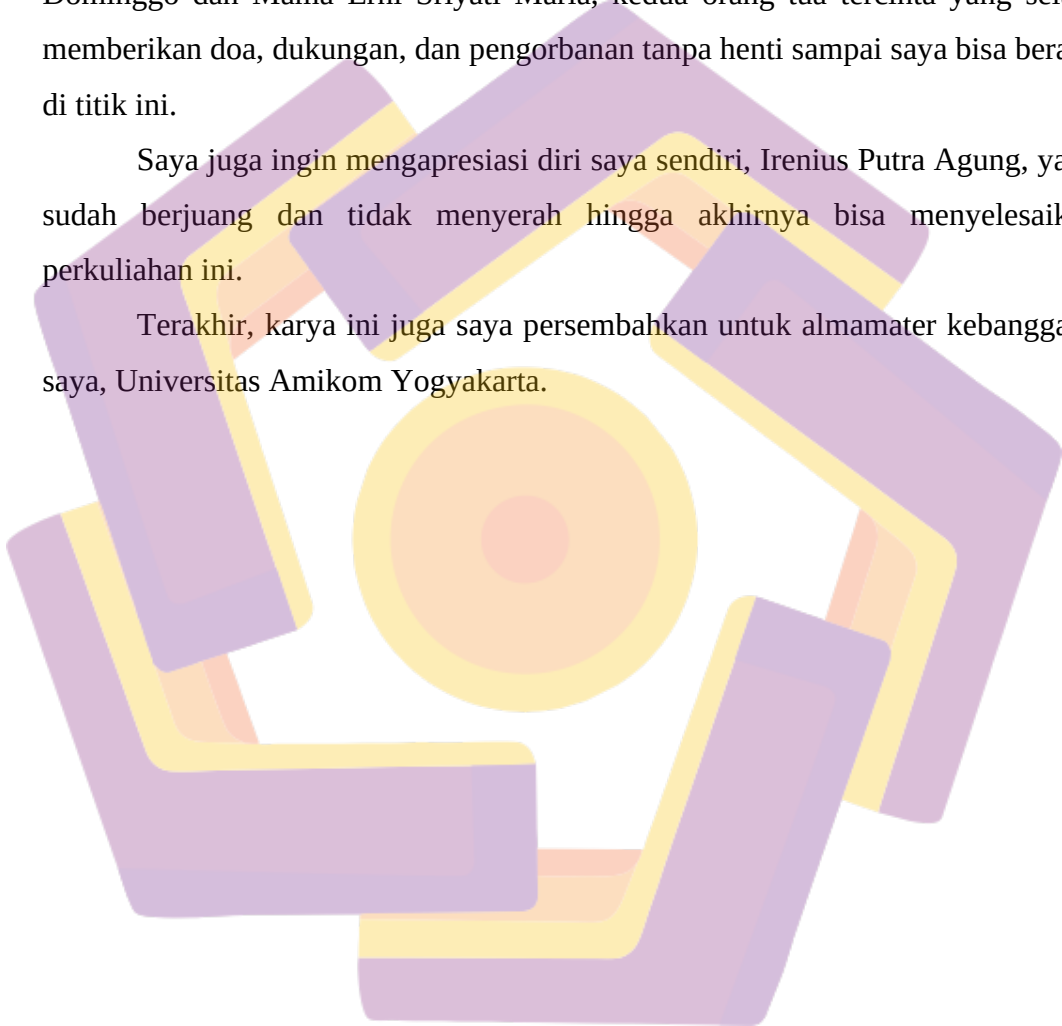
Irenius putra agung

## HALAMAN PERSEMBAHAN

Puji syukur kepada Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga saya dapat menyelesaikan skripsi ini. Dengan penuh rasa syukur dan cinta kasih, karya sederhana ini saya persembahkan kepada Bapak Kristianus Domingo dan Mama Erni Sriyati Maria, kedua orang tua tercinta yang selalu memberikan doa, dukungan, dan pengorbanan tanpa henti sampai saya bisa berada di titik ini.

Saya juga ingin mengapresiasi diri saya sendiri, Irenius Putra Agung, yang sudah berjuang dan tidak menyerah hingga akhirnya bisa menyelesaikan perkuliahan ini.

Terakhir, karya ini juga saya persembahkan untuk almamater kebanggaan saya, Universitas Amikom Yogyakarta.



## KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa karena atas berkat dan rahmat-Nya, skripsi ini dapat diselesaikan dengan baik sebagai salah satu syarat menyelesaikan studi di Universitas Amikom Yogyakarta.

Penulis mengucapkan terima kasih kepada Bapak Risky Sukma Kharisma, M.Kom. selaku dosen pembimbing yang telah memberikan arahan dan dukungan selama proses penyusunan skripsi. Terima kasih juga kepada Ibu Rumini, M.Kom. selaku dosen pembimbing akademik atas bimbingannya selama masa perkuliahan.

Ucapan terima kasih yang sebesar-besarnya juga penulis sampaikan kepada kedua orang tua tercinta, Bapak Kristianus Domingo dan Ibu Erni Sriyati Maria, atas doa, dukungan, dan kasih sayang yang selalu diberikan.

Tidak lupa, penulis juga berterima kasih kepada seluruh dosen dan staf Universitas Amikom Yogyakarta yang telah membantu selama masa perkuliahan.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna karena keterbatasan pengetahuan dan pengalaman yang dimiliki. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari berbagai pihak demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan dan bagi para pembaca.

Yogyakarta, 22 April 2026

Penulis

## DAFTAR ISI

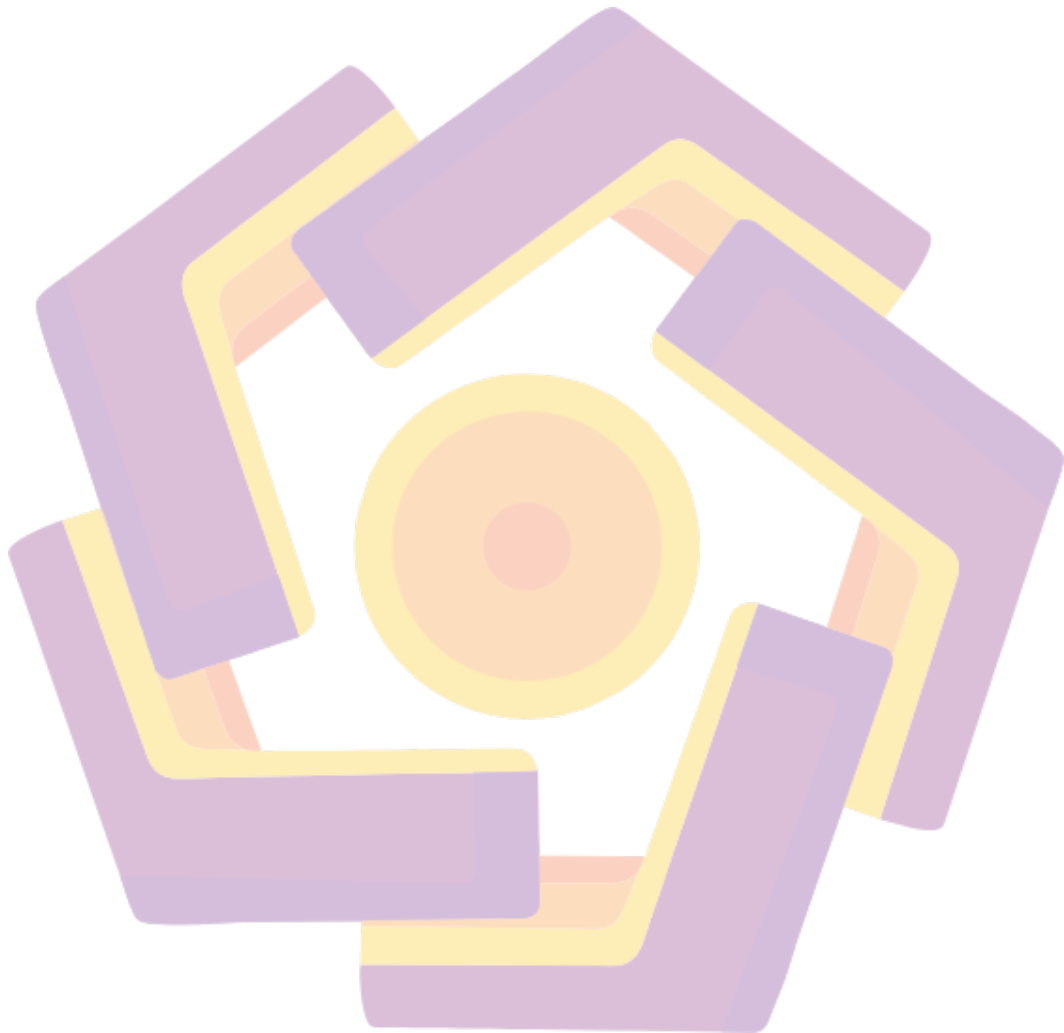
|                                    |      |
|------------------------------------|------|
| HALAMAN JUDUL .....                | i    |
| HALAMAN PERSETUJUAN.....           | ii   |
| HALAMAN PENGESAHAN .....           | iii  |
| .....                              | iv   |
| HALAMAN PERSEMBAHAN .....          | v    |
| KATA PENGANTAR .....               | vi   |
| DAFTAR ISI.....                    | vii  |
| DAFTAR TABEL.....                  | x    |
| DAFTAR GAMBAR.....                 | xi   |
| DAFTAR LAMBANG DAN SINGKATAN ..... | xii  |
| DAFTAR ISTILAH .....               | xiii |
| INTISARI .....                     | xiv  |
| ABSTRACT.....                      | xv   |
| BAB I PENDAHULUAN.....             | 1    |
| 1.1 Latar Belakang .....           | 1    |
| 1.2 Rumusan Masalah .....          | 2    |
| 1.3 Batasan Masalah.....           | 3    |
| 1.4 Tujuan Penelitian.....         | 3    |
| 1.5 Manfaat Penelitian.....        | 3    |
| BAB II TINJAUAN PUSTAKA .....      | 4    |
| 2.1 Studi litelatur .....          | 4    |
| 2.2 Dasar Teori .....              | 14   |
| 2.2.1 Analisis Sentimen .....      | 14   |

|                                                              |           |
|--------------------------------------------------------------|-----------|
| 2.2.2 Natural Language Processing.....                       | 14        |
| 2.2.3 Preprocessing Teks .....                               | 14        |
| 2.2.4 Ekstraksi Fitur Teks .....                             | 15        |
| 2.2.5 Algoritma Naive Bayes .....                            | 15        |
| 2.2.6 Evaluasi Model Klasifikasi .....                       | 16        |
| 2.2.7 Sentiment Scoring .....                                | 17        |
| <b>BAB III METODE PENELITIAN .....</b>                       | <b>18</b> |
| 3.1 Objek Penelitian.....                                    | 18        |
| 3.2 Alur Penelitian .....                                    | 19        |
| 3.2.1 Penjelasan tiap alur : .....                           | 20        |
| 3.2.1.1 Pengumpulan data.....                                | 20        |
| 3.2.1.2 Teks preprocessing.....                              | 20        |
| 3.2.1.3 Transformasi data TF-IDF .....                       | 21        |
| 3.2.1.4 Split data Test / training.....                      | 21        |
| 3.2.1.5 Training naive bayes .....                           | 21        |
| 3.2.1.6 Prediksi sentimen .....                              | 22        |
| 3.2.1.7 Evaluasi model.....                                  | 23        |
| 3.2.1.8 Sentiment scoring.....                               | 23        |
| 3.2.1.9 Analisis & Interpretasi Hasil.....                   | 23        |
| 3.3 Alat dan Bahan.....                                      | 24        |
| 3.3.1 Data Penelitian .....                                  | 24        |
| 3.3.2 Alat dan Perangkat Lunak.....                          | 24        |
| 3.3.3 Perangkat Keras .....                                  | 25        |
| 3.4 Metode Analisis Data.....                                | 26        |
| 3.4.1 Eksplorasi Data (Exploratory Data Analysis – EDA)..... | 26        |

|                             |                                          |    |
|-----------------------------|------------------------------------------|----|
| 3.4.2                       | Praproses Data (Text Preprocessing)      | 26 |
| 3.4.3                       | Ekstraksi Fitur (TF-IDF)                 | 27 |
| 3.4.4                       | Pelatihan Model (Training Model)         | 27 |
| 3.4.5                       | Evaluasi Model                           | 28 |
| 3.4.6                       | Interpretasi Hasil dan Sentiment Scoring | 28 |
| BAB IV HASIL DAN PEMBAHASAN |                                          | 30 |
| 4.1                         | Pengambilan data                         | 30 |
| 4.1.1                       | Karakteristik dataset                    | 30 |
| 4.2                         | Text Preprocessing                       | 32 |
| 4.3                         | Ekstraksi Fitur ( TF – IDF )             | 33 |
| 4.4                         | Split Data ( train & test )              | 33 |
| 4.5                         | Traning naïve bayes                      | 34 |
| 4.6                         | Evaluasi Model Klasifikasi               | 36 |
| 4.7                         | Sentimen Scoring                         | 39 |
| 4.8                         | Analisis dan interpretasi hasil          | 40 |
| BAB V PENUTUP               |                                          | 42 |
| 5.1                         | Kesimpulan                               | 42 |
| 5.2                         | Saran                                    | 42 |
| DAFTAR PUSTAKA              |                                          | 44 |
| LAMPIRAN                    |                                          | 46 |

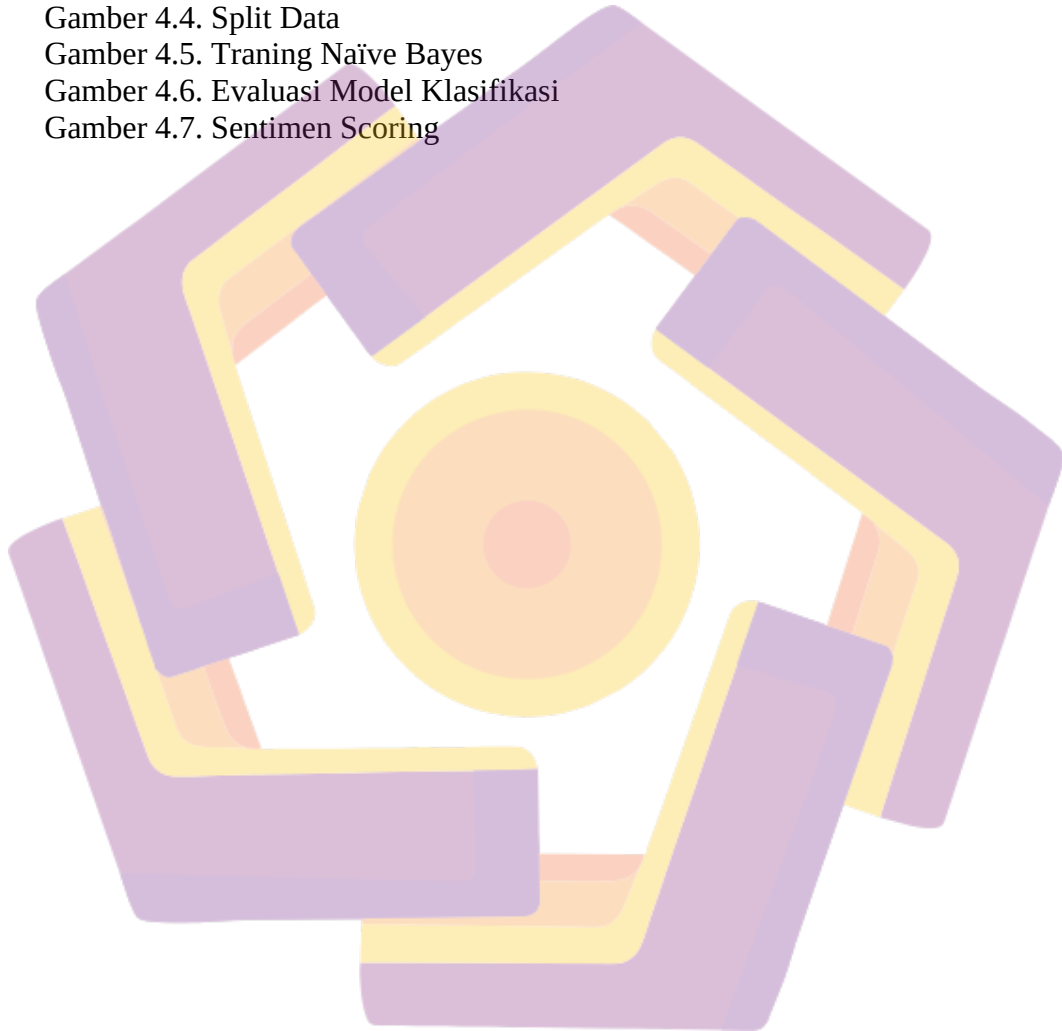
## DAFTAR TABEL

|                                     |    |
|-------------------------------------|----|
| Tabel 2.1. Keaslian Penelitian      | 8  |
| Tabel 3.1. Alat dan Perangkat Lunak | 25 |
| Tabel 3.2. Perangkat Keras          | 25 |

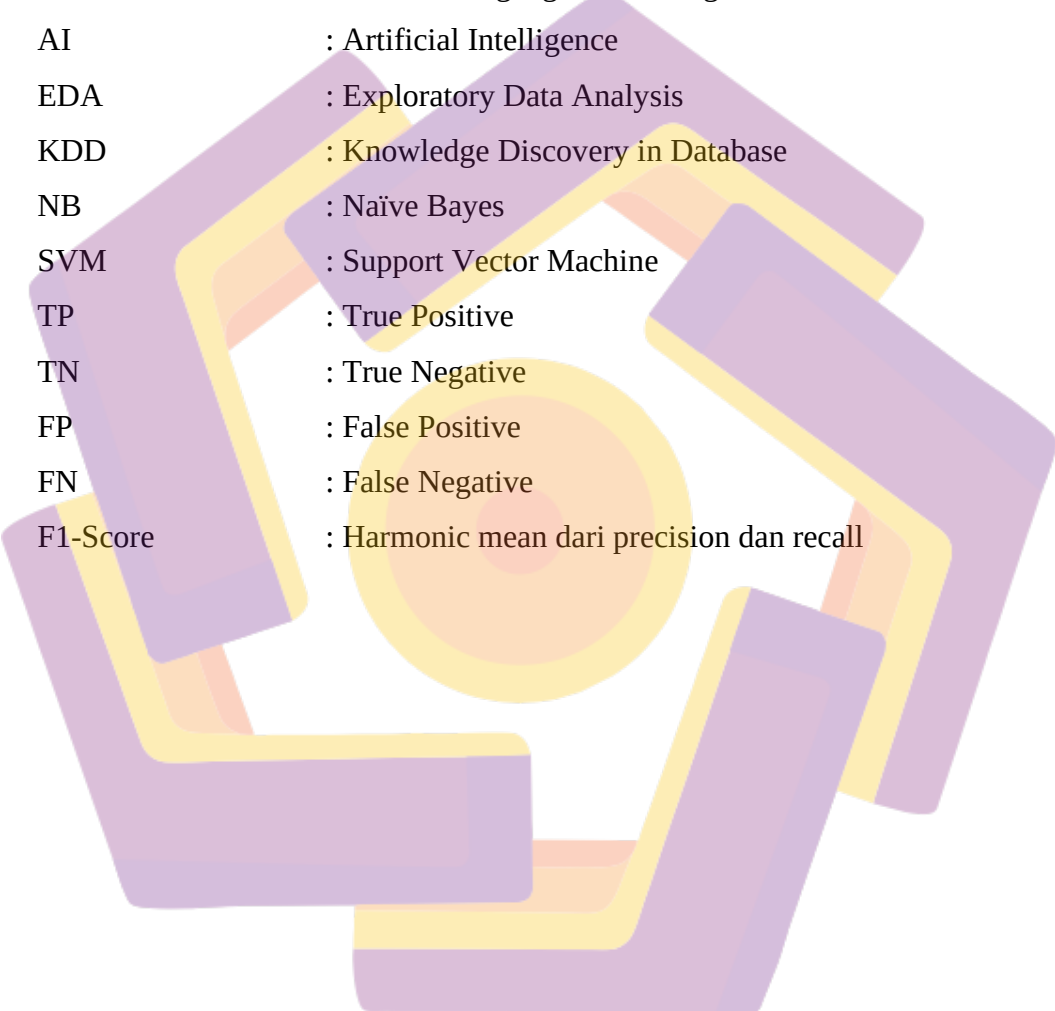


## DAFTAR GAMBAR

|                                        |    |
|----------------------------------------|----|
| Gambar 3.2. Alur Penelitian            | 19 |
| Gambar 4.1. Pengambilan Dataset        | 30 |
| Gambar 4.1.1. Karakteristik Dataset    | 31 |
| Gambar 4.2. Teks Preprocessing         | 32 |
| Gambar 4.3. Ekstraksi Fitur (TF-IDF)   | 33 |
| Gambar 4.4. Split Data                 | 33 |
| Gambar 4.5. Training Naïve Bayes       | 34 |
| Gambar 4.6. Evaluasi Model Klasifikasi | 36 |
| Gambar 4.7. Sentimen Scoring           | 37 |



## DAFTAR LAMBANG DAN SINGKATAN



|          |                                               |
|----------|-----------------------------------------------|
| TF       | : Term Frequency                              |
| IDF      | : Inverse Document Frequency                  |
| TF-IDF   | : Term Frequency – Inverse Document Frequency |
| NLP      | : Natural Language Processing                 |
| AI       | : Artificial Intelligence                     |
| EDA      | : Exploratory Data Analysis                   |
| KDD      | : Knowledge Discovery in Database             |
| NB       | : Naïve Bayes                                 |
| SVM      | : Support Vector Machine                      |
| TP       | : True Positive                               |
| TN       | : True Negative                               |
| FP       | : False Positive                              |
| FN       | : False Negative                              |
| F1-Score | : Harmonic mean dari precision dan recall     |

## DAFTAR ISTILAH



|                   |                                                                |
|-------------------|----------------------------------------------------------------|
| Analisis Sentimen | : proses mengidentifikasi opini atau emosi dari teks           |
| Dataset           | : kumpulan data yang digunakan dalam penelitian                |
| Preprocessing     | : tahapan pembersihan dan persiapan data teks                  |
| Tokenizing        | : proses memecah kalimat menjadi kata-kata                     |
| Stopword Removal  | : proses menghapus kata umum yang tidak memiliki makna penting |
| Stemming          | : proses mengubah kata menjadi bentuk dasar                    |
| Klasifikasi       | : proses pengelompokan data ke dalam kategori tertentu         |
| Model             | : representasi matematis yang digunakan untuk prediksi         |
| Training Data     | : data yang digunakan untuk melatih model                      |
| Testing Data      | : data yang digunakan untuk menguji model                      |
| Confusion Matrix  | : tabel untuk mengevaluasi kinerja model klasifikasi           |
| Akurasi           | : tingkat ketepatan prediksi model secara keseluruhan          |
| Precision         | : tingkat ketepatan prediksi pada kelas tertentu               |
| Recall            | : kemampuan model menemukan data yang relevan                  |
| Sentiment Scoring | : pemberian nilai numerik untuk mengukur kekuatan sentimen     |

## INTISARI

Perkembangan teknologi informasi di era digital menyebabkan peningkatan jumlah data, termasuk data umpan balik mahasiswa terhadap dosen yang umumnya berbentuk teks tidak terstruktur. Permasalahan yang muncul adalah sulitnya menganalisis data tersebut secara manual karena jumlahnya yang besar, sehingga berpotensi menghambat proses evaluasi kualitas pengajaran. Penelitian ini bertujuan untuk menganalisis sentimen umpan balik mahasiswa terhadap dosen menggunakan algoritma Naïve Bayes yang dipadukan dengan sentiment scoring. Metode yang digunakan meliputi tahapan preprocessing teks, ekstraksi fitur menggunakan TF-IDF, pembagian data menjadi training dan testing dengan rasio 80:20, pelatihan model Naïve Bayes, evaluasi menggunakan confusion matrix, serta penerapan sentiment scoring untuk mengukur intensitas sentimen. Hasil penelitian menunjukkan bahwa model memiliki performa yang baik dengan akurasi sebesar 83,18%, precision 82,25%, recall 97,42%, dan F1-score 89,19%. Selain itu, diperoleh rata-rata skor sentimen sebesar 0,69 yang menunjukkan tingkat kepuasan mahasiswa terhadap dosen cenderung tinggi. Penelitian ini memberikan kontribusi dalam menghasilkan metode evaluasi kinerja dosen yang lebih objektif, terukur, dan berbasis data, serta dapat dimanfaatkan oleh institusi pendidikan untuk meningkatkan kualitas pembelajaran. Penelitian selanjutnya disarankan untuk mengembangkan metode dengan algoritma lain atau dataset yang lebih beragam.

**Kata kunci:** analisis sentimen, naïve bayes, TF-IDF, sentiment scoring, umpan balik mahasiswa

## ABSTRACT

The rapid development of information technology in the digital era has led to a significant increase in data volume, including student feedback on lecturers, which is generally in the form of unstructured text. The main problem is the difficulty of analyzing such large-scale data manually, which may hinder the evaluation of teaching quality. This study aims to analyze student feedback sentiment toward lecturers using the Naïve Bayes algorithm combined with sentiment scoring. The research method includes text preprocessing, feature extraction using TF-IDF, data splitting into training and testing sets with an 80:20 ratio, Naïve Bayes model training, evaluation using a confusion matrix, and the application of sentiment scoring to measure sentiment intensity. The results show that the model achieved good performance with an accuracy of 83.18%, precision of 82.25%, recall of 97.42%, and F1-score of 89.19%. Additionally, the average sentiment score of 0.69 indicates a high level of student satisfaction toward lecturers. This study contributes to providing a more objective, measurable, and data-driven evaluation method for lecturer performance, which can be utilized by educational institutions to improve teaching quality. Future research is recommended to explore other algorithms or more diverse datasets.

**Keyword:** sentiment analysis, naïve bayes, TF-IDF, sentiment scoring, student feedback