

TUGAS AKHIR

KLASIFIKASI SENTIMEN ULASAN GAME BUS SIMULATOR INDONESIA DENGAN PENDEKATAN TF-IDF DAN NAÏVE BAYES

**Disusun untuk Melengkapi dan Memenuhi Syarat Kelulusan
Guna Meraih Gelar Sarjana Komputer
Teknik Informatika Universitas Muhammadiyah Jember**



SEPTA DAENG INDAR KURNIAWAN

2110651046

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER**

2026

LEMBAR PERSETUJUAN

TUGAS AKHIR

KLASIFIKASI SENTIMEN ULASAN GAME BUS SIMULATOR INDONESIA DENGAN PENDEKATAN TF-IDF DAN NAÏVE BAYES

Oleh
SEPTA DAENG INDAR KURNIAWAN
2110651046

Telah disetujui bahwa Laporan Tugas Akhir ini
untuk diajukan pada sidang Tugas Akhir sebagai salah satu syarat kelulusan
dan mendapatkan gelar Sarjana Komputer (S.Kom.)
di Universitas Muhammadiyah Jember

Menyetujui,

Pembimbing I



B. Satria Bakti, S.Si., M.Kom
0714107503

Pembimbing II



Ulya Anisatur Rosyidah, M.Kom
0710037903

LEMBAR PENGESAHAN

TUGAS AKHIR

KLASIFIKASI SENTIMEN ULASAN GAME BUS SIMULATOR INDONESIA DENGAN PENDEKATAN TF-IDF DAN NAÏVE BAYES

Oleh

Septa Daeng Indaer Kurniawan

2110651046

Telah mempertanggung jawabkan Laporan Tugas Akhirnya
pada sidang Tugas Akhir tanggal 29 Juli 2026, sebagai salah satu syarat kelulusan
dan mendapatkan gelar Sarjana Komputer (S.Kom)
di Universitas Muhammadiyah Jember

Menyetujui,

Pembimbing I



B. Satria Bakti, S.Si., M.Kom
NIDN. 0714107503

Pembimbing II



Ulya Anisatur Rosyidah, M.Kom
NIDN. 0710037903

Penguji I



Miftahur Rahman, S.Kom., M.Kom
NIDN. 0724039201

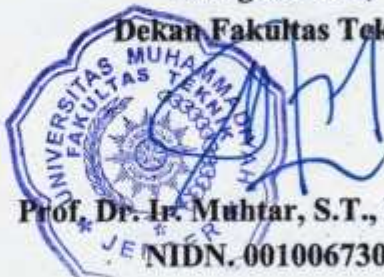
Penguji II



Qurrota A'yun, M.Pd
NIDN. 0703069002

Mengesahkan,

Dekan Fakultas Teknik



Prof. Dr. Ir. Muhtar, S.T., M.T., IPM
NIDN. 0010067301

Mengetahui,

Ketua Program Studi Teknik Informatika



Rosita Yanuarti, S.Kom., M.Cs
NIDN. 0629018601

PERNYATAAN ORISINALITAS

Saya yang bertanda tangan di bawah ini:

Nama : Septa Daeng Indar Kurniawan
NIM : 2110651046
Program Studi : Teknik Informatika
Judul Tugas Akhir : Klasifikasi Sentimen Ulasan Game Bus Simulator Indonesia
Dengan Pendekatan Naïve Bayes

dengan ini menyatakan bahwa karya ilmiah yang saya susun adalah murni hasil pemikiran, penelitian, dan penyusunan saya sendiri. Segala sumber informasi yang berasal atau dikutip dari karya orang lain, baik berupa publikasi, artikel, buku, maupun sumber lainnya, telah disebutkan dengan jelas di dalam daftar pustaka sesuai dengan kaidah penulisan ilmiah.

Saya menyadari bahwa dalam proses penyusunan Tugas Akhir ini, saya memanfaatkan bantuan teknologi kecerdasan buatan (AI) seperti *ChatGPT/Grammarly/Alat AI lain* hanya sebatas untuk:

1. Membantu dalam penyusunan tata bahasa, parafrasa, atau perbaikan redaksi.
2. Memberikan inspirasi ide awal atau kerangka berpikir, yang selanjutnya saya kembangkan secara mandiri.
3. Membantu dalam pengecekan konsistensi format, ejaan, dan tata tulis.

Saya menegaskan bahwa tidak ada bagian dari karya ilmiah ini yang seluruhnya dibuat oleh AI tanpa keterlibatan pemikiran kritis saya sendiri. Tanggung jawab penuh atas isi, keaslian, dan kebenaran karya ilmiah ini ada pada diri saya.

Apabila di kemudian hari terbukti bahwa pernyataan ini tidak benar, maka saya bersedia menerima segala konsekuensi sesuai dengan peraturan yang berlaku di Universitas Muhammadiyah Jember.

Jember, Agustus 2026
Yang membuat pernyataan,



Septa Daeng Indar Kurniawan
Nim: 2110651046

KATA PENGANTAR

Segala puji dan syukur saya panjatkan kepada Allah Subhanahu wa Ta'ala atas rahmat dan karunia-Nya sehingga karya ini dapat saya selesaikan dengan baik. Salawat serta salam semoga senantiasa tercurah kepada Nabi Muhammad Sallallahu 'alaihi wa sallam, teladan dalam menuntut ilmu dan mengamalkannya.

Proses penulisan tugas akhir ini berlangsung secara sistematis, dimulai dari perumusan tujuan dan identifikasi masalah, pengumpulan data, dan literatur yang relevan, hingga penyusunan kesimpulan. Setiap tahap penulisan memerlukan ketelitian, konsistensi, serta pemikiran kritis untuk memastikan akurasi dan relevansi informasi yang saya sajikan. Tantangan yang muncul, baik berupa keterbatasan sumber maupun waktu, saya hadapi dengan metode perencanaan dan evaluasi secara bertahap.

Melalui proses ini, saya memperoleh pemahaman yang lebih mendalam terkait topik yang dibahas, sekaligus mengasah kemampuan analisis, sintesis, dan penyampaian argumentasi secara ilmiah. Semoga hasil tugas akhir ini dapat memberikan kontribusi intelektual yang bermanfaat bagi pengembangan ilmu pengetahuan dan menjadi amal kebaikan yang diterima oleh Allah Subhanahu wa Ta'ala.

Jember, Juli 2026

Septa Daeng Indar Kurniawan

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN KEASLIAN	iv
MOTTO	v
LEMBAR PERSEMBAHAN	vi
UCAPAN TERIMA KASIH	vii
KATA PENGANTAR	ix
ABSTRAK	x
DAFTAR ISI	xii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvi
DAFTAR ISTILAH	xviii
BAB I: PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah	4
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian	5
BAB II: LANDASAN TEORI	6
2.1. Analisis Sentimen	6
2.1.1. Tujuan Analisis sentimen	6
2.1.2. Jenis-Jenis Sentimen	7
2.2. <i>Game</i> Bus Simulator Indonesia (BUSSID)	8
2.3. <i>Text Mining</i>	9
2.4. <i>Data Preprocessing</i>	9
2.5. Google Play Store	10
2.6. Penyeimbangan Data	10
2.7. <i>Hyper Sampling</i>	11
2.8. Klasifikasi	11
2.9. Hold-Out	12
2.10. TF-IDF	12
2.11. <i>Naïve Bayes</i>	14
2.12. Klasifikasi <i>Naïve Bayes</i>	16

2.13. <i>Confusion Matrix</i>	17
2.14. Penelitian Terdahulu	19
BAB III: METODOLOGI	22
3.1. Alur Penelitian	22
3.2. Studi Literatur	23
3.3. Pengumpulan Data	23
3.4. <i>Data Preprocessing</i>	23
3.5. Kombinasi TF-IDF dan <i>Naïve Bayes</i>	49
3.6. <i>Confusion Matrix</i>	58
BAB IV: HASIL DAN PEMBAHASAN	69
4.1. Data Ulasan Bus Simulator Indonesia	69
4.2. <i>Preprocessing</i> Data Ulasan	70
4.2.1. <i>Preprocessing: Data Cleaning</i>	71
4.2.2. <i>Preprocessing: Case Folding</i>	72
4.2.3. <i>Preprocessing: Normalisasi Data</i>	74
4.2.4. <i>Preprocessing: Tokenizing</i>	75
4.2.5. <i>Preprocessing: Stemming</i>	76
4.3. Rekapitulasi Perhitungan TF-IDF	78
4.4. Klasifikasi Data Uji	81
4.5. Evaluasi Kinerja Model	89
4.6. Penyeimbangan Data.	101
4.7. Performa Model Pasca Penyeimbangan Data	102
BAB V: PENUTUP	104
5.1. Kesimpulan	104
5.2. Saran	105
DAFTAR PUSTAKA	107
LAMPIRAN	112
1. Data Ulasan	112
2. <i>Data Preprocessing</i>	114
3. <i>Labeling</i> Data	116
4. Data Nilai <i>Term Frequency</i>	118
5. Data Nilai <i>Inverse Document Frequency</i>	120
6. Data Nilai <i>Inverse Document Frequency Smoothing</i>	121
7. Data Nilai <i>Term Frequency-Inverse Document Frequency</i>	122
8.1. Data Nilai <i>Likelihood Smoothing</i> Skenario 1	123
8.2. Data Nilai <i>Likelihood Smoothing</i> Skenario 2	1243
8.3. Data Nilai <i>Likelihood Smoothing</i> Skenario 3	125

9.1. Data Nilai Log <i>Likelihood</i> Skenario 1	126
9.2. Data Nilai Log <i>Likelihood</i> Skenario 2	127
9.3. Data Nilai Log <i>Likelihood</i> Skenario 3	128
10.1. Data Uji Klasifikasi Dokumen Skenario 1	129
10.2. Data Uji Klasifikasi Dokumen Skenario 2	130
10.3. Data Uji Klasifikasi Dokumen Skenario 3	131
11. Uji Coba Penyeimbang Data Dokumen Skenario 1	132
12. Uji Coba Penyeimbang Data Dokumen Skenario 2	133
13. Uji Coba Penyeimbang Data Dokumen Skenario 3	134
ARTIKEL ILMIAH	135
PROFIL PENULIS	141



DAFTAR LAMPIRAN

LAMPIRAN	112
1. Data Ulasan	112
2. <i>Data Preprocessing</i>	114
3. <i>Labeling Data</i>	116
4. Data Nilai <i>Term Frequency</i>	118
5. Data Nilai <i>Inverse Document Frequency</i>	120
6. Data Nilai <i>Inverse Document Frequency Smoothing</i>	121
7. Data Nilai <i>Term Frequency-Inverse Document Frequency</i>	122
8.1. Data Nilai <i>Likelihood Smoothing</i> Skenario 1	123
8.2. Data Nilai <i>Likelihood Smoothing</i> Skenario 2	1243
8.3. Data Nilai <i>Likelihood Smoothing</i> Skenario 3	125
9.1. Data Nilai Log <i>Likelihood</i> Skenario 1	126
9.2. Data Nilai Log <i>Likelihood</i> Skenario 2	127
9.3. Data Nilai Log <i>Likelihood</i> Skenario 3	128
10.1. Data Uji Klasifikasi Dokumen Skenario 1	129
10.2. Data Uji Klasifikasi Dokumen Skenario 2	130
10.3. Data Uji Klasifikasi Dokumen Skenario 3	131
11. Uji Coba Penyeimbang Data Dokumen Skenario 1	132
12. Uji Coba Penyeimbang Data Dokumen Skenario 2	133
13. Uji Coba Penyeimbang Data Dokumen Skenario 3	134