

PERBANDINGAN ALGORITMA NAÏVE BAYES DAN GRADIENT BOOSTING PADA ANALISIS SENTIMEN APLIKASI MOODLE

SKRIPSI SARJANA SISTEM INFORMASI



Oleh

Irsyad Humam Hariz

207006516021

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS NASIONAL**

Jl. Sawo Manila No.61, RT.14/RW.7, Pejaten Barat, Kec. Ps. Minggu, Kota Jakarta Selatan,
Daerah Khusus Ibukota Jakarta 12520

PERBANDINGAN ALGORITMA NAÏVE BAYES DAN GRADIENT BOOSTING PADA ANALISIS SENTIMEN APLIKASI MOODLE

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar

Sarjana Sistem Informasi dari Fakultas Teknologi Komunikasi dan Informatika

Oleh

Irsyad Humam Hariz

207006516021



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
UNIVERSITAS NASIONAL**

2023/2024

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

TUGAS AKHIR

**PERBANDINGAN ALGORITMA NAÏVE BAYES DAN GRADIENT
BOOSTING PADA ANALISIS SENTIMEN APLIKASI MOODLE**



Irsyad Humam Hariz

207006516021

Dosen Pembimbing

A handwritten signature in black ink, appearing to be 'Andrianingsih', is written over the bottom part of the logo.

Dr. Andrianingsih S.Kom., MMSI

NIDN. 0303097902

HALAMAN KEASLIAN TUGAS AKHIR

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

PERBANDINGAN ALGORITMA NAÏVE BAYES DAN GRADIENT BOOSTING PADA ANALISIS SENTIMEN APLIKASI MOODLE

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber Informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, Agustus 2024



Irsyad Humam Hariz

NPM : 207006516021

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Bilamana di kemudian hari ditemukan bahwasanya karya tulis ini menyalahi peraturan yang ada berkaitan etika serta kaidah penulisan karya ilmiah yang berlaku, maka saya bersedia dituntut serta diproses sesuai dengan ketentuan yang berlaku.

Yang menyatakan,

Nama : Irsyad Humam Hariz

NPM : 207006516021

Tanda Tangan :

(Irsyad Humam Hariz)

Tanggal : 26 Agustus 2024

Mengetahui

Pembimbing I : Dr. Andrianingsih S.Kom., MMSI

A

26/8/2024



LEMBAR PERSETUJUAN REVIEW AKHIR

LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

PERBANDINGAN ALGORITMA NAÏVE BAYES DAN GRADIENT BOOSTING PADA ANALISIS SENTIMEN APLIKASI MOODLE

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Akhir Semester Genap 2023-2024 pada tanggal 22 Agustus 2024



Dosen Pembimbing

Dr. Andrianingsih, S. Kom., MMSI.

NIDN. 0303097902

Ketua Program Studi

The image shows an official blue circular stamp. The outer ring contains the text 'UNIVERSITAS NASIONAL' at the top and 'FAKULTAS TEKNOLOGI KOMUNIKASI DAN INFORMATIKA' at the bottom. Inside the ring, there is a smaller emblem and the text 'PROGRAM STUDI SISTEM INFORMASI'. A handwritten signature is written across the stamp.

Dr. Andrianingsih, S. Kom., MMSI.

NIDN. 0303097902

HALAMAN PENGESAHAN
TUGAS SARJANA
PERBANDINGAN ALGORITMA NAÏVE BAYES DAN
GRADIENT BOOSTING PADA ANALISIS SENTIMEN
APLIKASI MOODLE



Penguji 1

[Signature]
Dr. Fauziah S.Kom, MMSI
NIDN. 0304107503

Penguji 2

[Signature]
Winarsih, S.Si, MMSI
NIDN. 0310057602

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Irsyad Humam Hariz
NPM : 207006516021
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
Program Studi : Sistem Informasi
Tanggal Sidang : 22 Agustus 2024

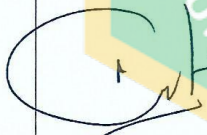


JUDUL DALAM BAHASA INDONESIA :

PERBANDINGAN ALGORITMA NAÏVE BAYES DAN GRADIENT BOOSTING PADA ANALISIS SENTIMEN APLIKASI MOODLE

JUDUL DALAM BAHASA INGGRIS :

COMPARISON OF NAÏVE BAYES AND GRADIENT BOOSTING ALGORITHM IN SENTIMENT ANALYSIS OF MOODLE APPLICATIONS

TANDA TANGAN DAN TANGGAL

Pembimbing	Ka. Prodi	Mahasiswa
TGL: 24.08.2024	TGL: 22.08.2024	TGL: 24.08.2024
		

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Irsyad Humam Hariz
NPM : 207006516021
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
Program Studi : Sistem Informasi
Tanggal Sidang : 22 Agustus 2024

JUDUL DALAM BAHASA INDONESIA :

PERBANDINGAN ALGORITMA NAÏVE BAYES DAN GRADIENT BOOSTING PADA ANALISIS SENTIMEN APLIKASI MOODLE

JUDUL DALAM BAHASA INGGRIS :

COMPARISON OF NAÏVE BAYES AND GRADIENT BOOSTING ALGORITHM IN SENTIMENT ANALYSIS OF MOODLE APPLICATIONS

TANDA TANGAN DAN TANGGAL

Pembimbing	Ka. Prodi	Mahasiswa
TGL: 24.08.2024	TGL: 22.08.2024	TGL: 24.08.2024
		

KATA PENGANTAR

Puji serta syukur peneliti panjatkan kepada Allah SWT atas segala rahmat serta karunia yang telah diberikan sehingga peneliti dapat menyelesaikan skripsi dengan judul “Perbandingan Algoritma Naïve Bayes serta Gradient Boosting Pada Analisis Sentimen Aplikasi Moodle” sebagai bagian dari persyaratan untuk menyelesaikan Program Sarjana Jurusan Sistem Informasi Fakultas Teknologi Komunikasi serta Informatika.

Peneliti menyampaikan rasa terima kasih yang mendalam kepada semua pihak yang sudah mendedikasikan dukungan dalam penelitian serta penyusunan skripsi ini. Khususnya kepada dosen pembimbing Tugas Akhir, Ibu Dr. Andrianingsih, S.Kom., MMSI, atas dedikasi pikiran, tenaga, serta waktu, serta bimbingan, arahan, serta motivasi yang diberikan, serta juga pengertian terhadap segala keterbatasan peneliti selama proses penelitian serta penulisan skripsi. Terima kasih juga haturkan kepada:

1. Allah SWT yang telah memberikan nikmat Panjang umur serta Kesehatansehingga diberikan kesempatan untuk menimba ilmu di Universitas Nasional.
2. Ayah serta Ibu selaku orang tua peneliti yang telah memberikan dukungan dalam berbagai bentuk yang tak terhitung.
3. Seluruh dosen pengajar di Program Studi Sistem Informasi FTKI maupun dosen di Program Studi lain yang telah memberikan banyak ilmu.
4. Ibu Dr. Andrianingsih, S.Kom., MMSI., selaku dosen pembimbing 1.
5. Kawan seangkatan serta sehimpuan bermacam angkatan yang sudah memberikan bantuan serta dukungan.

Akhir kata, semoga Allah SWT membalas segala kebaikan serta bantuan yang telah diberikan dengan hal yang lebih baik. Meskipun skripsi telah selesai iv dibuat, akan tetapi skripsi ini masih jauh dari kata sempurna. Oleh sebab itu, peneliti sangat mengharapkan arahan baik berupa saran maupun kritik dari pembaca.

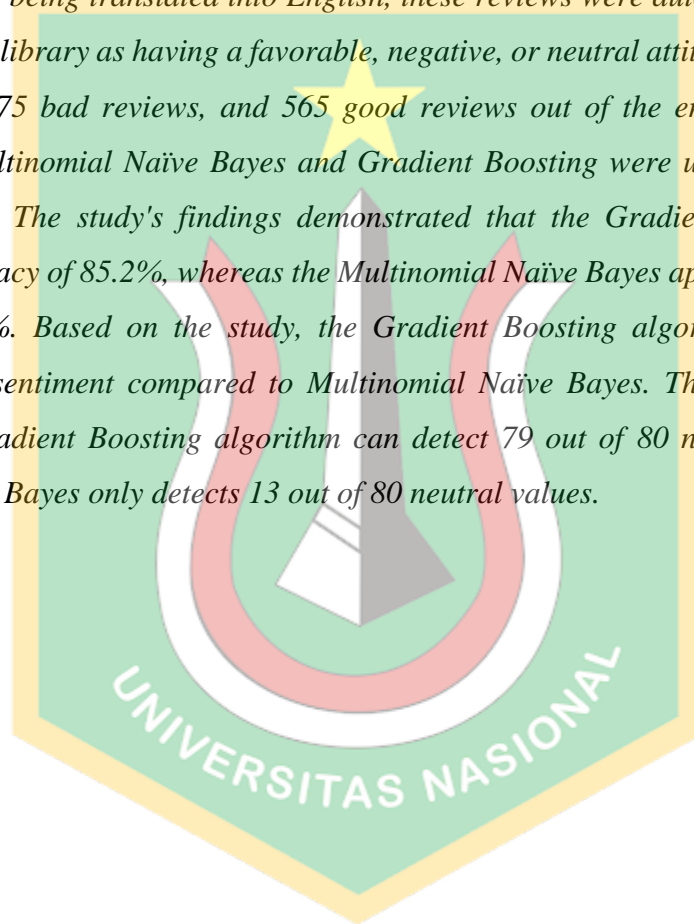
Jakarta, 15 Agustus 2024



Irsyad Humam Hariz

ABSTRACT

There are concerns over the applicability of the star rating when it comes to e-learning apps on the Google Play Store because they are downloaded with a similar amount of stars. Moodle is an e-learning program with 153 thousand ratings and a star rating of 4.4 out of 5. Applications undoubtedly have benefits and drawbacks of their own, which prompt users to respond and share their opinions in the review pool. This study used the Gradient Boosting and Naïve Bayes algorithms to assess reviews that were found on the Moodle program. After preprocessing 1592 review data from the Moodle application, 1346 data were obtained for this investigation. After being translated into English, these reviews were automatically classified using the TextBlob library as having a favorable, negative, or neutral attitude. There were 406 neutral reviews, 375 bad reviews, and 565 good reviews out of the entire data. Next, the algorithms for Multinomial Naïve Bayes and Gradient Boosting were utilized to categorize sentiment reviews. The study's findings demonstrated that the Gradient Boosting method produced an accuracy of 85.2%, whereas the Multinomial Naïve Bayes approach produced an accuracy of 62.6%. Based on the study, the Gradient Boosting algorithm is superior in detecting neutral sentiment compared to Multinomial Naïve Bayes. The Confusion Matrix shows that the Gradient Boosting algorithm can detect 79 out of 80 neutral values, while Multinomial Naïve Bayes only detects 13 out of 80 neutral values.



ABSTRAK

Aplikasi *e-learning* yang ada di Google Play Store mendapatkan jumlah unduhan pemberian bintang yang relatif sama, ini menimbulkan pertanyaan mengenai relevansi dari nilai bintang tersebut. Moodle merupakan aplikasi e-learning yang mempunyai nilai bintang 4,4 dari 5 bintang dengan 153 ribu ulasan. Suatu aplikasi tentunya mempunyai kekurangan serta kelebihan tersendiri yang mengakibatkan responsi dari pengguna yang kemudian akan diekspresikan ke kolam ulasan. Penelitian ini menganalisis ulasan yang terdapat pada aplikasi Moodle menggunakan algoritma Naïve Bayes serta Gradient Boosting. Penelitian ini menghasilkan sebanyak 1592 data ulasan diambil dari aplikasi Moodle serta kemudian dipreproses menjadi 1346 data. Ulasan-ulasan tersebut secara otomatis diberikan label sentimen netral, negatif, maupun positif memakai *library TextBlob* setelah diterjemahkan ke dalam Bahasa Inggris. Dari total data, terdapat 565 ulasan positif, 375 ulasan negatif, serta 406 ulasan netral. Algoritma Multinomial Naïve Bayes serta Gradient Boosting akan diterapkan untuk mengklasifikasikan sentimen ulasan. Perolehan studi mengindikasikan algoritma *Multinomial Naïve Bayes* menggapai akurasi 62,6%, sementara algoritma Gradient Boosting mencapai akurasi sebesar 85,2%. Berdasarkan penelitian, algoritma Gradient Boosting lebih unggul dalam mendeteksi sentimen netral dibandingkan dengan Multinomial Naïve Bayes. Confusion Matrix mengindikasikan bahwasanya algoritma Gradient Boosting dapat mendeteksi 79 nilai dari 80 nilai netral, sementara itu Multinomial Naïve Bayes hanya mendeteksi 13 dari 80 nilai netral.



DAFTAR ISI

BAB I	1
PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian.....	3
1.4. Batasan Masalah.....	3
1.5. Manfaat Penelitian.....	4
BAB II	5
TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Landasan Teori.....	9
2.2.1 Analisis Sentimen	9
2.2.2 <i>Natural Language Processing</i>	9
2.2.3 <i>Moodle</i>	10
2.2.4 <i>TextBlob</i>	10
2.2.5 Naïve Bayes.....	10
2.2.6 Multinomial Naïve Bayes.....	11
2.2.7 Gradient Boosting.....	11
2.2.8 Text Mining	12
2.2.9 Text Preprocessing	12
2.2.10 TF-IDF.....	13
2.2.11 <i>Confusion Matrix</i>	14
BAB III	16
METODOLOGI PENELITIAN	16
3.1 Tahapan Penelitian	16
3.2 Pengumpulan data	16

3.3 Text Preprocessing	17
3.4 Penerjemahan Teks	18
3.5 Labeling.....	18
3.6 Pembobotan TF-IDF	18
3.7 Klasifikasi <i>Naïve Bayes</i>	19
3.8 Klasifikasi <i>Gradient Boosting</i>	21
BAB IV	23
HASIL serta PEMBAHASAN	23
4.1 Pengumpulan data	23
4.2 Preprocessing	23
4.2.1 Data Cleaning	24
4.2.2 Normalisasi.....	25
4.2.3 Stopword Removal	26
4.2.4 Tokenisasi.....	27
4.2.5 Stemming.....	27
4.3 Penerjemahan Teks	28
4.4 Labeling.....	29
4.5 Klasifikasi <i>Naïve Bayes</i>	30
4.6 Klasifikasi <i>Gradient Boosting</i>	32
4.7 Evaluasi	33
BAB V	40
KESIMPULAN serta SARAN	40
5.1 Kesimpulan.....	40
5.2 Saran.....	40



DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian.....	16
Gambar 3.2 Contoh Ulasan.....	17
Gambar 3.3 Contoh Data Cleaning.....	17
Gambar 3.4 Contoh Normalisasi.....	17
Gambar 3.5 Contoh Stopword Removal.....	17
Gambar 3.6 Contoh Tokenization.....	17
Gambar 3.7 Contoh Stemming.....	18
Gambar 4.1 Proses Data Scraping.....	23
Gambar 4.2 Hasil Data Scraping.....	23
Gambar 4.3 Tahapan Preprocessing.....	24
Gambar 4.4 Proses Data Cleaning.....	24
Gambar 4.5 Hasil Case Folding.....	25
Gambar 4.6 Proses Normalisasi.....	25
Gambar 4.7 Hasil Normalisasi.....	26
Gambar 4.8 Proses Stopword Removal.....	26
Gambar 4.9 Hasil Stopword Removal.....	26
Gambar 4.10 Proses Tokenisasi.....	27
Gambar 4.11 Hasil Tokenisasi.....	27
Gambar 4.12 Proses Stemming.....	28
Gambar 4.13 Hasil Stemming.....	28
Gambar 4.14 Proses Penerjemahan Teks.....	28
Gambar 4.15 Hasil Penerjemahan Teks.....	29
Gambar 4.16 Labeling Text dengan Textblob.....	29
Gambar 4.17 Hasil Labeling Text.....	30
Gambar 4.18 Model MultinomialNB.....	30
Gambar 4.19 Confusion Matrix MultinomialNB.....	31

Gambar 4.20 Model Gradient Boosting.....	32
Gambar 4.21 Confusion Matrix Gradient Boosting.....	33
Gambar 4.22 Wordcloud Data Penelitian	34
Gambar 4.23 Diagram Batang Klasifikasi Sentimen.....	35
Gambar 4.24 Confusion Matrix MultinomialNB.....	36
Gambar 4.25 Confusion Matrix Gradient Boosting.....	38



DAFTAR TABEL

Tabel 2.1 Studi Literatur	9
Tabel 2.2 Confusion Matrix	14
Tabel 3.1 Sampel Data Latih	19
Tabel 3.2 Sampel Data Uji.....	19

