

**ANALISIS SENTIMEN *REVIEW* APLIKASI SPOTIFY UNTUK
REKOMENDASI PENINGKATAN LAYANAN
MENGUNAKAN METODE *SUPPORT VECTOR MACHINE*
DAN *NAÏVE BAYES***



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

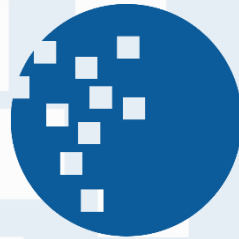
Kumala Dewi Chandra Asih

00000043238

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2024

***SENTIMENT ANALYSIS OF SPOTIFY REVIEW FOR SERVICE
IMPROVEMENT RECOMMENDATIONS USING SUPPORT
VECTOR MACHINE AND NAÏVE BAYES METHOD***



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Komputer (S.Kom.)

Kumala Dewi Chandra Asih

00000043238

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNIK DAN INFORMATIKA

UNIVERSITAS MULTIMEDIA NUSANTARA

TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Kumala Dewi Chandra Asih

Nomor Induk Mahasiswa : 00000043238

Program studi : Sistem Informasi

Skripsi dengan judul:

ANALISIS SENTIMEN REVIEW APLIKASI SPOTIFY UNTUK REKOMENDASI PENINGKATAN LAYANAN MENGGUNAKAN METODE SUPPORT VECTOR MACHINE DAN NAÏVE BAYES merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas Akhir yang telah saya tempuh.

Tangerang, 27 Mei 2024



(Kumala Dewi Chandra Asih)

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

HALAMAN PENGESAHAN

Skripsi dengan judul

Analisis Sentimen Review Aplikasi Spotify Untuk Rekomendasi Peningkatan Layanan Menggunakan Metode *Support Vector Machine* Dan *Naïve Bayes*


Oleh

Nama : Kumala Dewi Chandra Asih
NIM : 00000043238
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

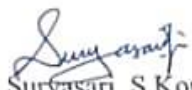
Telah diujikan pada hari Senin, 27 Mei 2024
Pukul 13.00 s.d 15.00 dan dinyatakan
LULUS

Dengan susunan penguji sebagai berikut.


Ketua Sidang


Ririn Ikana Desanti, S.KOM., M.Kom.
0313058001


Penguji


Suryasari, S.Kom., M.T.
0323088301

Pembimbing


Monika Evelin Jolhan, S.Kom., M.M.S.I.
0327059501

Ketua Program Studi Sistem Informasi


Ririn Ikana Desanti, S.KOM., M.Kom.

LEMBAR PERSETUJUAN PUBLIKASI
KARYA ILMIAH MAHASISWA

Yang bertanda tangan dibawah ini:

Nama : Kumala Dewi Chandra Asih
Nomor Induk Mahasiswa : 00000043238
Program Studi : Sistem Informasi
Jenjang : S1
Judul Karya Ilmiah : Analisis Sentimen Review Aplikasi Spotify Untuk
Rekomendasi Peningkatan Layanan Menggunakan
Metode *Support Vector Machine* Dan *Naïve Bayes*

Menyatakan dengan sesungguhnya bahwa saya bersedia:

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial. Saya tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun. Saya tidak bersedia, dikarenakan:
- Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: Dalam proses pengajuan penerbitan ke dalam jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)*.

Tangerang, 27 Mei 2024



(Kumala Dewi Chandra Asih)

KATA PENGANTAR

Dengan penuh rasa syukur, saya ingin mengucapkan terima kasih kepada Tuhan Yesus Kristus karena telah memberikan berkat dan karunia-Nya, yang memungkinkan saya menyelesaikan penulisan Skripsi dengan judul: "Analisis Sentimen Review Aplikasi Spotify Untuk Rekomendasi Peningkatan Layanan Menggunakan Metode Support Vector Machine Dan Naïve Bayes". Penulisan ini merupakan syarat penting dalam memenuhi gelar S1 di Jurusan Sistem Informasi, Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari betapa sulitnya menyelesaikan tugas akhir ini tanpa dukungan dan bimbingan dari berbagai pihak. Oleh karena itu, saya ingin menyampaikan rasa terima kasih kepada:

1. Bapak Dr. Ninok Leksono, M.A., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom, selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Monika Evelin Johan, S.Kom., M.M.S.I, sebagai Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga karya ilmiah ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 15 Mei 2024


(Kumala Dewi Chandra Asih)

**ANALISIS SENTIMEN *REVIEW* APLIKASI SPOTIFY UNTUK
REKOMENDASI PENINGKATAN LAYANAN
MENGUNAKAN METODE *SUPPORT VECTOR MACHINE*
DAN *NAÏVE BAYES***

Kumala Dewi Chandra Asih

ABSTRAK

Perkembangan teknologi telah membawa perubahan mendasar dalam cara masyarakat menikmati musik, terutama melalui kemunculan aplikasi *streaming* seperti Spotify yang telah meraih popularitas luar biasa di kalangan generasi milenial. Peningkatan pengguna Spotify mempengaruhi jumlah ulasan, baik yang bersifat positif maupun negatif. Ulasan pengguna berguna untuk memahami kepuasan pengguna terhadap layanan Spotify dan meningkatkan kualitas layanan dengan merespons masukan pengguna. Penelitian ini bertujuan untuk melakukan perbandingan metode antara metode *Support Vector Machine* dan *Naive Bayes* dalam analisis sentimen terhadap ulasan aplikasi Spotify. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna Spotify yang diambil dari Google Play Store, terdiri dari 2500 data. Ulasan dianalisis dari sentimen negatif. Perbandingan dilakukan antara dua model algoritma, yaitu *Support Vector Machine* dan *Naive Bayes*. SVM mencari hyperplane terbaik untuk memisahkan data ke dalam kelas-kelas. *Naive Bayes* mengasumsikan independensi antar fitur dan digunakan untuk analisis teks dan sentimen. Berdasarkan pengujian dan evaluasi yang dilakukan, ditemukan bahwa algoritma *Naive Bayes* memiliki akurasi yang lebih tinggi dibandingkan dengan *Support Vector Machine*. *Naive Bayes* memiliki Tingkat akurasi 82.40% sedangkan algoritma *Support Vector Machine* memiliki Tingkat akurasi sebesar 79.60%. Oleh karena itu, penggunaan algoritma *Naive Bayes* dapat diterapkan dalam analisis sentimen untuk mendukung peningkatan layanan aplikasi Spotify di masa mendatang. Hasil analisis ulasan tersebut akan digunakan untuk peningkatan kualitas aplikasi Spotify ke depannya agar terus memberikan kualitas terbaik bagi para pengguna

Kata kunci: Analisis Sentimen, *Naive Bayes*, Spotify, *Support Vector Machine*, *Streaming music*

SENTIMENT ANALYSIS OF SPOTIFY REVIEW FOR SERVICE IMPROVEMENT RECOMMENDATIONS USING SUPPORT VECTOR MACHINE AND NAIVE BAYES METHOD

Kumala Dewi Chandra Asih

ABSTRACT (English)

The development of technology has brought about fundamental changes in the way people enjoy music, especially through the emergence of streaming applications such as Spotify, which has gained tremendous popularity among the millennial generation. The increase in Spotify users affects the number of reviews, both positive and negative. User reviews are useful for understanding user satisfaction with the Spotify service and improving service quality by responding to user feedback. This study aims to compare methods between Support Vector Machine and Naive Bayes in sentiment analysis of Spotify application reviews. This research aims to analyze the sentiment of user reviews of Spotify taken from the Google Play Store, consisting of 2500 data points. The reviews are analyzed for negative sentiment. A comparison is made between two algorithmic models, namely Support Vector Machine and Naive Bayes. SVM seeks the best hyperplane to separate data into classes, while Naive Bayes assumes independence between features and is used for text and sentiment analysis. Based on the testing and evaluation conducted, it was found that the Naive Bayes algorithm has a higher accuracy compared to Support Vector Machine. Naive Bayes has an accuracy rate of 82.40%, while the Support Vector Machine algorithm has an accuracy rate of 79.60%. Therefore, the use of the Naive Bayes algorithm can be applied in sentiment analysis to support the improvement of Spotify application services in the future. The results of the review analysis will be used to improve the quality of the Spotify application in the future to continue providing the best quality for users.

Keywords: *Sentiment Analysis, Naive Bayes, Spotify, Support Vector Machine, Streaming music*

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PENGESAHAN.....	iii
LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA	iv
KATA PENGANTAR.....	v
ABSTRAK	vi
<i>ABSTRACT (English)</i>	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR RUMUS	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah	5
1.4 Tujuan dan Manfaat Penelitian.....	5
1.4.1 Tujuan Penelitian.....	5
1.4.2 Manfaat Penelitian.....	6
1.5 Sistematika Penulisan	6
BAB II LANDASAN TEORI	8
2.1 Penelitian Terdahulu.....	8
2.2 Teori tentang Topik Skripsi	13
2.2.1 Spotify	13
2.2.2 Analisis Sentimen.....	13
2.2.3 <i>CountVectorizer</i>	14
2.2.4 <i>Confusion Matrix</i>	15
2.3. Teori tentang <i>Framework</i> / Algoritma yang digunakan	17
2.3.1 <i>CRISP-DM</i>	17
2.3.2 <i>Naïve Bayes</i>	19

2.3.3	<i>Support Vector Machine</i>	20
2.3.4	<i>Text Mining</i>	21
2.4	Teori tentang <i>Tools / Software</i> yang digunakan	23
2.4.1	Google Play Store.....	23
2.4.2	Python	24
2.4.3	Google Colab	25
BAB III	METODOLOGI PENELITIAN	27
3.1	Gambaran Umum Objek Penelitian	27
3.2	Metode Penelitian	27
3.3	Variabel Penelitian	28
3.3.1	Variabel Independen	28
3.3.2	Variabel Dependen	28
3.4	Teknik Pengumpulan Data	29
3.5	Teknik Pengambilan Sampel	29
3.6	Teknik Analisis Data	29
BAB IV	ANALISIS DAN HASIL PENELITIAN	35
4.1	<i>Business Understanding</i>	35
4.2	<i>Data Understanding</i>	36
4.3	<i>Data Preparation</i>	38
4.3.1	<i>Case Folding dan Cleaning</i>	41
4.3.2	<i>Tokenization</i>	42
4.3.3	<i>Stopword Removal</i>	43
4.3.4	<i>Stemming</i>	45
4.3.5	<i>Labeling</i>	46
4.3.6	<i>Splitting Data</i>	49
4.3.7	<i>CountVectorizer</i>	51
4.3.7	<i>Wordcloud Analysis</i>	52
4.4	<i>Modeling</i>	55
4.4.1	<i>Support Vector Machine</i>	55
4.4.2	<i>Naïve Bayes</i>	56
4.5	<i>Evaluation</i>	57
4.5.1	<i>Support Vector Machine</i>	58

4.5.2 <i>Naïve Bayes</i>	61
4.6 <i>Deployment</i>	64
4.7 Analisis Hasil dan Diskusi	66
BAB V SIMPULAN DAN SARAN	69
5.1 Simpulan	69
5.2 Saran.....	70
DAFTAR PUSTAKA	72
LAMPIRAN	76



DAFTAR TABEL

Tabel 2 1 <i>Confusion Matrix</i>	15
Tabel 3 1 Perbandingan <i>Support Vector Machine</i> dan <i>Naïve Bayes</i>	28
Tabel 3 2 Perbandingan <i>Framework Data Mining</i>	29
Tabel 4 1 Variabel pada <i>dataset</i>	37
Tabel 4 2 Hasil Case folding dan Cleaning data.....	41
Tabel 4 3 Hasil <i>Tokenization</i>	43
Tabel 4 4 Hasil <i>Stopwords Removal</i>	45
Tabel 4 5 Hasil <i>Stemming</i>	46
Tabel 4 6 Hasil Data Labelling	48
Tabel 4 7 Hasil Performa Model SVM dan <i>Naïve Bayes</i>	66
Tabel 4 8 Perbandingan Penelitian Terdahulu	67



DAFTAR GAMBAR

Gambar 1.1 Pelanggan Spotify <i>Premium</i> Kuartal II 2023	3
Gambar 2 2 Tahapan <i>CRISP-DM</i>	17
Gambar 2 3 Hyperlane SVM	20
Gambar 3 1 Alur Penelitian	31
Gambar 3 2 Tahapan <i>Data preprocessing</i>	33
Gambar 4. 1 Scraping data	36
Gambar 4. 2 Mengganti nama variabel data	37
Gambar 4. 3 Data Spotify	38
Gambar 4. 4 Membaca data file excel.....	38
Gambar 4. 5 Info data Spotify.....	39
Gambar 4. 6 Mengecek missing values	39
Gambar 4. 7 Mengecek duplicate values	40
Gambar 4. 8 Mengecek data shape	40
Gambar 4. 9 Kode Tahapan Case folding dan Cleaning data	41
Gambar 4. 10 Tahapan Tokenization	42
Gambar 4. 11 Tahapan Import library nltk stopwords.....	43
Gambar 4. 12 Tahapan Stopwords	44
Gambar 4. 13 Tahapan Stemming.....	45
Gambar 4. 14 Install transformers.....	46
Gambar 4. 15 Install Googletrans	46
Gambar 4. 16 Import Translator dan Pipeline.....	47
Gambar 4. 17 Membersihkan teks Spotify.....	47
Gambar 4. 18 Proses penerjemahan label	47
Gambar 4. 19 Analisis sentimen teks menggunakan sentimen classifier	48
Gambar 4. 20 contoh sample acak 10 baris Dataframe.....	49
Gambar 4. 21 Splitting Data	50
Gambar 4. 22 CountVectorizer	51
Gambar 4. 23 Vektorisasi vocabulary	51
Gambar 4. 24 Persentase Label Sentimen.....	52
Gambar 4. 25 Word Cloud Sentimen Negatif.....	53
Gambar 4. 26 Word Cloud Sentimen Positif	53
Gambar 4. 27 Frekuensi Lima Kata Positif Teratas.....	54
Gambar 4. 28 Frekuensi Lima Kata Negatif Teratas	55
Gambar 4. 29 Potongan Kode Support Vector Machine	56
Gambar 4. 30 Potongan Kode Naïve Bayes.....	57
Gambar 4. 31 Classification Report Support Vector Machine Test size 0.3	58
Gambar 4. 32 Classification Report Support Vector Machine Test size 0.2	58

Gambar 4. 33 Classification Report Support Vector Machine Test size 0.1	59
Gambar 4. 34 Confusion Matrix Support Vector Machine.....	60
Gambar 4. 35 Classification Report Naïve Bayes test size 0.3.....	61
Gambar 4. 36 Classification Report Naïve Bayes test size 0.2.....	61
Gambar 4. 37 Classification Report Naïve Bayes test size 01	62
Gambar 4. 38 Confusion Matrix Naïve Bayes	63
Gambar 4. 39 Perbandingan Akurasi SVM dan Naïve Bayes	64
Gambar 4. 40 Prediksi Kalimat Positif dengan Naive Bayes.....	65
Gambar 4. 41 Prediksi Kalimat Negatif dengan Naive Bayes	66



DAFTAR RUMUS

Rumus 2 1 Perhitungan Nilai <i>Accuracy</i>	16
Rumus 2 2 Perhitungan Nilai <i>Recall</i>	16
Rumus 2 3 Perhitungan Nilai <i>Precision</i>	16
Rumus 2 4 Perhitungan Nilai <i>F1-Score</i>	16
Rumus 2 5 <i>Naïve Bayes</i>	19



DAFTAR LAMPIRAN

Lampiran A Formulir Konsultasi Skripsi Pembimbing	76
Lampiran B Hasil Turnitin	77

