

**ANALISIS ULASAN PENGGUNA APLIKASI
SPOTIFY PADA GOOGLE PLAY MENGGUNAKAN
ALGORITMA RANDOM FOREST**

SKRIPSI

ROSMAWATI
20200040096



**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS TEKNIK, KOMPUTER DAN DESAIN
UNIVERSITAS NUSA PUTRA
SUKABUMI
JUNI 2024**

**ANALISIS ULASAN PENGGUNA APLIKASI
SPOTIFY PADA GOOGLE PLAY MENGGUNAKAN
ALGORITMA RANDOM FOREST**

SKRIPSI

*Diajukan Untuk Memenuhi Salah Satu Syarat Dalam Menempuh Gelar Sarjana
Teknik Informatika*

ROSMAWATI
20200040096



**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS TEKNIK, KOMPUTER DAN DESAIN
UNIVERSITAS NUSA PUTRA
SUKABUMI
JUNI 2024**

PERNYATAAN PENULIS

JUDUL : ANALISIS ULASAN PENGGUNA APLIKASI SPOTIFY PADA
GOOGLE PLAY MENGGUNAKAN ALGORITMA RANDOM
FOREST

NAMA : ROSMAWATI

NIM : 20200040096

“Saya menyatakan dan bertanggungjawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Teknik Informatika saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut”.

Sukabumi, 28 Juni 2024



ROSMAWATI

Penulis

PENGESAHAN SKRIPSI

JUDUL : ANALISIS ULASAN PENGGUNA APLIKASI SPOTIFY PADA
GOOGLE PLAY MENGGUNAKAN ALGORITMA RANDOM
FOREST

NAMA : ROSMAWATI

NIM : 20200040096

Skripsi ini telah diujikan dan dipertahankan di depan Dewan Penguji pada
Sidang Skripsi tanggal 19 Juni 2024. Menurut pandangan kami,
Skripsi ini memadai dari segi kualitas untuk tujuan penganugerahan
gelar Sarjana Komputer (S.Kom)

Sukabumi, Juni 2024

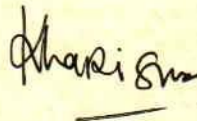
Pembimbing I



Gina Purnama Insany, S.Si.T., M.Kom

NIDN. 0417077908

Pembimbing II



Ivana Lucia Kharisma, M.Kom

NIDN. 0429038002

Ketua Penguji



Muhammad Ikhsan Thohir, M.Kom

NIDN. 0415049302

Ketua Program Studi



Ir. Sonantri S.T., M.Kom

NIDN. 0419128801

Plh. Dekan Fakultas Teknik, Komputer dan Desain

Ir. Paikun, S.T., M.T., IPM., ASEAN Eng.

NIDN. 0402037401

ABSTRACT

Technological developments have played a key role in the transformation of the music industry and driven the emergence of digital music streaming services. Spotify is one type of music service application, this application allows users to collect playlists of songs and store them neatly in one storage. With Google Play as the application distribution, Google Play also allows users to provide positive or negative reviews of the applications they use. The application of the Random Forest algorithm in research is expected to provide insight and information regarding user preferences for applications. From this background, this sentiment analysis research was carried out. Spotify app review data is collected from Google Play by applying web scrapping techniques. With 2000 review data, consisting of 1000 Indonesian language review data and 1000 English review data. From the modeling results using the Random Forest algorithm, the accuracy results for English language reviews were 93.6% and Indonesian language reviews were 88,4%. Modeling is carried out up to the deployment stage using streamlit, to share applications from research models that have been carried out into web-based applications.

Keywords : review, google play, spotify application, random forest, web scrapping.



ABSTRAK

Perkembangan teknologi berperan penting dalam transformasi industri musik dan mendorong munculnya layanan streaming musik digital. Spotify merupakan salah satu jenis aplikasi layanan musik, aplikasi ini memungkinkan penggunanya untuk mengumpulkan playlist lagu dan menyimpannya dengan rapi dalam satu penyimpanan. Dengan Google Play sebagai tempat distribusi aplikasi, Google Play juga memungkinkan pengguna untuk memberikan ulasan positif maupun negatif terhadap aplikasi yang mereka gunakan. Penerapan algoritma Random Forest dalam penelitian diharapkan dapat memberikan wawasan dan informasi mengenai preferensi pengguna terhadap aplikasi. Berdasarkan latar belakang tersebut, penelitian analisis sentimen ini dilakukan. Data ulasan aplikasi Spotify dikumpulkan dari Google Play dengan menerapkan teknik web scrapping. Dengan data review sebanyak 2000, terdiri dari 1000 data review berbahasa Indonesia dan 1000 data review berbahasa Inggris. Dari hasil pemodelan menggunakan algoritma Random Forest diperoleh hasil akurasi review bahasa Inggris sebesar 93,6% dan review bahasa Indonesia sebesar 88,4%. Pemodelan dilakukan hingga tahap deployment dengan menggunakan streamlit, untuk membagi aplikasi dari model penelitian yang telah dilakukan ke dalam aplikasi berbasis web.

Kata Kunci: ulasan, *google play*, aplikasi *spotify*, *random forest*, *web scrapping*.

KATA PENGANTAR

Puji syukur kami panjatkan ke hadirat Allah SWT, berkat rahmat dan karunia-Nya akhirnya skripsi "Analisis Ulasan Aplikasi Spotify Pada Google Play Menggunakan Algoritma *Random Forest*" dapat penulis selesaikan dengan baik.

Tujuan penulisan skripsi ini adalah memberikan gambaran terkait penelitian yang telah dilakukan sebelumnya. Skripsi ini juga merupakan tugas akhir untuk syarat memperoleh gelar Sarjana Komputer (S. Kom) pada program studi Teknik Informatika di Universitas Nusa Putra.

Sehubungan dengan itu penulis menyampaikan penghargaan dan ucapan terima kasih yang sebesar-besarnya kepada :

1. Bapak Rektor Universitas Nusa Putra Sukabumi, Bapak Dr. Kurniawan S.T, M.Si, M.M.
2. Wakil Rektor I Bidang Akademik Universitas Nusa Putra Sukabumi, Bapak Anggy Pradiftha Junfitharana, S.Pd., M.T.
3. Kepala Program Studi Teknik Informatika Universitas Nusa Putra Sukabumi, Bapak Ir. Somantri S.T, M.Kom.
4. Dosen Pembimbing I Universitas Nusa Putra Sukabumi, ibu Gina Purnama Insany, S.Si. T., M.Kom, yang telah meluangkan waktu untuk mengarahkan dan memberikan masukan serta saran untuk penyelesaian skripsi ini.
5. Dosen Pembimbing II Universitas Nusa Putra Sukabumi, ibu Ivana Lucia Kharisma, M.Kom, membantu dalam proses penyusunan skripsi.
6. Dosen Penguji, yang telah memberikan saran dan masukan untuk penyusunan skripsi ini menjadi lebih baik.
7. Para Dosen Program Studi Teknik Informatika Universitas Nusa Putra Sukabumi, yang telah membagikan ilmunya selama 4 tahun masa perkuliahan.
8. Orang tua dan keluarga, yang telah memberikan dukungan sehingga skripsi ini dapat terselesaikan.

Penulis menyadari bahwa skripisi ini masih jauh dari kesempurnaan, oleh karena itu kritik dan saran yang membangun dari berbagai pihak sangat kami harapkan demi perbaikan. Amin Yaa Rabbal 'Alamiin.

Sukabumi, 28 Juni 2024

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMI

Sebagai civitas akademik Universitas Nusa Putra saya yang
bertanda tangan dibawah ini :

Nama : Rosmawati

NIM : 20200040096

Program Studi: Teknik Informatika

Jenis karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada
Universitas Nusa Putra **Hak Bebas Royalti Noneksklusif** (*Non-exclusive
Royalty- Free Right*) atas karya ilmiah saya yang berjudul :

ANALISIS ULASAN PENGGUNA APLIKASI SPOTIFY PADA GOOGLE PLAY MENGGUNAKAN ALGORITMA *RANDOM FOREST*

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti
Noneksklusif ini Universitas Nusa Putra berhak menyimpan, mengalih
media/format-kan, mengelola dalam bentuk pangkalan data (database), merawat,
dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya
sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Sukabumi

Pada tanggal : 28 Juni 2024



ROSMAWATI

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
PERNYATAAN PENULIS	ii
PENGESAHAN SKRIPSI	iii
ABSTRACT	iv
ABSTRAK	v
KATA PENGANTAR	vi
PERNYATAAN PERSETUJUAN PUBLIKASI	vii
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
1.5.1 Bagi Mahasiswa.....	5
1.5.2 Bagi Masyarakat	5
1.6 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian Terkait.....	7
2.2 Landasan Teori	11
2.2.1 Analisis sentimen	11
2.2.2 Google Play	12
2.2.3 Aplikasi Spotify.....	12

2.2.4	<i>Machine Learning</i>	12
2.2.5	<i>Algoritma Random Forest</i>	13
2.2.6	<i>Text Mining</i>	15
2.2.7	<i>Scraping Data</i>	15
2.2.8	<i>Confusion Matrix</i>	15
2.2.9	<i>Python</i>	15
2.2.10	<i>Google Colab</i>	16
2.2.11	<i>Streamlit</i>	16
2.3	<i>Kerangka Pemikiran</i>	17
BAB III METODOLOGI PENELITIAN		19
3.1	<i>Tahapan Penelitian</i>	19
3.2	<i>Identifikasi masalah</i>	20
3.3	<i>Pengumpulan Data</i>	20
3.4	<i>Pelabelan data</i>	22
3.5	<i>Preprocessing</i>	22
3.6	<i>Pembobotan TF-IDF</i>	35
3.7	<i>Oversampling</i>	36
3.8	<i>Split Data</i>	36
3.9	<i>Pemodelan Algoritma</i>	36
3.10	<i>Evaluasi dan visualisasi</i>	36
3.11	<i>Model Deployment</i>	38
BAB IV HASIL DAN PEMBAHASAN		39
4.1	<i>Scraping Data</i>	39
4.2	<i>Pelabelan Data</i>	42
4.3	<i>Preprocessing</i>	44
4.4	<i>Pembobotan TF-IDF</i>	48
4.5	<i>Oversampling</i>	49
4.6	<i>Split data</i>	50

4.7	Pemodelan Algoritma	51
4.8	Evaluasi Hasil Analisis	51
4.9	Pembahasan Hasil Analisis.....	52
4.10	Visualisasi Hasil Analisis	55
4.11	<i>Model deployment</i>	56
BAB V	PENUTUP	58
5.1	Kesimpulan.....	58
5.2	Saran	58
DAFTAR PUSTAKA		60



DAFTAR TABEL

Tabel 2.1	Penelitian Terkait.....	7
Tabel 3.1	Tahap <i>Case Folding</i> (Indonesia).....	23
Tabel 3.2	Tahap <i>Case Folding</i> (Inggris)	24
Tabel 3.3	Tahap <i>Cleaning</i> (Indonesia)	25
Tabel 3.4	Tahap <i>Cleaning</i> (Inggris)	26
Tabel 3.5	Tahap <i>Tokenizing</i> (Indonesia).....	28
Tabel 3.6	Tahap <i>Tokenizing</i> (Inggris).....	29
Tabel 3.7	Tahap <i>Stopwords Removal</i> (Indonesia).....	31
Tabel 3.8	Tahap <i>Stopwords Removal</i> (Inggris).....	32
Tabel 3.9	Tahap <i>Stemming</i> (Indonesia)	34
Tabel 3.10	Tahap <i>Stemming</i> (Inggris)	34
Tabel 4.1	Jumlah Data Ulasan Indonesia.....	39
Tabel 4.2	Jumlah Data ulasan Inggris.....	39
Tabel 4.3	Dataset Hasil <i>Web Scraping</i> Ulasan Indonesia.....	39
Tabel 4.4	Dataset Hasil <i>Web Scraping</i> Ulasan Inggris.....	40
Tabel 4.5	Dataset Ulasan Hasil <i>Labeling</i>	41
Tabel 4.6	Dataset Hasil <i>Preprocessing</i> Ulasan Indonesia.....	45
Tabel 4.7	Dataset Hasil <i>Preprocessing</i> Ulasan Inggris.....	46
Tabel 4.8	Jumlah Data <i>Oversampling</i> (Indonesia).....	48
Tabel 4.9	Jumlah Data <i>Oversampling</i> (Inggris).....	49
Tabel 4.10	Hasil <i>confusion matrix</i>	54

DAFTAR GAMBAR

Gambar 1.1	Spotify di play store.....	1
Gambar 1.2	Pendapatan Spotify.....	2
Gambar 1.3	Ulasan pengguna di Google Play (Inggris).....	2
Gambar 1.4	Ulasan pengguna di Google Play (Indonesia).....	3
Gambar 2.1	Google Play	12
Gambar 2.2	Spotify.....	12
Gambar 2.3	<i>Random Forest structure</i> sumber (dqlab.id).....	13
Gambar 2.4	<i>Python</i>	16
Gambar 2.5	<i>Google Colab</i>	16
Gambar 2.6	Streamlit	16
Gambar 2.7	Kerangka Pemikiran.....	17
Gambar 3.1	Tahapan Penelitian.....	19
Gambar 3.2	Data Ulasan Bahasa Indonesia	20
Gambar 3.3	Data Ulasan Bahasa Inggris.....	21
Gambar 3.4	Proses Pengumpulan Data Ulasan.....	21
Gambar 3.5	Proses Pelabelan Data Ulasan (Indonesia)	22
Gambar 3.6	Proses Pelabelan Data Ulasan (Inggris).....	22
Gambar 3.7	Tahapan <i>Preprocessing</i>	23
Gambar 3.8	<i>Confusion Matrix</i>	37
Gambar 4.1	<i>Syntax Web Scraping</i>	39
Gambar 4.2	Hasil <i>Labeling</i> Sentimen (Indonesia).....	43
Gambar 4.3	Hasil <i>Labeling</i> Sentimen (Inggris).....	44
Gambar 4.4	<i>Library Preprocessing</i>).....	44
Gambar 4.5	Tahap <i>Case Folding</i>	45
Gambar 4.6	Tahap <i>Cleaning</i>	45
Gambar 4.7	Tahap <i>Tokenizing</i>	45
Gambar 4.8	Tahap <i>Stopwords Removal</i>	46
Gambar 4.9	Tahap <i>Stemming</i>	46
Gambar 4.10	Proses Pembobotan TF-IDF.....	49
Gambar 4.11	Hasil <i>Over Sampling</i> (Indonesia).....	49

Gambar 4.12 Hasil <i>Over Sampling</i> (Inggris).....	50
Gambar 4.13 Tahap <i>Split Data</i>	50
Gambar 4.14 <i>Syntax</i> Pelatihan.....	51
Gambar 4.15 Prediksi Data Tes.....	51
Gambar 4.16 <i>Confusion Matrix</i>	51
Gambar 4.17 Kalkulasi <i>Matrix</i>	52
Gambar 4.18 <i>Confusion Matrix</i> 70:30 (Indonesia).....	53
Gambar 4.19 <i>Confusion Matrix</i> 70:30 (Inggris).....	54
Gambar 4.20 <i>Wordcloud</i> Sentimen Negatif.....	55
Gambar 4.21 <i>Wordcloud</i> Sentimen Positif.....	56
Gambar 4.22 Tampilan <i>Home Streamlit</i>	56
Gambar 4.23 Prediksi Komentar Positif.....	57
Gambar 4.24 Prediksi Komentar Negatif.....	57



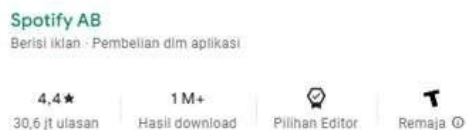
BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi, khususnya internet dan perangkat *mobile* telah memainkan peran kunci dalam transformasi industri musik dan mendorong munculnya layanan *streaming* musik. Dalam perkembangan tersebut toko aplikasi seperti *google play store* menjadi pusat utama bagi pengguna untuk mengunduh berbagai jenis aplikasi. Sebelum adanya layanan *streaming* musik seperti *spotify*, orang-orang biasanya mendengarkan musik melalui radio, membeli CD atau kaset musik, atau pun mengunduh lagu secara ilegal melalui internet. *Spotify* sebagai salah satu aplikasi *streaming* yang memungkinkan pengguna untuk mendengarkan musik, *podcast* atau pun konten audio lainnya dari seluruh dunia secara *online* [1].

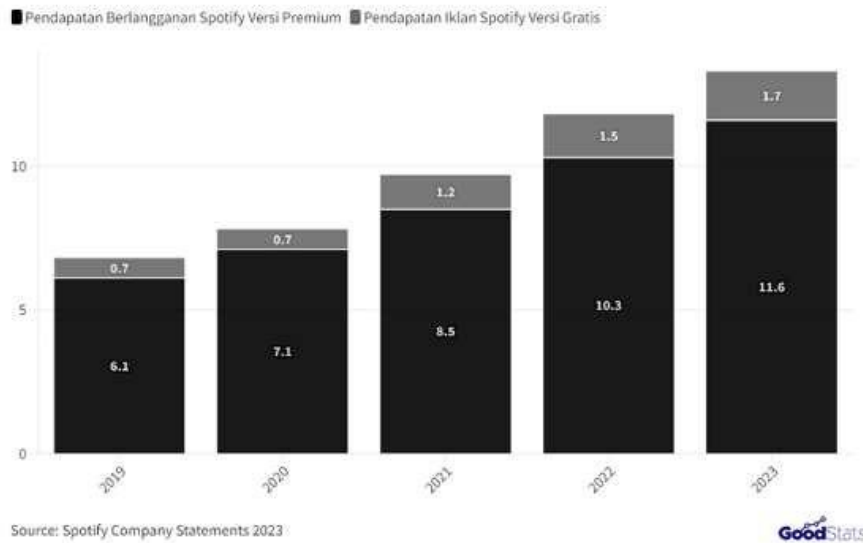
Spotify: Putar Musik & Podcast



Gambar 1.1 Spotify di Playstore

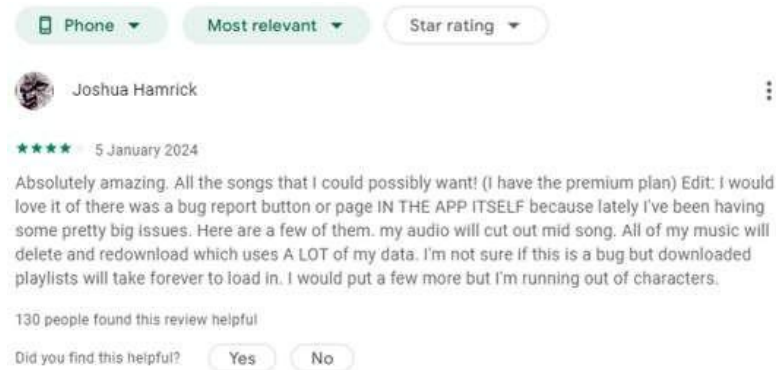
Pada Gambar 1.1 aplikasi ini mendapat penilaian rata-rata 4.4, menjadi salah satu aplikasi *streaming* musik yang banyak digunakan dan paling populer di dunia serta bisa diakses secara premium atau pun gratis. Peningkatan dari sisi pengguna, sejalan dengan pendapatan serta permintaan yang akan mendorong penyedia layanan untuk mengembangkan teknologi baru cepat dan efisien [2]. Maka pemahaman tentang preferensi pengguna menjadi kunci untuk meningkatkan kualitas dan daya saing suatu aplikasi.

Sumber Pendapatan Spotify Berdasarkan Tipe Berlangganan 5 Tahun Terakhir (dalam Miliar Euro)

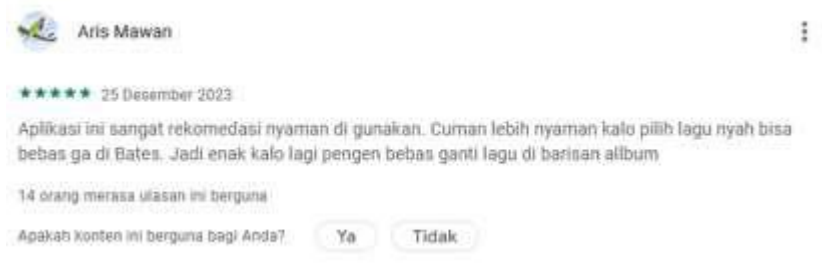


Gambar 1.2 Pendapatan Spotify

Spotify telah dirilis sejak tahun 2008 [3]. Berdasarkan pada Gambar 1.2 spotify mengalami kenaikan pendapatan dalam 5 tahun terakhir yang terhitung sejak 2019 hingga 2023. Menurut data yang dirilis spotify pada akhir tahun 2023 1,7 miliar euro atau 13% pendapatan berasal dari iklan yang ditayangkan pada aplikasi versi gratis dan 11,6 miliar euro dari pelanggan aplikasi premium dihitung dari total pendapatan perusahaan tahunan sebesar 13,3 miliar euro[3]. Aplikasi ini memungkinkan pengguna untuk mendengarkan jutaan lagu dari berbagai genre musik secara daring. Pengguna dapat membuat daftar putar pribadi, menelusuri berbagai artis dan album serta menemukan lagu atau konten audio berdasarkan riwayat lagu-lagu yang telah didengarkan dan disukai sebelumnya.



Gambar 1.3 Ulasan Pengguna di Google Play (Inggris)



Gambar 1.4 Ulasan Pengguna di Google Play (Indonesia)

Google play sebagai *platform* distribusi aplikasi tidak hanya menyediakan fitur untuk mengunduh, namun dapat melakukan pencarian aplikasi dan memberi ulasan untuk aplikasi yang telah digunakan. Seperti contoh ulasan pengguna pada Gambar 1.3 dan Gambar 1.4 ulasan tidak hanya menggambarkan bentuk kepuasan pengguna, tetapi juga berisi informasi penting yang dapat membantu pengembang meningkatkan kualitas aplikasi. Penelitian juga dilakukan menggunakan ulasan bahasa Indonesia dan bahasa Inggris sehingga informasi yang didapatkan akan lebih bervariasi untuk membuat keputusan lebih baik dan strategis berdasarkan sentimen pengguna yang lebih luas. Penelitian ini bertujuan untuk menjelajahi ulasan pengguna dengan pendekatan analisis sentimen. Tujuan analisis sentimen ulasan pengguna adalah untuk menentukan apakah suatu ulasan memiliki ulasan positif atau negatif [2]. Pemahaman mendalam terhadap sentimen ulasan dapat memberikan wawasan bermanfaat dalam meningkatkan kualitas dan fungsionalitas dari aplikasi serta aspek-aspek yang mempengaruhi penggunaan dan penerimaan aplikasi tersebut.

Melalui analisis sentimen penulis menentukan sentimen dari setiap ulasan dengan hasil akhir ditampilkan dalam diagram, *wordcloud*, atau persentase probabilitas untuk kelas sentimen positif dan negatif [4]. Untuk mencapai tujuan tersebut, algoritma *Random Forest* akan digunakan dalam membangun model klasifikasi. Random forest digunakan karena cenderung mudah diimplementasikan untuk dataset besar dan mampu menggabungkan hasil dari beberapa pohon keputusan yang berbeda untuk meningkatkan kinerja model. Tujuan akhir dari penelitian ini, memberikan wawasan kepada pembaca mengenai preferensi dan harapan pengguna terhadap aplikasi. Serta penelitian diharapkan dapat memberikan informasi terkait penerapan algoritma dalam proses analisis sentimen. Hasil

penelitian dapat digunakan oleh pihak terkait sebagai bentuk perbaikan dan informasi bagi pengguna terkait aplikasi Spotify. Dengan latar belakang yang ada, maka dari itu judul penelitian ini adalah "ANALISIS ULASAN PENGGUNA APLIKASI SPOTIFY PADA GOOGLE PLAY MENGGUNAKAN ALGORITMA *RANDOM FOREST*".

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, penelitian ini akan merumuskan masalah yang akan diteliti sebagai berikut :

1. Bagaimana tahapan proses dari analisis ulasan aplikasi spotify di google play dengan menggunakan algoritma *Random Forest*?
2. Bagaimana pencapaian akurasi yang didapatkan dari penerapan algoritma *Random Forest*?
3. Bagaimana performa model klasifikasi ulasan pengguna spotify menggunakan algoritma *Random Forest* dalam memprediksi sentimen positif dan negatif?

1.3 Batasan Masalah

Dalam proses penelitian, untuk lebih terarah berdasarkan perumusan masalah dan sesuai dengan kemampuan penulis, maka dibuat batasan dari masalah penelitian sebagai berikut :

1. Total data yang digunakan pada penelitian ini sebanyak 2000 data, terdiri dari 1000 ulasan berbahasa Indonesia dan 1000 ulasan berbahasa Inggris.
2. Penelitian menggunakan data ulasan relevan pengguna pada aplikasi spotify yang ada di google play.
3. Proses pelabelan data dilakukan untuk membagi ulasan dalam dua kategori label, yaitu positif dan negatif.
4. Metode yang digunakan dalam penelitian ini adalah algoritma *Random Forest*.

1.4 Tujuan Penelitian

Berdasarkan pada permasalahan diatas, maka tujuan dari penelitian ini sebagai berikut:

1. Memahami penerapan algoritma *Random Forest* dalam proses analisis ulasan pengguna berdasarkan ulasan paling relevan.

2. Menganalisis tingkat akurasi dari penerapan algoritma *Random Forest* dengan memanfaatkan ulasan pengguna spotify sebagai data awal.
3. Mengevaluasi keakuratan dan keandalan model klasifikasi algoritma *Random Forest* dalam memprediksi sentimen dari ulasan pengguna.

1.5 Manfaat Penelitian

Berdasarkan tujuan penelitian, maka diharapkan dapat memperoleh manfaat penelitian sebagai berikut :

1.5.1 Bagi Mahasiswa

1. Menambah pengetahuan penerapan algoritma dalam mengklasifikasikan ulasan pengguna aplikasi Spotify dengan algoritma *Random Forest*.
2. Mengetahui tanggapan dari pengguna pada aplikasi yang ada di google play, khususnya untuk aplikasi Spotify.
3. Mendapat gambaran terkait kendala yang banyak terjadi pada aplikasi melalui hasil analisis ulasan pengguna aplikasi.

1.5.2 Bagi Masyarakat

1. Memberikan gambaran kepada pembaca terkait ulasan opini positif dan negatif yang terdapat pada aplikasi, khususnya aplikasi Spotify.
2. Menjadi referensi untuk peneliti mendatang dengan tema atau model algoritma terkait.
3. Menambah wawasan kepada pembaca mengenai penerapan algoritma *Random Forest* dalam melakukan analisis data ulasan aplikasi.

1.6 Sistematika Penulisan

Sebagai acuan untuk penulis dalam menyusun skripsi ini, maka akan disusun dalam sistematika penulisan sebagai berikut :

- **BAB I PENDAHULUAN**

Pada bab ini berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

- **BAB II TINJAUAN PUSTAKA**

Pada bab ini menjelaskan terkait teori-teori yang berhubungan dengan konsep dari penelitian dan berisi tentang penelitian terkait.

- **BAB III METODOLOGI PENELITIAN**

Pada bab ini memberikan penjelasan tentang penelitian, metode pengumpulan data, jenis dan sumber data, teknik analisis data dan variabel penelitian.

- **BAB IV HASIL DAN PEMBAHASAN**

Pada bab ini memberikan penjelasan terkait penelitian yang terdiri dari objek penelitian, pengumpulan data dan deskripsi hasil penelitian.

- **BAB V PENUTUP**

Pada bab ini memberikan penjelasan mengenai kesimpulan dan saran berdasarkan hasil penelitian yang telah dilakukan.

- **DAFTAR PUSTAKA**





BAB V

PENUTUP

1.1 Kesimpulan

Penelitian yang sudah dilakukan pada ulasan aplikasi Spotify dengan menggunakan algoritma *Random Forest* terdapat beberapa kesimpulan yang didapatkan, diantaranya sebagai berikut :

1. Penelitian diawali dengan pengumpulan data pada google play menggunakan teknik *scraping* dengan jumlah 2000 data, terdiri dari 1000 ulasan bahasa Indonesia dan 1000 ulasan bahasa Inggris. Proses dilanjutkan dengan pelabelan data, *preprocessing*, pembobotan TF-IDF, *oversampling*, *split* data, pemodelan algoritma, evaluasi dan visualisasi hasil analisis.
2. Hasil penelitian yang dilakukan dengan menggunakan algoritma *Random Forest*, mendapatkan nilai akurasi sebesar 88,4% untuk ulasan bahasa indonesia. Sedangkan, untuk ulasan bahasa Inggris mendapatkan akurasi sebesar 93,2%.
3. Berdasarkan penelitian yang dilakukan dengan pengujian model *confusion Matrix* model mendapatkan hasil sebagai berikut :
 - a. Untuk ulasan bahasa Indonesia, hasil yang didapatkan yaitu : accuracy 88,4%, precision 84,1%, recall 91,8% dan f-1 score 87,7%.
 - b. Untuk ulasan bahasa Inggris, hasil dari pemodelan yang dilakukan yaitu: accuracy 93,2%, precision 93,6%, recall 93,6%, dan f-1 score 93,9%.
4. Dari hasil yang didapatkan setelah proses evaluasi dengan pengujian model *confusion matrix*, penerapan model algoritma *Random Forest* dapat dikatakan bekerja lebih baik pada penelitian untuk ulasan bahasa Inggris dengan nilai akurasi 93,6% daripada nilai akurasi ulasan bahasa Indonesia sebesar 88,4%.

1.2 Saran

Setelah penelitian dilakukan dan mendapatkan hasil berserta kesimpulan, maka saran untuk penelitian selanjutnya sebagai berikut :

1. Proses penelitian selanjutnya dapat menggunakan algoritma yang berbeda sebagai bentuk perbandingan dengan penelitian yang sudah dilakukan sebelumnya.
2. Tahap *preprocessing* dan pelabelan data dilakukan lebih detail, terutama untuk ulasan bahasa Indonesia.
3. Peneliti dapat menyesuaikan jumlah dataset dengan karakteristik model yang akan digunakan dalam proses penelitian untuk mendapatkan hasil lebih baik.



DAFTAR PUSTAKA

- [1] S. Y. M. Netti and Irwansyah, "Spotify: Aplikasi Music Streaming untuk Generasi Milenial," 2018. Accessed: Mar. 15, 2024. [Online]. Available: <https://journal.untar.ac.id/index.php/komunikasi/article/download/1102/1268>
- [2] A. M. Ndapamuri, D. Manongga, and A. Iriani, "Analisis Sentimen Ulasan Aplikasi Tripadvisor Dengan Metode Support Vector Machine, K-Nearest Neighbor, Dan Naive Bayes," *INOVTEK Polbeng - Seri Informatika*, vol. 8, no. 1, p. 127, Jun. 2023, doi: <https://doi.org/10.35314/isi.v8i1.3260>.
- [3] C. N. Harjadi, "Spotify Catat Rekor Pelanggan Premium Meningkatkan, Tren Streaming Musik Kini Ada Harganya - GoodStats." Accessed: Mar. 15, 2024. [Online]. Available: <https://goodstats.id/article/spotify-catat-rekor-pelanggan-premium-meningkat-tiap-tahun-tren-streaming-musik-kini-ada-harganya%20-wUBMv>
- [4] D. E. Sondakh, S. W. Taju, M. G. Tene, and A. E. T. Pangaila, "Sistem Analisis Sentimen Ulasan Aplikasi Belanja Online Menggunakan Metode Ensemble Learning," *CogITO Smart Journal*, vol. 9, no. 2, pp. 280–291, Dec. 2023, doi: <https://doi.org/10.31154/cogito.v9i2.525.280-291>.
- [5] M. Y. Aldean, P. Paradise, and N. A. Setya Nugraha, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 di Twitter Menggunakan Metode Random Forest Classifier (Studi Kasus: Vaksin Sinovac)," *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, vol. 4, no. 2, pp. 64–72, Jun. 2022, doi: <https://doi.org/10.20895/inista.v4i2.575>.
- [6] W. M. Baihaqi, I. R. Yunita, A. S. T. Damayanti, and L. Akhaerunnisa, "Analysis of Classification Algorithm Performance on User Review Sentiment of the Muamalat DIN Application," *CogITO Smart Journal*, vol. 9, no. 2, pp. 241–251, Dec. 2023, doi: <https://doi.org/10.31154/cogito.v9i2.511.241-251>.
- [7] A. S. Rahayu, A. Fauzi, and R. Rahmat, "Komparasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM) Pada Analisis Sentimen Spotify,"

- Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 2, p. 349, Dec. 2022, doi: <http://dx.doi.org/10.30865/json.v4i2.5398>.
- [8] A. D. A. Putra and S. Juanita, “Analisis Sentimen pada Ulasan pengguna Aplikasi Bibit Dan Bareksa dengan Algoritma KNN,” *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 8, no. 2, pp. 636–646, Jun. 2021, doi: <https://doi.org/10.35957/jatisi.v8i2.962>.
- [9] “Sentiment analysis: Understanding Public Opinion with Social Data Insights - FasterCapital.” Accessed: Mar. 21, 2024. [Online]. Available: <https://fastercapital.com/content/Sentiment-analysis--Understanding-Public-Opinion-with-Social-Data-Insights.html>
- [10] J. Asian, M. Dholah Rosita, and T. Mantoro, “Sentiment Analysis for the Brazilian Anesthesiologist Using Multi-Layer Perceptron Classifier and Random Forest Methods,” *Jurnal Online Informatika*, vol. 7, no. 1, p. 132, Sep. 2022, doi: [10.15575/join.v7i1.900](https://doi.org/10.15575/join.v7i1.900).
- [11] “Apa itu Machine Learning? Beserta Pengertian dan Cara Kerjanya - Dicoding Blog.” Accessed: Jan. 13, 2024. [Online]. Available: <https://www.dicoding.com/blog/machine-learning-adalah/>
- [12] I. L. Kharisma, D. A. Septiani, A. Fergina, and K. Kamdan, “Penerapan Algoritma Decision Tree untuk Ulasan Aplikasi Vidio di Google Play,” *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 9, no. 2, pp. 218–226, Sep. 2023, doi: [10.25077/TEKNOSI.v9i2.2023.218-226](https://doi.org/10.25077/TEKNOSI.v9i2.2023.218-226).
- [13] D. Irawan, E. B. Perkasa, Y. Yurindra, D. Wahyuningsih, and E. Helmud, “Perbandingan Klassifikasi SMS Berbasis Support Vector Machine, Naive Bayes Classifier, Random Forest dan Bagging Classifier,” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 10, no. 3, pp. 432–437, Dec. 2021, doi: <https://doi.org/10.32736/sisfokom.v10i3.1302>.
- [14] M. R. Givari, M. R. Sulaeman, and Y. Umaidah, “Perbandingan Algoritma SVM, Random Forest Dan XGBoost Untuk Penentuan Persetujuan Pengajuan Kredit,” *NUANSA INFORMATIKA*, vol. 16, no. 1, pp. 141–149, Jan. 2022, doi: [10.25134/nuansa.v16i1.5406](https://doi.org/10.25134/nuansa.v16i1.5406).

- [15] Hartatik, M. E. Koibur, G. P. Insany, and et all, *SAINS DATA: Strategi, Teknik, dan Model Analisis Data*, Cetakan Pertama. Bandung: Kaizen Media Publishing, 2023.
- [16] F. A. Larasati, D. E. Ratnawati, and B. T. Hanggara, “Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest,” 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [17] A. Sujjada and A. Fergina, “Implementasi Metode Vector Space Model Untuk Deteksi Emosi Menggunakan Data Teks Twitter,” *Jurnal Restikom : Riset Teknik Informatika dan Komputer*, vol. 3, no. 3, pp. 116–129, 2021, [Online]. Available: <https://restikom.nusaputra.ac.id>
- [18] “Python: Pengertian, Contoh Penggunaan, dan Manfaat Mempelajarinya.” Accessed: Jan. 13, 2024. [Online]. Available: <https://www.dicoding.com/blog/python-pengertian-contoh-penggunaan-dan-manfaat-mempelajarinya/>
- [19] R. Maulid, “Mengenal Streamlit, Tools Favorit Data Scientist,” DQLab. Accessed: Jan. 12, 2024. [Online]. Available: <https://dqlab.id/mengenal-streamlit-tools-favorit-data-scientist>
- [20] I. Afdhal, R. Kurniawan, I. Iskandar, R. Salambue, E. Budianita, and F. Syafria, “Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia,” *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 1, 2022.
- [21] A. Wijaya, Meilinda, and M. Maharani, “Analisis Sentimen Ulasan Aplikasi Shazam Di Google Play Store Menggunakan Support Vector Machine,” *ZONasi: Jurnal Sistem Informasi*, vol. 6, no. 1, pp. 197–207, Feb. 2024, doi: 10.31849/zn.v6i1.17994.
- [22] H. H. Sinaga and S. Agustian, “Pebandingan Metode Decision Tree dan XGBoost untuk Klasifikasi Sentimen Vaksin Covid-19 di Twitter,” *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 8, no. 3, pp. 107–114, Dec. 2022, doi: 10.25077/TEKNOSI.v8i3.2022.107-114.
- [23] R. Maulana, A. Voutama, and T. Ridwan, “Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma

- NBC,” *Jurnal Teknologi Terpadu*, vol. 9, no. 1, pp. 42–48, Jul. 2023, doi: 10.54914/jtt.v9i1.609.
- [24] S. Kurnia, H. G. Prabowo, and B. Suharmanto, “Klasifikasi Tweet Cyberbullying dengan Menggunakan Algoritma Random Forest,” in *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, Jakarta, Indonesia: upnvj, 2022. Accessed: Jan. 13, 2024. [Online]. Available: <https://conference.upnvj.ac.id/index.php/senamika/article/view/2207>
- [25] D. Sierra, “Algoritma TF — IDF,” Medium. Accessed: Apr. 28, 2024. [Online]. Available: <https://dltsierra.medium.com/algoritma-tf-idf-633e17d10a80>
- [26] F. Ihsan Septian and I. Lucia Kharisma, “Implementasi Metode Bidirectional Encoder Representations from Transformers (BERT) untuk Analisis Sentimen Komentar Pengguna Aplikasi Dana di Instagram,” 2023.
- [27] J. Asian, M. Dholah Rosita, and T. Mantoro, “Sentiment Analysis for the Brazilian Anesthesiologist Using Multi-Layer Perceptron Classifier and Random Forest Methods,” *Jurnal Online Informatika*, vol. 7, no. 1, p. 132, Sep. 2022, doi: 10.15575/join.v7i1.900.
- [28] A. Sujjada, J. N. Novianti, and I. G. T. Isa, “Analisis Sentimen Terhadap Review Bank Digital Pada Google Play Store Menggunakan Metode Support Vector Machine (SVM),” vol. 9, no. 2, pp. 122–135, 2023, [Online]. Available: <https://rekayasa.nusaputra.ac.id/index>
- [29] G. Ginabila and A. Fauzi, “Analisis Sentimen Terhadap Pemutar Musik Online Spotify Dengan Algoritma Naive Bayes dan Support Vector Machine,” *Jurnal Ilmiah ILKOMINFO - Ilmu Komputer & Informatika*, vol. 6, no. 2, pp. 111–122, Jul. 2023, doi: 10.47324/ilkominfo.v6i2.180.
- [30] R. Nurlaely, D. Sartika Simatupang, and I. Lucia Kharisma, “Analisis Sentimen Twitter Terhadap Cyberbullying Menggunakan Metode Support Vector Machine (SVM),” vol. 4, no. 2, pp. 376–384, 2023, doi: 10.37859/coscitech.v4i2.5161.

