

Analisis Sentimen Menggunakan Metode Naive Bayes dan Support Vector Machine pada Ulasan Aplikasi Joox Music

Wayan Pondra Lesmana, Andri Wijaya

Sistem Informasi, Fakultas Sains Dan Teknologi, Universitas Katolik Musi Charitas

Email: wayanpondra3@gmail.com

Abstract

In this digital era, user reviews of mobile applications play an important role in shaping people's perception of a product or service. This study presents an analysis of opinion reviews of Joox music application users using the Bayesian Naïve classification method. The Naive Bayes method is used to classify user ratings into positive, negative, or neutral opinion categories. The dataset contains a large number of Joox music user reviews, added manually and used as examples in this study. The results of sentiment analysis show that the Bayesian Naive method can identify and classify the sentiment of Joox music user reviews very accurately. These findings provide developers with valuable information to understand user opinions and respond more effectively to customer feedback. The study also helps to understand the effectiveness of the Bayesian Naive classification method in handling large and complex user review data.

Keywords: *Sentiment analysis, Support Vector Machine, Naïve Bayes*

Abstrak

Di era digital ini, ulasan pengguna terhadap aplikasi seluler berperan penting dalam membentuk persepsi masyarakat terhadap suatu produk atau layanan. Penelitian ini menyajikan analisis opini review pengguna aplikasi musik Joox Music dengan menggunakan metode klasifikasi Naïve Bayesian. Metode Naive Bayes digunakan untuk mengklasifikasikan penilaian pengguna ke dalam kategori opini positif, negatif, atau netral. Kumpulan data berisi sejumlah besar ulasan pengguna Joox Music, ditambahkan secara manual dan digunakan sebagai contoh dalam penelitian ini. Hasil analisis sentimen menunjukkan bahwa metode Naive Bayesian dapat mengidentifikasi dan mengklasifikasikan sentimen ulasan pengguna Joox Music dengan sangat akurat. Temuan ini memberikan informasi berharga kepada pengembang untuk memahami opini pengguna dan merespons masukan pelanggan dengan lebih efektif. Studi ini juga membantu untuk memahami efektivitas metode klasifikasi Naive Bayesian dalam menangani data ulasan pengguna yang besar dan kompleks dalam konteks aplikasi seluler. Secara ringkas dapat disimpulkan bahwa analisis opini dengan metode klasifikasi Naive Bayes membuka jalan bagi pengelolaan ulasan pengguna yang lebih efektif dan dapat membantu meningkatkan pengalaman pengguna saat menggunakan aplikasi Joox Music dan program serupa lainnya.

Kata Kunci: *Analysis sentiment, Support Vector Machine, Naïve Bayes*

1. PENDAHULUAN

1.1 Latar Belakang

Dalam era digital ini, aplikasi musik streaming telah menjadi bagian integral dari kehidupan sehari-hari pengguna internet. Joox Music, sebagai salah satu aplikasi musik streaming terkemuka, telah mencuri perhatian pengguna dengan berbagai fitur inovatif dan koleksi musik yang luas. Namun, sejalan dengan popularitasnya, meningkatnya volume ulasan pengguna juga memberikan tantangan dalam memahami sentimen dan umpan balik pengguna terhadap aplikasi ini. Di Indonesia, berdasarkan survei penetrasi dan perilaku penggunaan internet tahun 2022 yang dilakukan oleh APJII (Asosiasi Penyedia Jasa Internet Indonesia), pengguna internet di dominasi oleh kelompok usia 19-34 tahun sebesar 98,64% dari total pengguna internet di Indonesia yaitu 210 juta orang. Mereka mengakses platform streaming musik sebesar 38,51% dengan menggunakan perangkat gawai sebesar 88,2% dibandingkan mengakses aplikasi tv berbasis internet sebesar 11,10% [9]. Industri streaming musik global di Q2 tahun 2021 telah berkembang pesat sebesar 26% dibandingkan Q2 2020 berdasarkan penelitian yang dilakukan oleh (Midia, 2021). Platform streaming musik seperti Spotify, Apple Music, dan Youtube Music mengalami peningkatan yang besar dari segi layanan dan sejumlah inovasi yang menarik. Pada Q2 2021 Spotify menguasai 31% pangsa pasar industri streaming musik skala global [8].

Pada penelitian ini akan dilakukan analisis sentimen terhadap ulasan aplikasi Joox Music. Data yang dikumpulkan adalah kalimat ulasan dan rating skala lima bintang aplikasi Joox Music pada layanan GooglePlay Store. Kalimat ulasan di analisis melalui beberapa tahapan pembersihan data, seperti pelabelan, pembobotan, dan klasifikasi. Klasifikasi yang digunakan dalam penelitian ini adalah Naive Bayes (NB) & Support Vector Machine (SVM). Banyak penelitian tentang sentimen analisis yang menggunakan klasifikasi NB dikarenakan memiliki beberapa kelebihan seperti sederhana, cepat dan memiliki akurasi tinggi. SVM diaplikasikan pada kasus klasifikasi teks dan efisien dalam menangani jumlah data yang besar. Hasil akhir dari penelitian ini adalah melihat perbandingan nilai F1-Score terbaik untuk kasus analisis sentiment pada ulasan aplikasi Joox Music.

1.2 Tujuan atau Ruang Lingkup Tulisan

Tujuan utama dari penelitian ini adalah mengaplikasikan metode klasifikasi Naive Bayes pada ulasan pengguna Joox Music untuk melakukan analisis sentimen. Dalam konteks ini, Saya bertujuan untuk:

1. Mengidentifikasi sentimen umum pengguna terhadap aplikasi Joox Music, baik itu positif, negatif, atau netral.
2. Menganalisis aspek-aspek tertentu dari aplikasi (seperti antarmuka pengguna, kualitas streaming, dan pilihan musik) yang mungkin menjadi fokus sentimen pengguna.
3. Mengevaluasi keefektifan metode klasifikasi Naive Bayes dalam mengklasifikasikan sentimen pengguna Joox Music.
4. Memberikan rekomendasi perbaikan atau peningkatan berdasarkan hasil analisis sentimen yang telah dilakukan.

1.3 Perumusan Masalah

Berdasarkan latar belakang diatas, rumusan masalah yang di angkat pada penelitian ini adalah untuk mendapatkan model klasifikasi menggunakan algoritma Naive Bayes pada ulasan aplikasi Joox Music dan membandingkan hasil nilai F1-Score.

1.4 Batasan Masalah

1. Data yang dikumpulkan adalah kalimat ulasan dan rating bintang satu sampai Bintang lima aplikasi Joox Music pada layanan GooglePlay Store.
2. Pelabelan data rating di asumsikan menjadi;rating 5,4,3 menjadi positif dan rating 2,1menjadi negatif.
3. Akurasi F1-Score digunakan sebagai ukuran yang lebih baik terhadap keakuratan data yang tidak seimbang.

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen adalah bidang studi yang menganalisis pendapat, sentimen, penilaian, sikap dan emosi seseorang terhadap suatu entitas dan atributnya yang diungkapkan dalam sebuah teks tertulis (*Liu, 2015*). Entitas yang dimaksud dapat berupa produk, layanan, organisasi, individu, kejadian, isu, atau sebuah topik. [1]

Menurut Huang (2009), Analisis sentimen atau opinion mining merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini. Besarnya pengaruh dan manfaat dari analisis sentimen menyebabkan penelitian dan aplikasi berbasis analisis sentimen berkembang pesat. Bahkan di Amerika terdapat sekitar 20-30 perusahaan yang memfokuskan pada layanan analisis sentiment.[2]

Pada penelitian ini dilakukan analisis sentimen untuk mengetahui pendapat seseorang yang menggunakan aplikasi Trebel. Pendapat ini dapat digolongkan positif, netral atau negatif. Tingkat persepsi terhadap aplikasi Trebel dapat dijadikan parameter untuk menilai baik buruknya tingkat layanan yang diberikan.

2.2 Metode Naive Bayes

Metode Naive Bayes adalah metode klasifikasi yang berbasis pada teorema probabilitas Bayes. Dalam konteks analisis sentimen, metode ini dapat digunakan untuk mengklasifikasikan teks ke dalam kategori sentimen berdasarkan probabilitas yang terkait dengan kata-kata yang muncul dalam teks tersebut.[1]

Menurut Oslan dan Delen dalam (Kurniawan, Afif F, 2013) Naïve Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes yaitu dengan memprediksi masa depan berdasarkan data dari masa lalu. Naïve Bayes untuk setiap kelas keputusan menghitung probabilitas dengan syarat bahwa kelas keputusan adalah benar, mengingat vector informasi objek. Algoritma ini mengasumsikan bahwa atribut objek adalah independen. Probabilitas yang terlibat dalam memproduksi perkiraan akhir dihitung sebagai jumlah frekuensi dari “master” tabel keputusan.[1].

2.3 Support Vector Machine

Support Vector Machine (SVM) adalah suatu teknik untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi (Santosa, 2007). SVM memiliki prinsip dasar linier classifier yaitu kasus klasifikasi yang secara linier dapat dipisahkan, namun SVM telah dikembangkan agar dapat bekerja pada problem non-linier dengan memasukkan konsep kernel pada ruang kerja berdimensi tinggi. Pada ruang berdimensi tinggi, akan dicari hyperplane (hyperplane) yang dapat memaksimalkan jarak (margin) antara kelas data.[7]

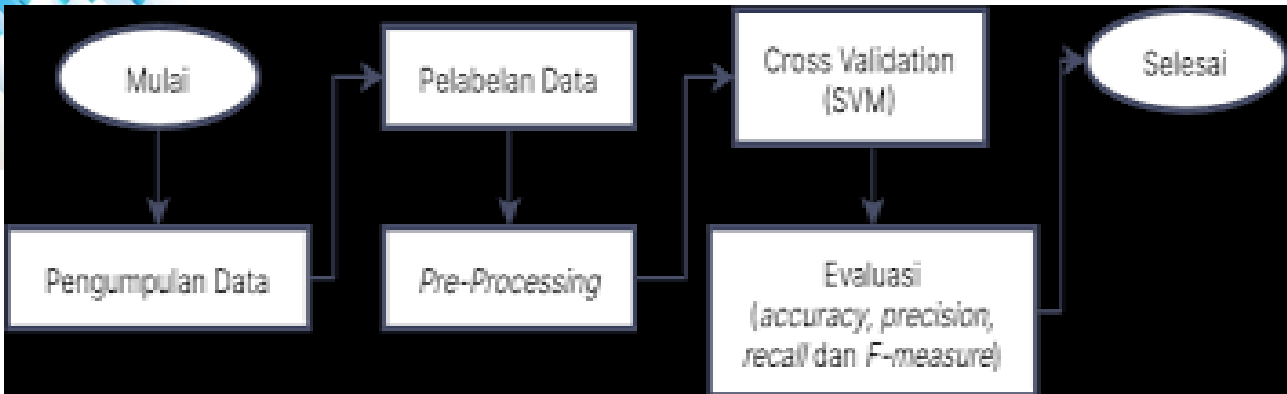
2.4 Joox Music

JOOX Music adalah aplikasi streaming musik freemium dari Tencent Mobility Limited yang memungkinkan pengguna untuk mendengarkan lebih dari 40 juta lagu dari seluruh dunia. Aplikasi ini tidak hanya memungkinkan pemutaran lagu favorit, tetapi juga menyertakan fitur media sosial seperti ruang obrolan video, berbagi video pendek, dan karaoke dengan teman-teman. JOOX Music telah mendominasi pasar Asia, beroperasi di Hong Kong, Macau, Malaysia, Thailand, Myanmar, Indonesia, dan beberapa bagian dunia lainnya. Meskipun memiliki sekitar setengah jumlah lagu yang ada di Spotify, keberadaan tanpa biaya langganan berulang membuatnya menarik. Aplikasi ini juga menawarkan fitur-fitur menarik seperti rekomendasi lagu, playlist personal, dan JOOX Live yang memungkinkan pengguna mendekati artis favorit mereka. JOOX Music juga memungkinkan pengguna untuk memainkan lagu sesuai permintaan, mengunduh musik untuk didengarkan secara offline, dan menawarkan kualitas audio yang tinggi untuk streaming musik. Aplikasi ini tersedia di Android, iOS, dan desktop, serta dapat diakses di Hong Kong, Thailand, Malaysia, Indonesia, dan Myanmar.

3. METODE PENELITIAN

A. Gambaran umum sistem

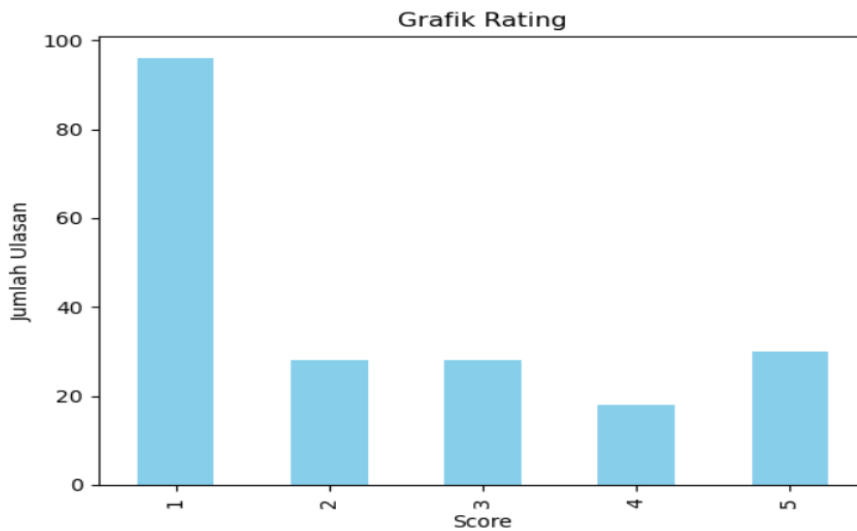
Pada bagian ini menjelaskan perancangan sistem yang akan digunakan dalam penelitian ini. Gambaran umum pada sistem sistem klasifikasi teks yang akan dibuat dari dataset ulasan aplikasi Joox Music adalah sebagai berikut :



Gambar 1 Diagram Alir Sistem Yang Dibangun

B. Dataset

Dataset yang digunakan berjumlah 200 data. Dataset bersumber dari ulasan para pengguna aplikasi Joox Music yang telah di scrapping menggunakan phyton di anaconda navigator.



Gambar 2 Grafik Perbandingan Ulasan pada Aplikasi Joox Music

Pelabelan data dibuat menjadi dua kelas yaitu kelas positif dan negatif yang berasal dari nilai rating. Dengan asumsi rating 5,4, dan 3 menjadi kelas positif, 2 dan 1 menjadi kelas negatif.

C. Preprocessing

Tahap data preprocessing dilakukan untuk mengurangi noise sebelum data diolah lebih lanjut untuk di ekstraksi sampai tahap klasifikasi. Tahap preprocessing yang dilakukan dalam penelitian ini adalah sebagai berikut:

Parah joox sekarang tidak seperti dulu tidak ada kesalahan atau BUG, kesalahan nya sekarang lagu yang saya taruh di PlayList & Favorit sebagian lagu tidak bisa di putar, 😞 lebih baik pindah saja kalau begini dan tidak ada perbaikan, fitur VIP sekarang di rubah untuk pengguna FREEPLAYER kasian tidak bisa ngakses VIP, joox dulu bisa akses FREE VIP, jadi inti nya kalau kedepan nya tidak di rubah atau tidak di perbaiki SAYA lebih baik pindah tempat, saya FANS joox dari awal rilis, TAPI yasudahlah !

Tabel 1 Kalimat Ulasan Sebelum Dilakukan Preprocessing

1. Case Folding dan Cleaning data

a. Case folding

Case folding adalah langkah di mana semua huruf dalam teks diubah menjadi huruf kecil (lowercase). Ini dilakukan untuk memastikan konsistensi dalam analisis teks, sehingga kata yang sama dengan huruf besar atau kecil dianggap identik.

parah joox sekarang tidak seperti dulu tidak ada kesalahan atau bug, kesalahan nya sekarang lagu yang saya taruh di playlist & favorit sebagian lagu tidak bisa di putar, 😞 lebih baik pindah saja kalau begini dan tidak ada perbaikan, fitur vip sekarang di rubah untuk pengguna freeplayer kasian tidak bisa ngakses vip, joox dulu bisa akses free vip, jadi inti nya kalau kedepan nya tidak di rubah atau tidak di perbaiki saya lebih baik pindah tempat, saya fans joox dari awal rilis, tapi yasudahlah !

Tabel 2 Case Folding

b. Cleaning data

Pembersihan data melibatkan penghilangan atau penggantian elemen-elemen yang tidak diinginkan atau dapat mengganggu analisis teks. Ini termasuk penghilangan karakter khusus, tanda baca, angka, atau kata yang tidak relevan.

parah joox sekarang tidak seperti dulu tidak ada kesalahan atau bug kesalahan nya sekarang lagu yang saya taruh di playlist favorit sebagian lagu tidak bisa di putar lebih baik pindah saja kalau begini dan tidak ada perbaikan fitur vip sekarang di rubah untuk pengguna freeplayer kasian tidak bisa ngakses vip joox dulu bisa akses free vip jadi inti nya kalau kedepan nya tidak di rubah atau tidak di perbaiki saya lebih baik pindah tempat saya fans joox dari awal rilis tapi yasudahlah

Tabel 3 Cleaning Data

2. Tokenisasi

Tokenisasi adalah proses memecah teks atau kalimat menjadi unit-unit yang lebih kecil yang disebut token. Token dapat berupa kata, frasa, atau karakter, tergantung pada tingkat detail yang dibutuhkan. Tujuan tokenisasi adalah untuk memecah teks menjadi elemen-elemen yang lebih mudah diolah atau dianalisis.

```
[ 'parah', 'joox', 'sekarang', 'tidak', 'seperti', 'dulu', 'tidak', 'ada', 'kesalahan', 'atau', 'bug', 'kesalahan', 'nya', 'sekarang', 'lagu', 'yang', 'saya', 'taruh', 'di', 'playlist', 'favorit', 'sebagian', 'lagu', 'tidak', 'bisa', 'di', 'putar', 'lebih', 'baik', 'pindah', 'saja', 'kalau', 'begini', 'dan', 'tidak', 'ada', 'perbaikan', 'fitur', 'vip', 'sekarang', 'di', 'rubah', 'untuk', 'pengguna', 'freeplayer', 'kasian', 'tidak', 'bisa', 'ngakses', 'vip', 'joox', 'dulu', 'bisa', 'akses', 'free', 'vip', 'jadi', 'inti', 'nya', 'kalau', 'kedepan', 'nya', 'tidak', 'di', 'rubah', 'atau', 'tidak', 'di', 'perbaiki', 'saya', 'lebih', 'baik', 'pindah', 'tempat', 'saya', 'fans', 'joox', 'dari', 'awal', 'rilis', 'tapi', 'yasudahlah' ]
```

Tabel 4 Tahap Tokenisasi

3. Stopword Removal

Stopword removal adalah proses menghapus kata-kata umum yang sering muncul dalam teks, tetapi tidak memberikan informasi signifikan dalam analisis. Dalam banyak kasus, kata-kata seperti "dan", "atau", "yang", dll., dianggap sebagai stopwords dan dihapus dari teks.

```
['parah', 'joox', 'kesalahan', 'bug', 'kesalahan', 'nya', 'lagu',
'taruh', 'playlist', 'favorit', 'lagu', 'putar', 'pindah',
'perbaiki', 'fitur', 'vip', 'rubah', 'pengguna', 'freeplayer',
'kasian', 'ngakses', 'vip', 'joox', 'akses', 'free', 'vip', 'inti',
'nya', 'kedepan', 'nya', 'rubah', 'perbaiki', 'pindah', 'fans',
'joox', 'rilis', 'yasudahlah']
```

Tabel 5 Kalimat Ulasan Setelah Stopword Removal

4. Stemming

Stemming adalah proses menghilangkan akhiran atau awalan kata sehingga hanya menyisakan akar kata atau bentuk dasarnya. Tujuan dari stemming adalah untuk mengurangi kata-kata ke bentuk dasarnya agar kata-kata yang semantiknya serupa dapat dianggap sama.

```
['parah', 'joox', 'sekarang', 'tidak', 'seperti', 'dulu', 'tidak',
'ada', 'salah', 'atau', 'bug', 'salah', 'nya', 'sekarang', 'lagu',
'yang', 'saya', 'taruh', 'di', 'playlist', 'favorit', 'bagi',
'lagu', 'tidak', 'bisa', 'di', 'putar', 'lebih', 'baik', 'pindah',
'saja', 'kalau', 'begini', 'dan', 'tidak', 'ada', 'baik', 'fitur',
'vip', 'sekarang', 'di', 'rubah', 'untuk', 'guna', 'freeplayer',
'kasi', 'tidak', 'bisa', 'ngakses', 'vip', 'joox', 'dulu', 'bisa',
'akses', 'free', 'vip', 'jadi', 'inti', 'nya', 'kalau', 'depan',
'nya', 'tidak', 'di', 'rubah', 'atau', 'tidak', 'di', 'baik',
'saya', 'lebih', 'baik', 'pindah', 'tempat', 'saya', 'fans',
'joox', 'dari', 'awal', 'rilis', 'tapi', 'yasudahlah']
```

Tabel 6 Kalimat Ulasan Setelah Proses Stemming

D. TF-IDF (Term Frequency – Inverse Document Frequency)

Pembobotan kata adalah proses menambahkan nilai pada seluruh kata dalam kalimat ulasan yang telah melalui tahap preprocessing. Pembobotan menggunakan metode TF-IDF dituliskan dalam persamaan (1) dan (2) lalu hasilnya akan dilanjutkan pada tahap klasifikasi.[3] Langkah menghitung pembobotan kata dimulai dari menghitung frekuensi kemunculan kata keluar dalam kalimat (TF). Lalu menghitung angka frekuensi kata (DF) pada dokumen dan menghitung nilai invers (IDF). Setelah mendapatkan nilai TF, tahap selanjutnya menghitung nilai invers. Persamaan menghitung nilai IDF adalah sebagai berikut :

$$IDF(w) = \log\left(\frac{N}{DF(w)}\right)$$

Tahap selanjutnya dalam TF-IDF adalah dengan mengkalikan nilai TF dengan IDF seperti persamaan

$$W_{ij} = tf_{ij} \log\left(\frac{D}{df_j}\right)$$

E. Klasifikasi

1. Naïve Bayes

Metode Naive Bayes merupakan metode supervised learning yang ditemukan oleh Thomas Bayes pada abad ke-18[4]. Klasifikasi Naive Bayes digunakan untuk memprediksi probabilitas keanggotaan dalam satu kelas. Kelebihan algoritma Naive Bayes yaitu mudah diimplementasikan, efisiensi waktu, dan dapat menangani data yang besar[5]. Salah satu model Naive Bayes yang sering digunakan untuk klasifikasi teks adalah Multinomial Naive Bayes. Perumusan Multinomial Naive Bayes menggunakan pembobotan TF-IDF dituliskan pada persamaan.[1]

$$\hat{P}(t|c) = \frac{W_{ct}+1}{(\sum_{w \in V} W_{ct})+B}$$

2. Support Vector Machine

Support Vector Machine (SVM) adalah salah satu algoritma machine learning yang biasa digunakan dalam proses klasifikasi. Algoritma SVM bertujuan untuk mencari hyperlane terbaik untuk memisahkan kedalam kelas positif dan negatif, dan memaksimalkan margin diantara dua kelas tersebut.[6] Hyperlane adalah sebuah fungsi yang digunakan untuk pemisah antar kelas. Proses klasifikasi pada penelitian ini menggunakan bahasa pemrograman python dan library scikit learn svm. Fungsi kernel yang umum digunakan pada SVM adalah kerner linear, polynomial, dan RBF.

F. Evaluasi

Tahap terakhir dalam proses analisis sentimen pada penelitian ini adalah pengujian menggunakan scikit-learn untuk mendapatkan nilai F1- Score dari hasil klasifikasi algoritma NB dan SVM.

4. Hasil dan Pembahasan

A. Hasil Pengujian

Pengujian dilakukan dengan membagi dataset menjadi 80% data train & 20% data test yang telah melalui tahapan pembobotan dan data preprocessing.

1. Skenario pengujian pertama

Pada skenario pengujian pertama di metode SVM dilakukan tanpa menggunakan parameter pada kernel linear, polynomial & RBF.

Akurasi SVM: 0.55				
Laporan Klasifikasi SVM:				
	precision	recall	f1-score	support
1	0.54	0.95	0.69	21
2	0.00	0.00	0.00	5
3	0.00	0.00	0.00	6
4	0.00	0.00	0.00	3
5	0.67	0.40	0.50	5
accuracy			0.55	40
macro avg	0.24	0.27	0.24	40
weighted avg	0.37	0.55	0.42	40

Tabel 7 Laporan Klasifikasi SVM

a. Support Vector Machine

1. Akurasi SVM: 0.55:

Akurasi adalah proporsi total prediksi yang benar. Dalam kasus ini, akurasi adalah 55%, yang berarti model SVM benar sekitar 55% dari keseluruhan data pengujian.

2. Laporan Klasifikasi SVM:

- Precision: Proporsi instance yang diprediksi benar dari semua instance yang diprediksi positif.
- Recall: Proporsi instance yang diprediksi benar dari semua instance yang sebenarnya positif.
- F1-Score: Harmonic mean dari precision dan recall.

Dalam kasus ini, fokus pada kelas "1" menunjukkan bahwa SVM memiliki presisi yang relatif baik (54%), tetapi recall yang lebih tinggi (95%). Ini berarti SVM cenderung baik dalam mengidentifikasi instance positif sebenarnya, tetapi mungkin kehilangan beberapa instance positif. Untuk kelas lainnya, model memiliki kinerja yang rendah.

Akurasi Naive Bayes: 0.525				
Laporan Klasifikasi Naive Bayes:				
	precision	recall	f1-score	support
1	0.53	1.00	0.69	21
2	0.00	0.00	0.00	5
3	0.00	0.00	0.00	6
4	0.00	0.00	0.00	3
5	0.00	0.00	0.00	5
accuracy			0.53	40
macro avg	0.11	0.20	0.14	40
weighted avg	0.28	0.53	0.36	40

Tabel 8 Laporan Klasifikasi Naive Bayes

b. Naive Bayes

1. Akurasi Naive Bayes: 0.525:

Akurasi adalah proporsi total prediksi yang benar. Dalam kasus ini, akurasi Naive Bayes adalah 52.5%.

2. Laporan Klasifikasi Naive Bayes: Precision, Recall, F1-Score: Sama seperti pada SVM.

Dalam kasus ini, model Naive Bayes menunjukkan akurasi yang sedikit lebih rendah daripada SVM. Fokus pada kelas "1" menunjukkan bahwa Naive Bayes memiliki presisi yang relatif baik (53%) dan recall yang tinggi (100%) untuk kelas tersebut. Namun, performa model buruk untuk kelas lainnya.

2. Skenario pengujian kedua

a) Kernel Linier

Akurasi SVM dengan Kernel Linier: 0.55				
Laporan Klasifikasi SVM dengan Kernel Linier:				
	precision	recall	f1-score	support
1	0.54	0.95	0.69	21

	2	0.00	0.00	0.00	5
	3	0.00	0.00	0.00	6
	4	0.00	0.00	0.00	3
	5	0.67	0.40	0.50	5
	accuracy			0.55	40
	macro avg	0.24	0.27	0.24	40
	weighted avg	0.37	0.55	0.42	40

Tabel 9 Kernel Linier

Akurasi SVM dengan Kernel Linier: 0.55:

- a. Akurasi adalah proporsi total prediksi yang benar. Dalam kasus ini, akurasi SVM dengan kernel linier adalah 55%, yang berarti model SVM dengan kernel linier benar sekitar 55% dari keseluruhan data pengujian.

Laporan Klasifikasi SVM dengan Kernel Linier:

- b. Precision: Proporsi instance yang diprediksi benar dari semua instance yang diprediksi positif.
- c. Recall: Proporsi instance yang diprediksi benar dari semua instance yang sebenarnya positif.
- d. F1-Score: Harmonic mean dari precision dan recall.

Dalam kasus ini, fokus pada kelas "1" menunjukkan bahwa SVM dengan kernel linier memiliki presisi yang relatif baik (54%) dan recall yang tinggi (95%) untuk kelas tersebut. Ini berarti SVM dengan kernel linier cenderung baik dalam mengidentifikasi instance positif sebenarnya, tetapi mungkin kehilangan beberapa instance positif. Untuk kelas lainnya, model memiliki kinerja yang rendah, seperti yang diindikasikan oleh precision, recall, dan F1-Score yang rendah.

b) Kernel Polynomial

Akurasi SVM dengan Kernel Polinomial: 0.525					
Laporan Klasifikasi SVM dengan Kernel Polinomial:					
	precision	recall	f1-score	support	
	1	0.53	1.00	0.69	21
	2	0.00	0.00	0.00	5
	3	0.00	0.00	0.00	6
	4	0.00	0.00	0.00	3
	5	0.00	0.00	0.00	5
	accuracy			0.53	40
	macro avg	0.11	0.20	0.14	40
	weighted avg	0.28	0.53	0.36	40

Tabel 10 Kernel Polynomial

Akurasi SVM dengan Kernel Polynomial: 0.525:

- a. Akurasi adalah proporsi total prediksi yang benar. Dalam kasus ini, akurasi SVM dengan kernel polinomial adalah 52.5%, yang berarti model SVM dengan kernel polinomial benar sekitar 52.5% dari keseluruhan data pengujian.

Laporan Klasifikasi SVM dengan Kernel Polinomial:

- b. Precision: Proporsi instance yang diprediksi benar dari semua instance yang diprediksi positif.
- c. Recall: Proporsi instance yang diprediksi benar dari semua instance yang sebenarnya positif.
- d. F1-Score: Harmonic mean dari precision dan recall.

Dalam kasus ini, fokus pada kelas "1" menunjukkan bahwa SVM dengan kernel polinomial memiliki presisi yang relatif baik (53%) dan recall yang tinggi (100%) untuk kelas tersebut. Ini berarti SVM dengan kernel polinomial cenderung baik dalam mengidentifikasi instance positif sebenarnya untuk kelas tersebut. Namun, untuk kelas lainnya, model memiliki kinerja yang rendah, seperti yang diindikasikan oleh precision, recall, dan F1-Score yang rendah.

c) Kernel RBF

Akurasi SVM dengan Kernel RBF: 0.525				
Laporan Klasifikasi SVM dengan Kernel RBF:				
	precision	recall	f1-score	support
1	0.53	1.00	0.69	21
2	0.00	0.00	0.00	5
3	0.00	0.00	0.00	6
4	0.00	0.00	0.00	3
5	0.00	0.00	0.00	5
accuracy			0.53	40
macro avg	0.11	0.20	0.14	40
weighted avg	0.28	0.53	0.36	40

Tabel 11 Kernel RBF

Akurasi SVM dengan Kernel RBF: 0.525:

Akurasi adalah proporsi total prediksi yang benar. Dalam kasus ini, akurasi SVM dengan kernel RBF adalah 52.5%, yang berarti model SVM dengan kernel RBF benar sekitar 52.5% dari keseluruhan data pengujian.

Laporan Klasifikasi SVM dengan Kernel RBF:

- a. Precision: Proporsi instance yang diprediksi benar dari semua instance yang diprediksi positif.
- b. Recall: Proporsi instance yang diprediksi benar dari semua instance yang sebenarnya positif.
- c. F1-Score: Harmonic mean dari precision dan recall.

Dalam kasus ini, fokus pada kelas "1" menunjukkan bahwa SVM dengan kernel RBF memiliki presisi yang relatif baik (53%) dan recall yang tinggi (100%) untuk kelas tersebut. Ini berarti SVM dengan kernel RBF cenderung baik dalam mengidentifikasi instance positif sebenarnya untuk kelas tersebut. Namun, untuk kelas lainnya, model memiliki kinerja yang rendah, seperti yang diindikasikan oleh precision, recall, dan F1-Score yang rendah.

Kesimpulan

Berdasarkan hasil pengujian Hasil klasifikasi menggunakan SVM dengan kernel polinomial menunjukkan presisi yang relatif baik (53%) dan recall yang tinggi (100%) untuk kelas tertentu, sementara hasil klasifikasi menggunakan SVM dengan kernel RBF menunjukkan akurasi sebesar

52.5%. Namun, untuk kelas lainnya, model memiliki kinerja yang rendah, seperti yang diindikasikan oleh precision, recall, dan F1-Score yang rendah.

Saran

Peneliti selanjutnya perlu memilih metode yang tepat untuk mengolah data teks, mengurangi noise, dan menganalisis teks dengan lebih rinci. Selain itu, penting untuk melakukan evaluasi terhadap hasil analisis dan menyajikannya dalam bentuk yang mudah dipahami.

Referensi

- [1] M. Rifqi and F. Ramdhani, "Analisis Sentimen Menggunakan Metode Naive Bayes dan Support Vector Machine pada Ulasan Aplikasi Spotify," 2023.
- [2] G. A. Buntoro, "Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter," 2017. [Online]. Available: <https://t.co/jrvaMsgBdH>
- [3] Kosasih, R., & Alberto, A. (2021). Sentiment analysis of game product on shopee using the TF-IDF method and naive bayes classifier. LKOM Jurnal Ilmiah, 13(2), 101- 109
- [4] Suyanto. 2017. Data mining Untuk Klasifikasi Dan Klasterisasi Data. Bandung: Informatika Bandung.
- [5] Wibawa, A. P., Kurniawan, A. C., Della Murbarani Prawidya Murti, Adiperkasa, R. P., Putra, S. M., Kurniawan, S. A., & Nugraha, Y. R. (2019). Naïve Bayes Classifier for Journal Quartile Classification. Int. J. Recent Contributions Eng. Sci. IT, 7(2), 91-99.
- [6] Pratama, A., Wihandika, R. C., & Ratnawati, D. E. (2018). Implementasi algoritme support vector machine (SVM) untuk prediksi ketepatan waktu kelulusan mahasiswa. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer E-ISSN, 2548, 1704-1708
- [7] Putri, D. A. 2020. Comparison of Naive Bayes Algorithm and Support Vector Machine using PSO Feature Selection for Sentiment Analysis on E-Wallet Review. Journal of Physics: Conference Series, vol. 1641, No. 1, p.012085.
- [8] APJII (Asosiasi Penyedia Jasa Internet Indonesia). 2022. Profil Internet Indonesia 2022. Survei.
- [9] Stephen, Global Streaming music subscription market Q2 2021 MIDiA Research, [Online]. <https://musicindustryblog.wordpress.com/2022/01/18/music-subscriber-market-shares-q2-2021/> [Diakses 25 April 2022].