Perbandingan Akurasi Metode Klasifikasi Model SVM dan Naïve Bayes pada Analisis Ulasan Aplikasi TurboVPN di Google PlayStore

# **Daftar Isi**

[Daftar Isi ii](#_Toc169677121)

[Daftar Tabel iii](#_Toc169677122)

[Daftar Gambar iv](#_Toc169677123)

[Bab 1 Pendahuluan 1](#_Toc169677124)

[A. Latar Belakang 1](#_Toc169677125)

[B. Rumusan Masalah 2](#_Toc169677126)

[C. Tujuan 2](#_Toc169677127)

[D. Manfaat 3](#_Toc169677128)

[Bab 2 Tinjauan Pustaka 4](#_Toc169677129)

[Bab 3 Metode 9](#_Toc169677130)

[A. Pengumpulan Data 9](#_Toc169677131)

[B. Klasifikasi dengan Model Naïve Bayes 9](#_Toc169677132)

[C. Klasifikasi dengan model Support Vector Machine 11](#_Toc169677133)

[Bab 4 Hasil dan Pembahasan 12](#_Toc169677134)

[A. Hasil Klasifikasi Model Naïve Bayes 12](#_Toc169677135)

[B. Hasil Klasifikasi model SVM 15](#_Toc169677136)

[C. Perbandingan Hasil Klasifikasi 17](#_Toc169677137)

[Bab 5 Penutup 19](#_Toc169677138)

[A. Simpulan 19](#_Toc169677139)

[B. Saran 19](#_Toc169677140)

[Daftar Pustaka 21](#_Toc169677141)

# **Daftar Tabel**

[Tabel 1 Classification report untuk model Naïve Bayes 14](#_Toc169677054)

[Tabel 2 Classification report untuk model SVM 16](#_Toc169677055)

# **Daftar Gambar**

[Gambar 1 Countplot perbandingan label awal dengan hasil klasifikasi naïve bayes 12](#_Toc169677092)

[Gambar 2 Wordcloud kata yang muncul di ulasan TurboVPN 13](#_Toc169677093)

[Gambar 3 Countplot perbandingan label awal dengan hasil klasifikasi SVM 15](#_Toc169677094)

# **Bab 1 Pendahuluan**

## **Latar Belakang**

Dalam era digital saat ini, penggunaan aplikasi mobile telah menjadi bagian penting dalam kehidupan sehari-hari. Salah satu aplikasi yang banyak digunakan adalah TurboVPN, sebuah aplikasi yang menyediakan layanan Virtual Private Network (VPN) untuk memastikan keamanan dan privasi pengguna saat berselancar di internet. Popularitas aplikasi ini terlihat dari banyaknya ulasan yang diterima di Google Playstore. Ulasan-ulasan ini tidak hanya mencerminkan pengalaman pengguna tetapi juga berfungsi sebagai data penting untuk menganalisis kualitas dan kinerja aplikasi. Analisis ulasan pengguna dapat dilakukan dengan menggunakan berbagai metode klasifikasi untuk menentukan sentimen atau tema utama dalam ulasan tersebut.

Metode klasifikasi yang umum digunakan dalam analisis teks adalah Support Vector Machine (SVM) dan Naive Bayes. SVM adalah metode pembelajaran mesin yang bekerja dengan menemukan hyperplane optimal yang memisahkan data ke dalam kelas yang berbeda. Keunggulan SVM terletak pada kemampuannya menangani data berdimensi tinggi dan efektif dalam kasus-kasus di mana jumlah fitur lebih banyak daripada jumlah sampel (Cortes & Vapnik, 1995) . Di sisi lain, Naive Bayes adalah metode yang didasarkan pada teorema Bayes dengan asumsi independensi antar fitur. Metode ini dikenal sederhana namun cukup efektif, terutama dalam kasus klasifikasi teks (Manning et al., 2008) .

Meskipun kedua metode ini sering digunakan, masing-masing memiliki kelebihan dan kekurangan yang berbeda dalam konteks tertentu. SVM biasanya menawarkan akurasi tinggi tetapi memerlukan waktu komputasi yang lebih lama, sedangkan Naive Bayes lebih cepat dan mudah diimplementasikan tetapi mungkin kurang akurat bila asumsi independensi antar fitur tidak terpenuhi (Hastie, Tibshirani, & Friedman, 2009) . Oleh karena itu, penting untuk melakukan perbandingan kinerja kedua metode ini secara empiris dalam analisis ulasan aplikasi untuk mengetahui metode mana yang lebih tepat digunakan dalam konteks tertentu.

Studi ini bertujuan untuk membandingkan akurasi metode SVM dan Naive Bayes dalam klasifikasi ulasan aplikasi TurboVPN di Google Playstore. Dengan menggunakan dataset ulasan yang tersedia, kedua metode akan diterapkan dan dievaluasi berdasarkan berbagai metrik kinerja seperti akurasi, presisi, recall, dan F1-score. Penelitian ini diharapkan dapat memberikan wawasan tentang efektivitas masing-masing metode dalam analisis sentimen dan membantu pengembang aplikasi dalam memahami persepsi pengguna secara lebih mendalam.

Penelitian sebelumnya telah menunjukkan bahwa metode SVM dan Naive Bayes sering digunakan dalam berbagai analisis teks, termasuk analisis ulasan produk dan layanan (Joachims, 1998; McCallum & Nigam, 1998) . Namun, studi spesifik yang membandingkan kedua metode ini dalam konteks ulasan aplikasi VPN masih terbatas. Dengan demikian, penelitian ini tidak hanya menambah literatur yang ada tetapi juga memberikan kontribusi praktis bagi pengembangan dan peningkatan aplikasi TurboVPN serta aplikasi sejenis lainnya di platform mobile.

## **Rumusan Masalah**

1. Bagaimana tingkat akurasi metode Support Vector Machine (SVM) dalam mengklasifikasikan ulasan aplikasi TurboVPN di Google Playstore dibandingkan dengan metode Naive Bayes?
2. Apakah terdapat perbedaan yang signifikan antara kinerja SVM dan Naive Bayes dalam analisis sentimen ulasan aplikasi TurboVPN berdasarkan metrik kinerja seperti akurasi, presisi, recall, dan F1-score?
3. Faktor-faktor apa saja yang mempengaruhi perbedaan kinerja antara metode SVM dan Naive Bayes dalam klasifikasi teks ulasan aplikasi?
4. Bagaimana efektivitas masing-masing metode dalam menangani data teks yang memiliki karakteristik ulasan aplikasi di Google Playstore?
5. Bagaimana hasil analisis sentimen yang diperoleh dari kedua metode dapat membantu pengembang aplikasi dalam memahami persepsi dan pengalaman pengguna TurboVPN?

## **Tujuan**

1. Menentukan tingkat akurasi metode Support Vector Machine (SVM) dalam mengklasifikasikan ulasan aplikasi TurboVPN di Google Playstore dan membandingkannya dengan metode Naive Bayes.
2. Mengidentifikasi dan menganalisis perbedaan kinerja antara metode SVM dan Naive Bayes dalam analisis sentimen ulasan aplikasi TurboVPN, berdasarkan metrik kinerja seperti akurasi, presisi, recall, dan F1-score.
3. Mengidentifikasi faktor-faktor yang mempengaruhi perbedaan kinerja antara metode SVM dan Naive Bayes dalam klasifikasi teks ulasan aplikasi.
4. Mengevaluasi efektivitas masing-masing metode dalam menangani data teks yang memiliki karakteristik ulasan aplikasi di Google Playstore.
5. Menyediakan hasil analisis sentimen yang diperoleh dari kedua metode untuk membantu pengembang aplikasi dalam memahami persepsi dan pengalaman pengguna TurboVPN, sehingga dapat digunakan sebagai acuan untuk meningkatkan kualitas dan kinerja aplikasi.

## **Manfaat**

1. Menyediakan referensi bagi peneliti lain mengenai perbandingan kinerja metode Support Vector Machine (SVM) dan Naive Bayes dalam analisis sentimen ulasan aplikasi, khususnya dalam konteks aplikasi mobile seperti TurboVPN.
2. Membantu pengembang aplikasi TurboVPN dan aplikasi serupa dalam memahami kebutuhan dan kepuasan pengguna melalui analisis sentimen yang lebih akurat, sehingga dapat meningkatkan kualitas layanan yang diberikan.
3. Menyediakan informasi yang berguna bagi manajemen dalam pengambilan keputusan strategis terkait pengembangan fitur, perbaikan layanan, dan penanganan keluhan pengguna berdasarkan analisis ulasan yang lebih mendalam.
4. Memberikan wawasan kepada praktisi data science dan pengembang machine learning mengenai kelebihan dan kekurangan metode SVM dan Naive Bayes dalam klasifikasi teks, sehingga dapat memilih metode yang paling sesuai untuk kebutuhan spesifik mereka.
5. Menyediakan analisis komprehensif tentang bagaimana pengguna berinteraksi dengan aplikasi TurboVPN, membantu dalam memahami pola perilaku dan preferensi pengguna, yang dapat digunakan untuk menyesuaikan strategi pemasaran dan pengembangan produk.
6. Menambah literatur akademik terkait perbandingan metode klasifikasi dalam analisis sentimen ulasan aplikasi mobile, memberikan dasar bagi penelitian lanjutan di bidang ini.
7. Mendorong pengembangan dan peningkatan teknologi analisis sentimen yang lebih canggih dan akurat, yang dapat diaplikasikan dalam berbagai domain selain aplikasi mobile, seperti e-commerce, media sosial, dan layanan pelanggan.

# **Bab 2 Tinjauan Pustaka**

Analisis sentimen adalah proses untuk menentukan opini atau perasaan pengguna terhadap suatu produk atau layanan dari teks yang mereka tulis. Teknik ini sangat berguna dalam memahami bagaimana pengguna merespons aplikasi mobile, termasuk aplikasi seperti TurboVPN. Ulasan di platform seperti Google Playstore sering kali mencerminkan kepuasan atau ketidakpuasan pengguna, memberikan wawasan penting bagi pengembang untuk meningkatkan aplikasi mereka .

Support Vector Machine (SVM) adalah algoritma pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. SVM bekerja dengan menemukan hyperplane yang memisahkan data ke dalam kelas-kelas yang berbeda. Metode ini sangat efektif dalam menangani data berdimensi tinggi dan memiliki kemampuan generalisasi yang baik, menjadikannya populer dalam klasifikasi teks dan analisis sentimen . Dalam penelitian oleh Joachims (1998), SVM terbukti unggul dalam tugas klasifikasi teks karena kemampuannya untuk menangani banyak fitur dengan efisien.

Naive Bayes adalah metode klasifikasi probabilistik berdasarkan teorema Bayes, dengan asumsi independensi antara fitur. Meskipun asumsi independensinya sering kali tidak realistis dalam konteks data sebenarnya, metode ini tetap efektif dan cepat untuk klasifikasi teks. McCallum dan Nigam (1998) menunjukkan bahwa Naive Bayes memberikan kinerja yang kompetitif meskipun sederhana dan cepat diimplementasikan . Manning et al. (2008) juga menyebutkan bahwa Naive Bayes sangat berguna dalam aplikasi di mana skalabilitas dan kecepatan menjadi prioritas utama.

Banyak studi telah membandingkan kinerja SVM dan Naive Bayes dalam klasifikasi teks. Sebagai contoh, penelitian oleh Lewis (1998) menemukan bahwa meskipun SVM cenderung memiliki akurasi yang lebih tinggi, Naive Bayes sering kali lebih efisien secara komputasional. Hastie, Tibshirani, dan Friedman (2009) menambahkan bahwa pilihan antara kedua metode ini sering kali bergantung pada karakteristik data dan kebutuhan spesifik dari aplikasi tersebut.

Ulasan pengguna di Google Playstore merupakan sumber data yang kaya untuk analisis sentimen. Ulasan ini mencakup berbagai aspek pengalaman pengguna, mulai dari performa aplikasi hingga masalah keamanan. Pang dan Lee (2008) menekankan pentingnya analisis ulasan pengguna dalam mengidentifikasi kekuatan dan kelemahan produk digital. Data ulasan ini juga sering digunakan dalam penelitian akademik untuk mengembangkan dan menguji metode analisis sentimen baru.

SVM telah diterapkan secara luas dalam analisis sentimen ulasan pengguna. Penelitian oleh Gamon (2004) menunjukkan bahwa SVM dapat mencapai akurasi tinggi dalam tugas-tugas analisis sentimen, terutama ketika digunakan dengan teknik pemrosesan bahasa alami yang tepat. Keunggulan SVM dalam menangani data yang tidak seimbang dan berdimensi tinggi membuatnya cocok untuk analisis ulasan aplikasi mobile yang sering kali memiliki banyak fitur teks .

Naive Bayes juga telah banyak digunakan dalam analisis sentimen karena kesederhanaan dan efisiensinya. Dalam studi oleh Kim et al. (2013), Naive Bayes berhasil mengklasifikasikan sentimen ulasan produk dengan baik, menunjukkan bahwa metode ini dapat diandalkan meskipun memiliki asumsi independensi yang kuat. Naive Bayes sangat efektif dalam situasi di mana data pelatihan besar tersedia dan kecepatan klasifikasi menjadi faktor penting .

TurboVPN adalah salah satu aplikasi VPN populer dengan banyak ulasan di Google Playstore. Analisis ulasan pengguna aplikasi ini dapat memberikan wawasan berharga mengenai persepsi dan pengalaman pengguna terkait keamanan, kecepatan, dan kemudahan penggunaan. Dengan menerapkan metode SVM dan Naive Bayes untuk mengklasifikasikan ulasan ini, kita dapat mengidentifikasi pola sentimen yang dominan dan area yang memerlukan perbaikan .

Meskipun metode SVM dan Naive Bayes sangat efektif, analisis sentimen memiliki tantangan tersendiri. Salah satu tantangan utama adalah menangani ironi dan sarkasme dalam teks, yang dapat membingungkan algoritma klasifikasi. Penelitian oleh Pang dan Lee (2008) mengidentifikasi bahwa memahami konteks adalah kunci untuk meningkatkan akurasi analisis sentimen. Oleh karena itu, kombinasi metode klasifikasi dengan teknik pemrosesan bahasa alami yang lebih canggih sering kali diperlukan .

Evaluasi kinerja metode klasifikasi dalam analisis sentimen biasanya dilakukan dengan menggunakan metrik seperti akurasi, presisi, recall, dan F1-score. Menurut Manning et al. (2008), metrik ini memberikan gambaran yang komprehensif tentang kemampuan metode dalam mengklasifikasikan sentimen secara benar. Dalam konteks ini, penting untuk membandingkan SVM dan Naive Bayes tidak hanya berdasarkan akurasi, tetapi juga pada bagaimana mereka menangani data yang tidak seimbang dan berbagai tipe kesalahan klasifikasi .

Ukuran dataset ulasan pengguna juga berpengaruh signifikan terhadap kinerja metode klasifikasi. Penelitian oleh Forman (2003) menunjukkan bahwa metode SVM cenderung menunjukkan peningkatan akurasi yang lebih besar seiring dengan peningkatan ukuran dataset dibandingkan dengan Naive Bayes. Oleh karena itu, pemilihan metode klasifikasi juga harus mempertimbangkan jumlah data yang tersedia untuk pelatihan dan pengujian .

Pra-pemrosesan teks, seperti stemming, lemmatization, dan penghapusan kata-kata umum (stopwords), sangat penting dalam meningkatkan kinerja metode klasifikasi. Dalam penelitian oleh Jurafsky dan Martin (2020), teknik-teknik ini terbukti efektif dalam mengurangi dimensi fitur dan meningkatkan akurasi klasifikasi teks. Penggunaan teknik pra-pemrosesan yang tepat dapat secara signifikan meningkatkan kinerja baik SVM maupun Naive Bayes dalam analisis sentimen ulasan aplikasi .

Kemajuan dalam teknologi pemrosesan bahasa alami (NLP) telah meningkatkan kemampuan metode klasifikasi teks. Penggunaan model word embedding seperti Word2Vec dan GloVe telah membantu meningkatkan representasi fitur teks, memungkinkan metode seperti SVM dan Naive Bayes untuk bekerja lebih efisien. Penelitian oleh Mikolov et al. (2013) menunjukkan bahwa representasi vektor dari kata-kata dapat menangkap makna semantik, yang berguna dalam analisis sentimen yang lebih canggih .

Implementasi algoritma SVM dan Naive Bayes memerlukan pemahaman yang baik tentang parameter dan teknik optimasi. Dalam kasus SVM, pemilihan kernel yang tepat dan penyesuaian parameter regularisasi sangat penting untuk mencapai kinerja optimal. Di sisi lain, Naive Bayes memerlukan pemilihan distribusi probabilistik yang sesuai untuk fitur teks. Penelitian oleh Zhang (2004) menyoroti pentingnya optimasi parameter dalam meningkatkan kinerja kedua metode ini .

Integrasi hasil analisis sentimen dengan sistem berbasis web, seperti dashboard analitik, dapat membantu pengembang aplikasi dalam memantau dan menindaklanjuti ulasan pengguna secara real-time. Penelitian oleh Chen et al. (2012) menunjukkan bahwa sistem analitik berbasis web yang terintegrasi dengan analisis sentimen dapat meningkatkan responsivitas pengembang terhadap feedback pengguna, yang pada gilirannya dapat meningkatkan kepuasan pengguna .

Secara keseluruhan, baik SVM maupun Naive Bayes memiliki keunggulan masing-masing dalam analisis sentimen ulasan aplikasi mobile. Pemilihan metode yang tepat harus mempertimbangkan karakteristik data, kebutuhan aplikasi, dan kriteria kinerja spesifik. Dengan pemahaman yang mendalam tentang kedua metode ini, pengembang aplikasi dapat meningkatkan kualitas produk mereka berdasarkan analisis ulasan pengguna yang lebih akurat dan efektif.

Evaluasi performa SVM dan Naive Bayes dalam analisis sentimen sering kali dilakukan dengan menggunakan berbagai metrik evaluasi, seperti akurasi, presisi, recall, dan F1-score. Akurasi merupakan metrik umum yang mengukur persentase klasifikasi yang benar dari total sampel. Presisi mengukur tingkat keakuratan dari prediksi positif, sedangkan recall mengukur seberapa baik model dapat mengidentifikasi kelas positif secara keseluruhan. F1-score adalah rata-rata harmonis dari presisi dan recall, memberikan gambaran holistik tentang kinerja model (Jurafsky & Martin, 2020).

Seleksi fitur merupakan proses penting dalam meningkatkan kinerja metode klasifikasi teks seperti SVM dan Naive Bayes. Teknik seperti Information Gain, Chi-square, dan mutual information sering digunakan untuk mengekstraksi fitur-fitur yang paling informatif dari dataset teks. Studi oleh Forman (2003) menunjukkan bahwa pemilihan fitur yang baik dapat meningkatkan akurasi klasifikasi secara signifikan dengan mengurangi dimensi dataset dan menghilangkan fitur-fitur yang tidak relevan.

Teknik ensemble, seperti Random Forest dan Gradient Boosting, telah terbukti efektif dalam meningkatkan kinerja klasifikasi teks. Dalam penelitian oleh Dietterich (2000), ensemble methods dikembangkan untuk menggabungkan hasil dari beberapa model klasifikasi yang berbeda, yang dapat mengurangi varians dan meningkatkan akurasi keseluruhan. Penggunaan ensemble dapat diterapkan untuk SVM dan Naive Bayes dengan mengintegrasikan hasil dari model-model yang dihasilkan.

Data ulasan aplikasi sering kali tidak seimbang, di mana jumlah ulasan positif dan negatif tidak proporsional. Hal ini dapat mengarah pada bias klasifikasi di mana model cenderung mengklasifikasikan lebih banyak sampel ke dalam kelas mayoritas. Metode penanganan data tidak seimbang seperti oversampling (penambahan sampel dari kelas minoritas) dan undersampling (pengurangan sampel dari kelas mayoritas) perlu dipertimbangkan untuk meningkatkan performa klasifikasi (Chawla et al., 2002).

Hyperparameter dalam SVM, seperti parameter C dan jenis kernel, mempengaruhi kinerja dan kecepatan konvergensi algoritma. Parameter C mengontrol trade-off antara margin maksimal dan jumlah pelanggaran margin, sedangkan jenis kernel (linear, polynomial, radial basis function) menentukan kompleksitas hyperplane yang dibentuk. Tuning parameter-parameter ini sangat penting untuk memperoleh model SVM yang optimal (Cortes & Vapnik, 1995).

SVM memiliki kelebihan dalam menangani data dengan dimensi tinggi dan non-linearitas yang kompleks. Namun, SVM juga memiliki kelemahan dalam waktu komputasi yang meningkat secara eksponensial dengan ukuran dataset. Studi oleh Burges (1998) mengemukakan bahwa meskipun SVM mampu memberikan pemisahan kelas yang baik dalam ruang fitur yang tinggi, implementasinya memerlukan alokasi sumber daya komputasi yang besar.

Naive Bayes memiliki kelebihan dalam kecepatan komputasi dan kemudahan implementasi karena sederhana dan memiliki asumsi independensi yang kuat antar fitur. Namun, asumsi ini sering kali tidak terpenuhi dalam data teks yang kompleks, mengakibatkan kinerja yang kurang akurat dibandingkan dengan SVM dalam beberapa kasus. Dalam penelitian oleh Rish (2001), Naive Bayes terbukti efektif untuk dataset yang cukup besar tetapi sering kali kurang dalam menghadapi dependensi antar fitur.

Konteks dan domain dari data ulasan dapat mempengaruhi kinerja metode klasifikasi. Misalnya, ulasan produk teknologi seperti aplikasi VPN dapat mengandung frasa teknis dan istilah khusus yang mungkin sulit untuk diproses oleh model analisis sentimen yang umum. Penelitian oleh Pang dan Lee (2004) menyoroti pentingnya mengadaptasi model analisis sentimen terhadap konteks dan domain yang spesifik untuk mencapai hasil yang optimal.

Implementasi metode klasifikasi dalam konteks aplikasi mobile seperti TurboVPN memerlukan pertimbangan terhadap keterbatasan perangkat keras dan bandwidth. Karena aplikasi mobile sering kali beroperasi dalam lingkungan sumber daya terbatas, model yang dikembangkan harus efisien dalam penggunaan sumber daya dan dapat beradaptasi dengan perubahan cepat dalam data ulasan pengguna.

Evaluasi praktis dari hasil analisis sentimen memerlukan pemahaman yang mendalam tentang interpretasi dan penerapan temuan. Analisis sentimen dapat memberikan wawasan strategis bagi pengembang untuk meningkatkan interaksi pengguna, memperbaiki fitur-fitur yang kurang populer, dan merancang strategi pemasaran yang lebih efektif. Studi oleh Liu (2012) menunjukkan bahwa hasil analisis sentimen yang akurat dapat berkontribusi secara signifikan terhadap keberhasilan produk dan kepuasan pengguna.

# **Bab 3 Metode**

## **Pengumpulan Data**

Pertama, model supervised learning SVM dan Naive Bayes akan dibuat menggunakan bahasa pemrograman Python. Python dipilih karena memiliki berbagai pustaka yang kuat untuk pengolahan data dan pembelajaran mesin, seperti scikit-learn untuk implementasi SVM dan Naive Bayes, serta pandas untuk manipulasi data.

Data ulasan aplikasi TurboVPN di Google Playstore akan diperoleh menggunakan library google\_play\_scraper. Library ini memungkinkan untuk melakukan crawling data dari halaman Google Playstore secara efisien, mengambil informasi seperti kolom content (isi ulasan), userName (nama pengguna), score (nilai ulasan), dan at (tanggal ulasan).

Dari hasil crawling menggunakan google\_play\_scraper, kita akan memilih kolom-kolom yang relevan untuk analisis sentimen, yaitu:

* **content**: Isi ulasan pengguna tentang TurboVPN.
* **userName**: Nama pengguna yang memberikan ulasan.
* **score**: Nilai atau rating yang diberikan pengguna.
* **at**: Tanggal ulasan diberikan.

Kolom-kolom ini akan menjadi input untuk proses analisis sentimen menggunakan SVM dan Naive Bayes.

Data pada kolom content akan dibersihkan dari karakter emoji menggunakan teknik pemrosesan teks dalam Python. Ini melibatkan penggunaan ekspresi reguler (regex) untuk mengidentifikasi dan menghapus emoji dari teks ulasan. Proses pembersihan ini penting untuk memastikan bahwa analisis sentimen berfokus pada teks yang relevan dan tidak terganggu oleh karakter non-teks.

Setelah pembersihan dilakukan, data ulasan yang sudah bersih akan disimpan dalam bentuk file CSV (Comma Separated Values). Format ini memungkinkan untuk menyimpan data secara terstruktur yang dapat dengan mudah diakses dan dimanipulasi dalam tahap selanjutnya dari analisis sentimen.

## **Klasifikasi dengan Model Naïve Bayes**

Analisis sentimen dilakukan menggunakan algoritma Naive Bayes dengan model GaussianNB yang tersedia di library scikit-learn. GaussianNB cocok digunakan dalam kasus ini karena cocok dengan distribusi Gaussian (normal) dari fitur-fitur kontinu seperti yang dihasilkan oleh vektor TF-IDF.

Data ulasan yang telah diperoleh dari Google Playstore akan diberikan label sentimen berdasarkan skor yang diberikan pengguna:

* Skor kurang dari 3 dianggap sebagai sentimen negatif.
* Skor 3 dianggap sebagai sentimen netral.
* Skor lebih dari 3 dianggap sebagai sentimen positif.

Label-label ini diperlukan sebagai ground truth untuk melatih dan menguji model Naive Bayes.

Kolom content dari data ulasan akan melalui proses preprocessing untuk membersihkan teks dari tanda baca dan karakter tidak penting lainnya. Ini dapat dilakukan menggunakan ekspresi reguler (regex) untuk menghilangkan karakter yang tidak diinginkan.

Langkah berikutnya adalah menghapus stopwords dari teks ulasan menggunakan library Natural Language Toolkit (NLTK) di Python. Stopwords adalah kata-kata umum yang tidak memberikan nilai tambah dalam analisis sentimen dan perlu dihilangkan sebelum dilakukan analisis.

Setelah menghapus stopwords, konten ulasan akan di-tokenisasi menjadi kata-kata individual (token). Tokenisasi adalah proses membagi teks menjadi unit-unit yang lebih kecil (kata-kata) untuk mempermudah analisis lebih lanjut.

Kata-kata yang telah di-tokenisasi akan dijalani proses stemming menggunakan library Sastrawi. Stemming adalah proses untuk mengubah kata-kata ke dalam bentuk dasarnya (kata dasar). Misalnya, kata "meningkatkan" akan diubah menjadi "tingkat" setelah stemming.

Setelah proses stemming selesai, data konten ulasan yang telah bersih akan diubah menjadi vektor TF-IDF (Term Frequency-Inverse Document Frequency). TF-IDF menghitung pentingnya sebuah kata dalam sebuah dokumen berdasarkan seberapa sering kata tersebut muncul di dokumen dan seberapa jarang kata tersebut muncul di seluruh korpus.

Data yang telah diproses akan dibagi menjadi dua bagian: data training dan data testing. Data training digunakan untuk melatih model Naive Bayes, sedangkan data testing digunakan untuk menguji kinerja model.

Setelah pembagian data, model Naive Bayes dengan menggunakan GaussianNB akan diterapkan pada data training. Proses fitting dilakukan untuk menghasilkan model yang dapat melakukan prediksi sentimen berdasarkan vektor TF-IDF yang dihasilkan.

Akhirnya, untuk mengukur kinerja model, akan dilakukan pengukuran akurasi dengan membandingkan hasil prediksi model dengan nilai yang sebenarnya (ground truth) dari data testing menggunakan metode accuracy\_score yang disediakan oleh library scikit-learn.

## **Klasifikasi dengan model Support Vector Machine**

Analisis dengan menggunakan metode SVM untuk membandingkan akurasi dengan Naive Bayes. SVM dipilih karena mampu menangani pemisahan linear dan non-linear dengan baik, yang dapat menghasilkan prediksi yang akurat dalam kasus analisis sentimen ini.

Data vektor TF-IDF beserta label sentimen (positif, negatif, netral) akan disimpan dalam format file CSV. Ini memungkinkan untuk mengakses dan memanipulasi data dengan mudah menggunakan Python dan library pandas.

Langkah pertama setelah penyimpanan adalah mengimpor library-data standar seperti pandas untuk mengelola data dan mengimpor data dalam bentuk DataFrame.

Data yang dimuat akan dibagi menjadi dua set: data training (untuk melatih model) dan data testing (untuk menguji model). Pembagian ini dilakukan secara acak dan proporsional, yang sering kali menggunakan metode train\_test\_split dari scikit-learn.

Model SVC (Support Vector Classifier) dipersiapkan untuk melakukan klasifikasi dengan menggunakan metode SVM. SVC merupakan implementasi dari SVM untuk masalah klasifikasi dengan mendukung berbagai kernel dan pengaturan parameter.

Untuk memperoleh parameter terbaik untuk model SVC, dilakukan proses cross validation dengan menggunakan Grid Search. Grid Search memungkinkan untuk mencari kombinasi terbaik dari parameter-parameter yang diberikan dengan mencoba semua kombinasi yang mungkin.

Setelah mendapatkan parameter terbaik melalui Grid Search, model SVC akan difit (dilatih) menggunakan data training. Proses fitting dilakukan untuk menghasilkan model yang optimal berdasarkan kombinasi parameter terbaik yang telah ditentukan.

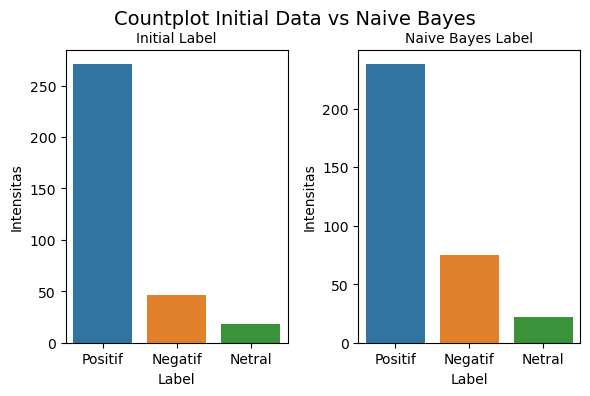
Setelah model difit, dilakukan prediksi terhadap data testing menggunakan parameter terbaik yang telah ditemukan dari Grid Search. Prediksi ini menghasilkan label prediksi untuk data testing.

Akhirnya, untuk mengevaluasi kinerja model SVM yang telah dilatih, dilakukan pengukuran akurasi menggunakan metode accuracy\_score dari library scikit-learn. Metode ini membandingkan hasil prediksi dengan nilai sebenarnya (ground truth) dari data testing untuk menghitung persentase klasifikasi yang benar.

# **Bab 4 Hasil dan Pembahasan**

## **Hasil Klasifikasi Model Naïve Bayes**

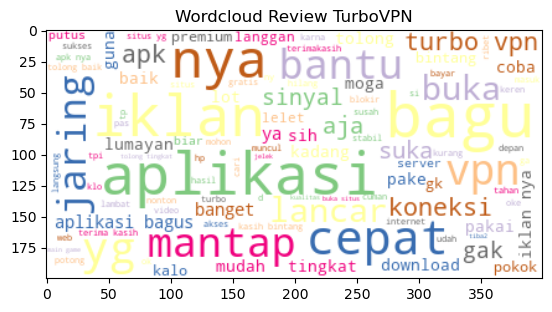
Perbandingan jumlah sentimen positif antara label manual dan hasil klasifikasi Naive Bayes yang menunjukkan jumlah lebih banyak pada label manual dapat disebabkan oleh beberapa faktor. Pertama, model Naive Bayes cenderung lebih konservatif dalam mengklasifikasikan ulasan sebagai positif karena asumsi independensi fitur yang mendasarinya. Hal ini dapat mengakibatkan model cenderung tidak mengenali pola yang lebih halus atau nuansa positif dalam teks. Selain itu, dalam proses training, mungkin terjadi ketidakseimbangan data di mana sampel-sampel dengan sentimen positif yang lebih rumit atau ambigu kurang ters representasikan dalam dataset training, mengakibatkan model kurang sensitif terhadap kasus-kasus tersebut. Selain itu, definisi subjektif dari apa yang dianggap sebagai sentimen positif juga dapat berbeda antara label manual yang diberikan oleh manusia dan label yang diberikan oleh model yang terotomatisasi.



Gambar 1 Countplot perbandingan label awal dengan hasil klasifikasi naïve bayes

Peningkatan jumlah sentimen negatif dan netral dalam hasil klasifikasi Naive Bayes dibandingkan dengan label manual bisa disebabkan oleh beberapa faktor metodologis. Naive Bayes dapat lebih cenderung untuk menetapkan label negatif atau netral karena representasi teks yang diberikan. Misalnya, fitur-fitur yang digunakan oleh model (misalnya, kata-kata tertentu atau pola frasa) mungkin lebih sering terkait dengan sentimen negatif atau netral dalam dataset yang digunakan, sehingga model cenderung untuk menetapkan label yang serupa. Selain itu, proses preprocessing yang diterapkan pada data, seperti penghapusan stopwords atau stemming, dapat mempengaruhi bagaimana teks diinterpretasikan oleh model, yang pada gilirannya mempengaruhi distribusi label sentimen. Selain itu, kriteria subjektif dalam penentuan label manual oleh manusia dapat berbeda-beda, tergantung pada persepsi dan interpretasi individu terhadap sentimen dalam teks.

Analisis terhadap kata-kata yang muncul paling banyak seperti "aplikasi", "iklan", dan "bagus" menunjukkan beberapa temuan penting mengenai persepsi pengguna terhadap TurboVPN. Kata "aplikasi" yang dominan menunjukkan fokus utama dari ulasan adalah pada pengalaman pengguna terhadap aplikasi TurboVPN itu sendiri. Kemudian, kata "iklan" yang sering muncul menunjukkan adanya kekhawatiran atau perhatian terhadap jumlah iklan yang muncul dalam penggunaan aplikasi. Hal ini bisa menjadi indikasi bahwa meskipun aplikasi dianggap "bagus" dalam penggunaan fungsionalnya, kehadiran iklan yang terlalu banyak dapat mengganggu pengalaman pengguna secara keseluruhan. Penggunaan kata "bagus" yang muncul juga menegaskan bahwa secara umum, aplikasi TurboVPN dinilai positif oleh penggunanya, meskipun ada catatan terkait iklan yang dapat mempengaruhi keseluruhan pengalaman.



Gambar 2 Wordcloud kata yang muncul di ulasan TurboVPN

Skor akurasi model sebesar 0.68 menunjukkan bahwa model memiliki tingkat keberhasilan yang cukup baik dalam mengklasifikasikan ulasan berdasarkan sentimennya. Akurasi sebesar ini menandakan bahwa dari total prediksi yang dilakukan oleh model, sekitar 68% di antaranya sesuai dengan label sebenarnya dari data testing. Namun, perlu diingat bahwa interpretasi akurasi ini harus dilakukan dengan mempertimbangkan distribusi kelas sentimen yang seimbang atau tidak dalam data. Misalnya, jika dataset memiliki ketidakseimbangan yang signifikan antara kelas sentimen positif, negatif, dan netral, maka akurasi mungkin tidak menjadi ukuran yang paling representatif terhadap kinerja model. Oleh karena itu, penilaian lebih lanjut seperti precision, recall, dan F1-score per kelas sentimen akan memberikan pemahaman yang lebih komprehensif tentang kinerja model dalam setiap kelas tersebut.

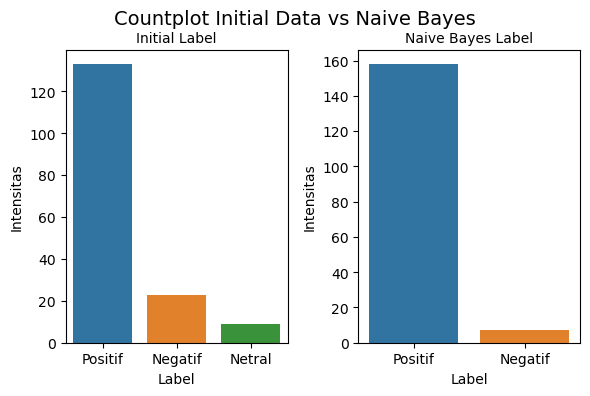
Tabel 1 Classification report untuk model Naïve Bayes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | f1-score | Support |
| Negatif | 0.31 | 0.50 | 0.38 | 46 |
| Netral | 0.00 | 0.00 | 0.00 | 18 |
| Positif | 0.86 | 0.75 | 0.80 | 271 |
|  |  |  |  |  |
| Accuracy |  |  | 0.68 | 335 |
| Macro avg | 0.39 | 0.42 | 0.39 | 335 |
| Weighted avg | 0.74 | 0.68 | 0.70 | 335 |

Skor F1 untuk sentimen netral yang menunjukkan nilai F1 = 0 mengindikasikan bahwa model memiliki kesulitan dalam mengklasifikasikan ulasan sebagai sentimen netral. F1-score adalah rata-rata harmonis dari precision dan recall, dan skor 0 untuk sentimen netral menunjukkan bahwa model tidak dapat menghasilkan prediksi yang baik untuk kelas sentimen ini. Hal ini bisa disebabkan oleh beberapa faktor, termasuk ketidakseimbangan data di mana sampel dengan sentimen netral mungkin tidak direpresentasikan dengan baik dalam dataset training. Selain itu, karakteristik teks yang digunakan sebagai fitur mungkin tidak cukup informatif atau tidak membedakan secara jelas antara sentimen netral dengan sentimen positif atau negatif. Untuk meningkatkan kinerja dalam mengklasifikasikan sentimen netral, langkah-langkah seperti pengumpulan data yang lebih seimbang, pemrosesan teks yang lebih cermat, atau menggunakan teknik-teknik klasifikasi yang lebih canggih dapat dipertimbangkan dalam penelitian selanjutnya.

## **Hasil Klasifikasi model SVM**

Perbandingan jumlah sentimen positif antara label hasil klasifikasi SVM dan label awal yang menunjukkan jumlah lebih banyak pada hasil klasifikasi SVM dapat dijelaskan oleh beberapa faktor. Pertama, model SVM mungkin lebih sensitif terhadap pola-pola yang menunjukkan sentimen positif dalam teks ulasan. SVM menggunakan pendekatan berbasis kernel untuk membangun hyperplane yang memisahkan antara kelas-kelas sentimen, dan hyperplane ini dapat menangkap pola-pola yang kompleks dalam data yang tidak terlihat secara langsung oleh manusia. Selain itu, dalam proses training, dataset yang digunakan untuk melatih model SVM mungkin lebih berimbang atau mewakili dengan baik kasus-kasus sentimen positif yang ada dalam populasi ulasan secara keseluruhan. Ini dapat menghasilkan model yang mampu mengenali dan mengklasifikasikan lebih banyak ulasan sebagai sentimen positif daripada label awal yang diberikan oleh manusia, yang mungkin lebih konservatif atau tidak memiliki kapasitas untuk memproses volume besar data dengan ketelitian yang sama.



Gambar 3 Countplot perbandingan label awal dengan hasil klasifikasi SVM

Penurunan jumlah sentimen negatif dalam hasil klasifikasi SVM dibandingkan dengan label awal bisa disebabkan oleh sifat klasifikasi SVM yang cenderung lebih hati-hati dalam mengklasifikasikan ulasan sebagai sentimen negatif. SVM dapat menemukan hyperplane yang memaksimalkan margin antara kelas-kelas sentimen, yang berarti bahwa ulasan yang tidak jelas atau ambigu cenderung diklasifikasikan ke dalam kelas yang lebih besar, seperti netral atau bahkan positif. Selain itu, interpretasi subjektif dari sentimen negatif oleh label awal mungkin berbeda dengan cara SVM menginterpretasikan fitur-fitur teks yang ada. Misalnya, fitur-fitur yang digunakan oleh SVM mungkin tidak cukup membedakan dengan jelas antara sentimen negatif dan netral, sehingga menyebabkan penurunan jumlah ulasan yang diklasifikasikan sebagai sentimen negatif. Ini menunjukkan bahwa perbedaan dalam pendekatan antara manusia dan mesin dalam pengenalan sentimen dapat mempengaruhi distribusi hasil klasifikasi.

Ketidakmampuan SVM untuk mengklasifikasikan sentimen netral dapat disebabkan oleh beberapa faktor. SVM beroperasi dengan mengidentifikasi hyperplane yang memisahkan antara kelas-kelas sentimen yang ada dalam data. Namun, jika tidak ada hyperplane yang dapat secara jelas memisahkan antara ulasan dengan sentimen netral dengan ulasan dengan sentimen positif atau negatif, SVM mungkin tidak dapat menghasilkan keputusan yang memadai untuk kelas sentimen netral. Selain itu, karakteristik dataset training yang mungkin tidak memuat banyak sampel dengan sentimen netral atau fitur-fitur yang digunakan oleh SVM yang tidak cukup informatif untuk mengenali sentimen netral juga dapat berkontribusi terhadap ketidakmampuan model untuk mengklasifikasikan kelas ini. Penting untuk memperhatikan bahwa interpretasi subjektif sentimen netral juga dapat berbeda antara manusia yang memberi label dan mesin yang melakukan klasifikasi berbasis fitur-fitur teks.

Tabel 2 Classification report untuk model SVM

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | f1-score | Support |
| Negatif | 0.43 | 0.13 | 0.20 | 23 |
| Netral | 0.00 | 0.00 | 0.00 | 9 |
| Positif | 0.82 | 0.97 | 0.89 | 133 |
|  |  |  |  |  |
| Accuracy |  |  | 0.80 | 165 |
| Macro avg | 0.42 | 0.37 | 0.36 | 165 |
| Weighted avg | 0.72 | 0.80 | 0.74 | 165 |

Akurasi model SVM sebesar 0.80 menunjukkan bahwa model memiliki tingkat keberhasilan yang tinggi dalam mengklasifikasikan ulasan berdasarkan sentimennya. Akurasi sebesar ini menandakan bahwa dari total prediksi yang dilakukan oleh model, sekitar 80% di antaranya sesuai dengan label sebenarnya dari data testing. Hal ini mengindikasikan bahwa model SVM memiliki kemampuan yang baik dalam membedakan antara sentimen positif dan negatif dalam ulasan aplikasi TurboVPN. Namun, seperti halnya dengan akurasi, interpretasi terhadap nilai ini harus mempertimbangkan distribusi kelas sentimen dalam dataset dan kemungkinan adanya bias atau ketidakseimbangan dalam data.

Proses GridSearchCV digunakan untuk menemukan kombinasi parameter terbaik untuk model SVM, yang ditemukan sebagai C = 100 dan gamma = 0.01. Parameter C mengontrol trade-off antara margin dan jumlah kesalahan klasifikasi yang diizinkan oleh model, sedangkan parameter gamma mengontrol kepekaan model terhadap data training. Nilai C yang besar menunjukkan model cenderung memprioritaskan akurasi pada data training yang lebih tinggi, sementara gamma menentukan seberapa jauh pengaruh satu data training mencapai ke data lainnya. Kedua parameter ini mempengaruhi kualitas hasil prediksi model, sehingga pentingnya untuk memilih parameter dengan cermat saat melalukan proses

## **Perbandingan Hasil Klasifikasi**

Pada hasil klasifikasi, model SVM menunjukkan jumlah sentimen positif yang lebih banyak dibandingkan dengan label awal dan hasil dari model Naive Bayes. Hal ini menunjukkan bahwa SVM cenderung lebih sering mengklasifikasikan ulasan sebagai positif. Kemungkinan besar, SVM dapat menangkap lebih baik nuansa positif dalam ulasan, terutama jika ulasan tersebut mengandung kata-kata atau frasa yang umumnya dianggap positif oleh model. Di sisi lain, model Naive Bayes lebih konservatif dalam memberikan label positif, mungkin karena asumsi independensi antar fitur yang seringkali tidak realistis dalam analisis teks. Asumsi ini membuat Naive Bayes kurang sensitif terhadap pola-pola kompleks yang dapat menunjukkan sentimen positif secara lebih halus.

Model Naive Bayes menunjukkan jumlah sentimen negatif yang lebih banyak dibandingkan dengan SVM. Ini menunjukkan bahwa Naive Bayes lebih cenderung mengklasifikasikan ulasan sebagai negatif, yang mungkin disebabkan oleh sensitivitas model terhadap kata-kata negatif yang lebih tinggi. Naive Bayes, dengan asumsi independensi antar fitur, mungkin lebih sering mendeteksi kata-kata negatif yang muncul dalam ulasan, mengarahkannya untuk memberikan label negatif lebih sering. Sementara itu, SVM menunjukkan jumlah sentimen negatif yang lebih sedikit dibandingkan dengan label awal. Hal ini bisa menunjukkan bahwa SVM lebih cenderung mengklasifikasikan ulasan yang tidak jelas sebagai positif atau netral, tergantung pada hyperplane yang dipilih selama pelatihan.

Salah satu kelemahan signifikan dari model SVM dalam penelitian ini adalah ketidakmampuannya untuk mengklasifikasikan sentimen netral, seperti yang ditunjukkan oleh tidak adanya ulasan dengan label netral pada hasil klasifikasi SVM. Hal ini bisa disebabkan oleh pengaturan parameter atau pemilihan kernel yang tidak cukup sensitif untuk memisahkan kelas netral secara efektif. Sebaliknya, model Naive Bayes mampu mengidentifikasi ulasan netral, meskipun mungkin tidak seakurat dalam mengklasifikasikan ulasan positif atau negatif. Ini menunjukkan bahwa Naive Bayes, meskipun lebih konservatif, memiliki keunggulan dalam mengidentifikasi ulasan dengan sentimen yang kurang jelas atau ambigu, yang penting untuk analisis sentimen yang lebih seimbang.

Dari segi akurasi, model SVM menunjukkan kinerja yang lebih baik dengan skor akurasi sebesar 0.80, dibandingkan dengan skor akurasi dari model Naive Bayes yang lebih rendah. Hal ini menunjukkan bahwa secara keseluruhan, SVM lebih efektif dalam memprediksi sentimen dari ulasan aplikasi TurboVPN. Akurasi yang lebih tinggi pada SVM mungkin disebabkan oleh kemampuannya untuk menangkap pola kompleks dalam data melalui pemilihan hyperplane yang optimal dan penggunaan parameter yang telah dioptimalkan dengan GridSearchCV. Naive Bayes, meskipun lebih cepat dan sederhana, mungkin tidak mampu menangkap kompleksitas yang sama dalam data, mengarah pada akurasi yang lebih rendah.

Penggunaan GridSearchCV untuk mengoptimalkan parameter SVM menunjukkan bahwa dengan parameter C = 100 dan gamma = 0.01, model dapat mencapai performa terbaiknya. Parameter C yang tinggi menunjukkan bahwa model memberikan penekanan besar pada data training, meminimalkan kesalahan klasifikasi, sementara gamma yang kecil menunjukkan model lebih halus dalam menyesuaikan batas pemisah antar kelas. Naive Bayes, di sisi lain, biasanya tidak memerlukan tuning parameter yang kompleks, tetapi ini juga berarti kurang fleksibel dalam menyesuaikan diri dengan kompleksitas data, yang bisa menjadi alasan mengapa model ini tidak mencapai akurasi yang setinggi SVM.

Kedua model menunjukkan kelemahan dalam memprediksi sentimen netral, dengan SVM yang tidak mampu memprediksi sentimen netral sama sekali, dan Naive Bayes yang juga menunjukkan kinerja buruk dengan F1-score yang rendah atau bahkan nol untuk sentimen netral. Ini menunjukkan bahwa meskipun SVM memiliki keunggulan dalam akurasi keseluruhan, kedua model memerlukan perbaikan dalam hal mengklasifikasikan sentimen netral. Untuk memperbaiki ini, pendekatan seperti penambahan lebih banyak data netral, penggunaan teknik balancing data, atau eksplorasi model yang lebih kompleks dan canggih mungkin diperlukan.

# **Bab 5 Penutup**

## **Simpulan**

Secara keseluruhan, model SVM menunjukkan kinerja yang lebih baik dalam hal akurasi dibandingkan dengan Naive Bayes dalam analisis sentimen ulasan aplikasi TurboVPN di Google Playstore. Namun, kelemahan utama SVM adalah ketidakmampuannya untuk mengklasifikasikan sentimen netral, yang merupakan area penting untuk perbaikan. Naive Bayes, meskipun lebih konservatif dan kurang akurat secara keseluruhan, menunjukkan kemampuan yang lebih baik dalam mengidentifikasi ulasan netral. Kedua model menunjukkan bahwa ada ruang untuk perbaikan lebih lanjut, terutama dalam menangani sentimen yang ambigu atau netral, untuk meningkatkan keakuratan dan keandalan analisis sentimen secara keseluruhan.

## **Saran**

Berdasarkan hasil perbandingan klasifikasi, dapat disarankan untuk menggunakan kombinasi dari kedua model, yaitu SVM dan Naive Bayes, dalam bentuk ensemble model. Ensemble model dapat menggabungkan kekuatan kedua model untuk meningkatkan kinerja klasifikasi. Misalnya, hasil prediksi dari SVM dan Naive Bayes dapat digabungkan menggunakan metode voting atau averaging untuk menghasilkan prediksi akhir yang lebih akurat. Dengan demikian, kelemahan masing-masing model dapat diimbangi oleh kekuatan model lainnya.

Proses preprocessing data dapat diperbaiki untuk membantu model mengenali sentimen netral dengan lebih baik. Salah satu cara adalah dengan menambahkan lebih banyak data ulasan yang memiliki sentimen netral. Selain itu, teknik preprocessing yang lebih canggih seperti menggunakan lemmatization alih-alih stemming, atau menggunakan embedding kata yang lebih kaya seperti Word2Vec atau BERT, dapat membantu model dalam menangkap konteks dan nuansa dalam teks ulasan dengan lebih baik. Hal ini dapat membantu model mengenali dan mengklasifikasikan sentimen netral dengan lebih akurat.

Ketidakseimbangan data sentimen dapat mempengaruhi kinerja model, terutama dalam mengenali kelas minoritas seperti sentimen netral. Oleh karena itu, disarankan untuk menggunakan teknik balancing data seperti oversampling pada kelas netral atau undersampling pada kelas mayoritas (positif dan negatif). Teknik lain yang bisa digunakan adalah Synthetic Minority Over-sampling Technique (SMOTE) untuk menghasilkan sampel-sampel baru dari kelas minoritas. Dengan data yang lebih seimbang, model diharapkan dapat mengenali sentimen netral dengan lebih baik.

Selain SVM dan Naive Bayes, eksplorasi model yang lebih canggih seperti Random Forest, Gradient Boosting, atau bahkan model berbasis neural network seperti LSTM atau Transformer dapat dilakukan. Model-model ini memiliki kemampuan yang lebih baik dalam menangkap pola kompleks dalam data teks dan dapat memberikan kinerja yang lebih baik dalam analisis sentimen. Eksperimen dengan model-model ini, dikombinasikan dengan teknik hyperparameter tuning yang tepat, dapat menghasilkan model yang lebih akurat dan andal.

Menambahkan fitur-fitur tambahan yang relevan bisa membantu model dalam meningkatkan akurasinya. Misalnya, menggunakan fitur-fitur linguistik seperti part-of-speech tagging, n-grams, atau sentiment lexicon. Fitur-fitur ini dapat memberikan informasi tambahan yang berguna bagi model dalam mengenali sentimen dari teks ulasan. Selain itu, metadata dari ulasan seperti panjang ulasan, frekuensi kata-kata tertentu, atau pola penggunaan kata bisa menjadi fitur tambahan yang berguna.

Dengan menerapkan saran-saran di atas, diharapkan dapat diperoleh model analisis sentimen yang lebih akurat dan andal dalam mengklasifikasikan ulasan aplikasi TurboVPN di Google Playstore. Hal ini akan membantu dalam mendapatkan wawasan yang lebih baik tentang persepsi pengguna terhadap aplikasi tersebut dan memungkinkan pengambilan keputusan yang lebih baik dalam pengembangan dan pemasaran aplikasi.

# **Daftar Pustaka**

Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, 2(2), 121-167.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321-357.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.

Dietterich, T. G. (2000). Ensemble methods in machine learning. Multiple Classifier Systems, 1857, 1-15.

Forman, G. (2003). An extensive empirical study of feature selection metrics for text classification. Journal of Machine Learning Research, 3, 1289-1305.

Gamon, M. (2004). Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. Proceedings of the ACL 2004 on Interactive poster and demonstration sessions, 9-12.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.

Joachims, T. (1998). Text categorization with Support Vector Machines: Learning with many relevant features. Proceedings of the 10th European Conference on Machine Learning, 137-142.

Jurafsky, D., & Martin, J. H. (2020). Speech and Language Processing (3rd ed.). Pearson.

Kim, S. M., & Hovy, E. (2004). Determining the sentiment of opinions. Proceedings of the 20th international conference on Computational Linguistics, 1367-1373.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press.

McCallum, A., & Nigam, K. (1998). A comparison of event models for Naive Bayes text classification. AAAI-98 workshop on learning for text categorization, 752(1), 41-48.

Lewis, D. D. (1998). Naive (Bayes) at forty: The independence assumption in information retrieval. Proceedings of the 10th European Conference on Machine Learning, 4-15.

Liu, B. (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 5(1), 1-167.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.

Rish, I. (2001). An empirical study of the naive Bayes classifier. IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, 41-46.

Zhang, H. (2004). The Optimality of Naive Bayes. Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference, 562-567.

Pang, B., & Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. Proceedings of ACL, 271-278.