PERBANDINGAN ALGORITMA NAÏVE BAYES DAN SVM TERHADAP ANALISIS SENTIMEN PENGGUNA APLIKASI FATSECRET PADA GOOGLE PLAY STORE

Ahmad Faiq Wicaksono, Rina Candra Noor Santi
Teknik Informatika, Universitas Stikubank Semarang
Jalan Kendeng V Semarang, Indonesia
wicakssono70@gmail.com

ABSTRAK

Pertumbuhan pesat aplikasi kesehatan digital seperti FatSecret memicu lonjakan ulasan pengguna di Google Play Store yang memuat informasi strategis bagi pengembang. Namun, volume data tekstual yang besar dan tidak terstruktur menyulitkan proses ekstraksi informasi secara manual. Penelitian ini bertujuan untuk mengomparasikan kinerja algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM) dalam mengklasifikasikan sentimen pengguna. Metode penelitian menggunakan pendekatan kuantitatif eksperimental yang mencakup pengumpulan 2.000 data ulasan melalui web scraping, pra-pemrosesan teks, dan ekstraksi fitur berbasis Term Frequency-Inverse Document Frequency (TF-IDF). Hasil evaluasi menunjukkan bahwa algoritma Naïve Bayes memiliki performa lebih unggul dengan akurasi mencapai 97.16%, sedangkan SVM dengan kernel linear menghasilkan akurasi sebesar 96.64%. Selisih akurasi ini mengindikasikan bahwa pendekatan probabilistik Naïve Bayes lebih efektif dalam menangani karakteristik data ulasan aplikasi FatSecret dibandingkan SVM.

Kata kunci : *Analisis Sentimen, FatSecret, Naïve Bayes, SVM, TF-IDF.*

1. PENDAHULUAN

Perkembangan teknologi internet di Indonesia yang pesat telah menjadikan aplikasi digital di Google Play Store bagian integral dari kehidupan sehari-hari. Ulasan pengguna di platform ini menjadi aset berharga bagi pengembang untuk mengevaluasi aspek teknis dan pengalaman pengguna.

Namun, volume ulasan yang sangat besar menyulitkan proses identifikasi sentimen secara manual, terlebih sering terjadi ketidaksesuaian antara rating bintang yang diberikan dengan isi ulasan tertulis [1]. Mengingat fungsi aplikasi yang vital bagi kesehatan, umpan balik pengguna sangat krusial, namun analisis manual menjadi tidak efisien dan rentan terhadap kesalahan interpretasi.

Berangkat dari permasalahan tersebut, penelitian ini bertujuan untuk membandingkan performa algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM) secara langsung pada data ulasan aplikasi FatSecret guna menentukan algoritma mana yang paling akurat. Pemilihan kedua algoritma ini didasarkan pada temuan penelitian terdahulu yang beragam, di mana sebagian mengunggulkan SVM pada ulasan aplikasi, sementara yang lain menemukan Naïve Bayes lebih baik pada data media social [2]

Untuk menjawab permasalahan tersebut, penelitian ini merancang rencana penyelesaian masalah melalui penerapan teknik data mining yang sistematis. Tahapan dimulai dengan web scraping untuk mengakuisisi data ulasan, diikuti oleh preprocessing untuk mereduksi noise. Pada tahap ekstraksi fitur, metode Term Frequency-Inverse Document Frequency (TF-IDF) diterapkan karena terbukti mampu meningkatkan akurasi dengan memberikan bobot yang relevan pada kata-kata penting dalam dokumen teks, sebagaimana dibuktikan oleh penelitian [3].

Strategi inti penyelesaian masalah difokuskan pada komparasi dua algoritma Naïve Bayes dan SVM. Pemilihan algoritma Naïve Bayes didasarkan pada studi [2] yang menyoroti efisiensi dan keunggulan akurasinya pada data ulasan media sosial dengan fitur independen yang kuat. Sementara itu, SVM dipilih sebagai pembanding merujuk pada penelitian [1] yang membuktikan ketangguhan SVM dalam memisahkan kelas sentimen pada data berdimensi tinggi melalui pemanfaatan fungsi kernel. Integrasi metode-metode yang telah teruji dalam penelitian terdahulu ini diharapkan dapat menghasilkan model klasifikasi yang paling optimal untuk kasus ulasan aplikasi FatSecret.

2. TINJAUAN PUSTAKA

Bagian ini menguraikan konsep-konsep teoretis dari riset yang telah dihasilkan oleh para peneliti sebelumnya. Landasan teoretis ini akan digunakan sebagai panduan dan rujukan dalam melaksanakan studi berjudul "Perbandingan Algoritma Naïve Bayes Dan SVM Pada Analisis Sentimen Pengguna Aplikasi FatSecret pada Google Play Store". Berikut beberapa konsep relevan dari studi-studi sebelumnya.

2.1. Penelitian Terdahulu

Dalam penelitian yang telah dilakukan [4] mengenai analisis sentimen terhadap aplikasi SatuSehat pada Twitter, dijelaskan bahwa perbandingan antara algoritma Naive Bayes dan Support Vector Machine (SVM) menunjukkan SVM menghasilkan akurasi yang lebih tinggi, yaitu 87.95%, dibandingkan Naive Bayes yang sebesar 81.65%, dengan mayoritas sentimen terdeteksi negatif.

Dalam penelitian yang telah dilakukan [5] tentang analisis sentimen ulasan aplikasi investasi Pluang di Google Play Store, dijelaskan bahwa perbandingan algoritma Naive Bayes dan Support

Vector Machine (SVM) menunjukkan model SVM bekerja lebih baik dengan akurasi 99.50%, dibandingkan Naive Bayes dengan akurasi 99.25%.

Dalam penelitian yang telah dilakukan [6] kuota internet Kemdikbudristek selama pandemi Covid-19 di Twitter, dijelaskan bahwa perbandingan antara Naïve Bayes dan Support Vector Machine (SVM) menunjukkan tingkat akurasi SVM lebih tinggi (80%) dibandingkan Naïve Bayes (64%).

Dalam penelitian yang telah dilakukan [7] mengenai analisis sentimen terhadap program Kampus Merdeka di Twitter, dijelaskan bahwa perbandingan algoritma Naïve Bayes dan SVM menunjukkan bahwa SVM dengan kernel linear menghasilkan akurasi lebih tinggi (93%) pada data pengujian dibandingkan Naïve Bayes (86%).

Dalam penelitian yang telah dilakukan[8]tentang analisis sentimen ulasan aplikasi TikTok di Google Play Store, dijelaskan bahwa perbandingan algoritma Naïve Bayes dan Support Vector Machine (SVM) menunjukkan nilai akurasi SVM (84%) lebih tinggi dibandingkan metode Naïve Bayes (79%).

Dalam penelitian yang telah dilakukan [9] mengenai perbandingan kinerja Naïve Bayes dan SVM pada analisis sentimen Twitter terkait Ibukota Nusantara (IKN), dijelaskan bahwa metode SVM memberikan tingkat akurasi analisis sentimen sebesar 94%, yang lebih tinggi dibandingkan metode Naïve Bayes yang mencapai 91%.

3. METODE PENELITIAN

3.1. Pengumpulan Data

yaitu proses pengambilan (scraping) data ulasan pengguna aplikasi FatSecret dari platform Google Play Store menggunakan library google-play-scraper.

3.2. Pelabelan Data

secara otomatis berdasarkan skor (rating) yang diberikan pengguna. Ulasan dengan skor 4 dan 5 diberi label 'Positif', sementara ulasan dengan skor kurang dari 3 (1 dan 2) diberi label 'Negatif'. Ulasan yang tidak termasuk dalam kategori tersebut (skor 3) akan dihapus (handling missing value).

3.3. Preprocessing.

Tahapan ini meliputi Case Folding (mengubah teks menjadi huruf kecil), Cleaning (membersihkan teks dari angka, tanda baca, dan karakter spesial), Tokenization (memecah kalimat menjadi token kata), Stopword Removal (menghapus kata-kata umum dalam bahasa Indonesia), dan Stemming (mengubah kata berimbuhan menjadi kata dasar menggunakan pustaka Sastrawi).

3.4. Pembobotan Kata

Pembobotan Kata menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency) untuk mengubah data teks menjadi vektor numerik . Data vektor ini kemudian dibagi menjadi dua bagian: 80% sebagai data latih (training) dan 20% sebagai data uji (testing).

Penghitungan nilai IDF bertujuan untuk mengurangi dominasi kata-kata umum yang tidak informatif, dengan rumus sebagai berikut:

$$IDF(t) = \log(\frac{N}{df(t)})$$

Selanjutnya, bobot akhir untuk setiap kata dihitung dengan mengalikan nilai TF dengan nilai IDF, sebagaimana persamaan berikut:

$$W_{t,d} = TF_{t,d} \times IDF_t$$

3.5. Klasifikasi Naïve Bayes

Naïve Bayes Classifier (NBC) merupakan salah satu algoritma klasifikasi yang populer dalam machine learning, bekerja berdasarkan prinsip probabilitas dan statistik yang didasarkan pada Teorema Bayes. Metode ini mengklasifikasikan data dengan menghitung probabilitas suatu objek termasuk dalam kelas tertentu, dengan asumsi utama bahwa setiap fitur (atribut) bersifat independen satu sama lain dalam mempengaruhi hasil klasifikasi[10].

Persamaan dasar dari Teorema Bayes yang digunakan dalam Algoritma *Naïve Bayes Classification* dijelaskan sebagai berikut:

$$P(c|X) = \frac{P(X|c).P(c)}{P(X)}$$

3.6. Klasifikasi Support Vector Machine (SVM)

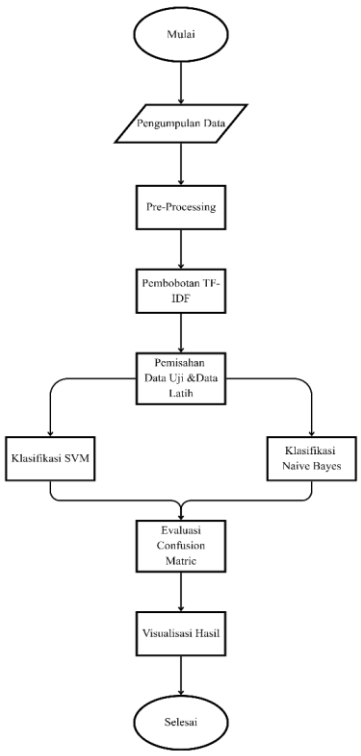
Support Vector Machine (SVM) adalah algoritma machine learning yang termasuk dalam kategori supervised learning dan banyak digunakan untuk tugas klasifikasi serta regresi. SVM bekerja dengan prinsip mencari hyperplane atau garis pemisah (decision boundary) terbaik yang dapat memisahkan data ke dalam kelas-kelas yang berbeda dengan jarak (margin) semaksimal mungkin antara titik data terdekat dari masing-masing kelas (yang disebut support vectors) ke hyperplane tersebut.[11]

Secara matematis, hyperplane pemisah dalam SVM dapat direpresentasikan oleh persamaan:

$$(w.x_i) + b = 0$$

3.7. Evaluasi Model

kedua model yang telah dilatih akan diuji kinerjanya menggunakan data uji. Hasil evaluasi akan dibandingkan menggunakan Confusion Matrix untuk mendapatkan nilai akurasi, presisi, recall, dan f1-score guna menentukan algoritma mana yang memiliki performa terbaik. Untuk lebih lengkapnya, flowchart penelitian dapat dilihat pada Gambar 1.



Gambar 1. Alur metode penelitian

4. HASIL DAN PEMBAHASAN

4.1. Hasil Pengumpulan Data

Tahap awal implementasi adalah pengumpulan data ulasan aplikasi FatSecret. Proses ini dilakukan menggunakan teknik web scraping dengan bantuan library google-play-scraper pada lingkungan Google Colab. Parameter yang digunakan meliputi app_id ('com.FatSecret.android'), bahasa (lang='id'), dan negara (country='id').

Tabel 1. Hasil scraping data

Nama	Waktu	Rating	Komentar
Shabrina Siregar	2025-10-28 10:35:50	4	AMAZING!! bisa membantu, sudah hampir 5 tahun memakai, hanya perlu upgrade banyak makanan/camilan kekenian dari segala merk untuk dapan lebih membatu lagi, semoga semakin langsing kita-kita yaakkkk
Rosita Sopwi	2025-11-18 20:46:57	5	Kalau konsisten, niat, dan benar-benar ngitung kalorinya Insyaallah bantu banget buat diet
y0uhax z3ro	2025-7-12 12:37:02	1	kenapa si astaga padahal jaringan bagus tapi buat buka aplikasi, buat input makanan gabisa, lucu banget
Ricegun Mehveen	2023-5-30 13:35:02	1	Istirahat kenapa secara default dihitung ngurangin kalori? ganggu banget sumpah, tambahin fitur buat matiin ini dong. cuma mau hitung pengurangan kalori dari olahraga jadi susah. sayang banget aplikasi bagus tapi bagian ini nya ga dibener"in.

Dari tabel diatas dapat kita lihat berhasil dikumpulkan sebanyak 2.000 data ulasan mentah. Data yang diambil mencakup informasi nama pengguna, waktu ulasan, rating (skor bintang), dan isi komentar (content).

4.2. Hasil Pelabelan

Sesuai dengan metode penelitian, ulasan dengan skor 4 dan 5 dikategorikan sebagai sentimen Positif, sedangkan ulasan dengan skor 1 dan 2 dikategorikan sebagai sentimen Negatif. Ulasan dengan skor 3 (netral) dihapus dari dataset untuk menghindari ambiguitas.

Tabel 2. Hasil pelabelan dataset

Komen	Label Sentimen
sangat membantu untuk diet, buat ngerem makanan yang seharusnya. perhitungan kalorinya mantap. manjur. josss.	Positif
AMAZING!! bisa membantu, sudah hampir 5 tahun memakai, hanya perlu upgrade banyak makanan/camilan kekenian dari segala merk untuk dapan lebih membatu lagi, semoga semakin langsing kita-kita yaakkkk	Positif
Gak tau kenapa udah 3 hari, mau nyatet kalori harian pas maghrib ke atas langsung lemot, sebelemnya gak pernah kaya gini. Ini udah uninstal terus reinstal tetep sama, malah gak bisa login saking loading mulu.	Negatif
Sebelumnya pernah pake, ngebantu banget turun hampir 10kilo, tapi secarang appnya entah kenapa ngak bisa kebuka stak di awal ngak mau masuk ke app nya.	Negatif

Tabel diatas Adalah Hasil dari proses pelabelan, ini adalah dataset yang siap untuk tahap preprocessing selanjutnya.

4.3. Hasil Preprocessing

Tahapan preprocessing dilakukan untuk membersihkan data teks agar siap diolah oleh algoritma. Tahapan ini meliputi Case Folding, Cleaning, Tokenization, Stopword Removal, dan Stemming . Berikut adalah rincian hasil dari setiap tahapan :

4.4. Case Folding Dan Cleaning

Pada tahap ini, seluruh huruf diubah menjadi huruf kecil (lowercase) dan karakter-karakter pengganggu seperti tanda baca, angka, emoji, serta username (@) dihapus.

Tabel 3. Hasil case folding dan cleaning

Komen	Case Folding
AMAZING!! bisa membantu, sudah hampir 5 tahun memakai, hanya perlu upgrade banyak makanan/camilan kekenian dari segala merk untuk dapan lebih membatu lagi, semoga semakin langsing kita-kita yaakkkk	amazing bisa membantu sudah hampir tahun memakai hanya perlu upgrade banyak makanancamilan kekenian dari segala merk untuk dapan lebih membatu lagi semoga semakin langsing kitakita yaakkkk
KERENN BGTT PLISS! tp coba perbaiki lagi pake kamera gitu biar bisa tau kalori yang kita makan tuh berapa, jdi di deteksi otomatis	kerenn bgtt pliss tp coba perbaiki lagi pake kamera gitu biar bisa tau kalori yang kita makan tuh berapa jdi di deteksi otomatis
Aplikasi yang detail, dan tepat untuk orang yg ingin mengatur pola makan	aplikasi yang detail dan tepat untuk orang yg ingin mengatur pola makan

Dari tabel diatas, proses case folding dan cleaning berhasil menstandarisasi data dengan mengonversi seluruh teks menjadi huruf kecil (lowercase) serta mengeliminasi noise berupa angka dan tanda baca.

4.5. Tokenization

Teks yang telah bersih kemudian dipecah menjadi potongan kata (token) yang berdiri sendiri.

Tabel 4. Hasil tokenization

Komen	Tokenization
amazing bisa membantu sudah hampir tahun memakai hanya perlu upgrade banyak makanancamilan kekenian dari segala merk untuk dapan lebih membatu lagi semoga semakin langsing kitakita yaakkkk	[amazing, membantu, memakai, upgrade, makanancamilan, kekonian, merk, dapan, membatu, semoga, langsing, kitakita, yaakkkk]
kerenn bgtt pliss tp coba perbaiki lagi pake kamera gitu biar bisa tau kalori yang kita makan tuh berapa jdi di deteksi otomatis	[kerenn, bgtt, pliss, tp, coba, perbaiki, pake, kamera, gitu, biar, tau, kalori, makan, tuh, jdi, deteksi ,otomatis]
aplikasi yang detail dan tepat untuk orang yg ingin mengatur pola makan	[aplikasi, detail ,orang, yg, mengatur, pola, makan]

Berdasarkan tabel diatas, proses tokenisasi menguraikan untaian teks yang telah dibersihkan (cleaned text) menjadi unit-unit kata diskrit atau token yang berdiri sendiri. Tahapan ini mentransformasi struktur data dari kalimat utuh menjadi senarai (list) kata terpisah, yang bertujuan untuk memisahkan setiap entitas leksikal agar dapat dianalisis secara individual pada tahapan pemrosesan selanjutnya.

4.6. Stopword Removal

Kata-kata umum yang tidak memiliki makna sentimen signifikan (seperti "yang", "dan", "tapi", "di") dihapus menggunakan kamus stopword bahasa Indonesia dari pustaka NLTK.

Tabel 5. Hasil stopword removal

Komen	Stopword Removal
amazing bisa membantu sudah hampir tahun memakai hanya perlu upgrade banyak makanancamilan kekenian dari segala merk untuk dapan lebih membatu lagi semoga semakin langsing kitakita yaakkkk	amazing membantu memakai upgrade makanancamilan kekenian merk dapan membatu semoga langsing kitakita yaakkkk
kerenn bgtt pliss tp coba perbaiki lagi pake kamera gitu biar bisa tau kalori yang kita makan tuh berapa jdi di deteksi otomatis	kerenn bgtt pliss tp coba perbaiki pake kamera gitu biar tau kalori makan tuh jdi deteksi otomatis
aplikasi yang detail dan tepat untuk orang yg ingin mengatur pola makan	aplikasi detail orang yg mengatur pola maka

Berdasarkan diatas dapat kita lihat, tahapan stopword removal berfungsi mereduksi dimensi data dengan mengeliminasi kata-kata umum yang tidak memiliki signifikansi sentimen (stopwords) menggunakan referensi pustaka NLTK. Sehingga menghasilkan himpunan kata kunci esensial yang lebih padat (seperti "amazing", "membantu", dan "aplikasi") guna meningkatkan efisiensi dan fokus analisis pada tahap klasifikasi selanjutnya.

4.7. Stemming

Kata-kata berimbuhan diubah menjadi kata dasar menggunakan pustaka Sastrawi. Misalnya kata "membantu" menjadi "bantu" dan "menggunakan" menjadi "guna.

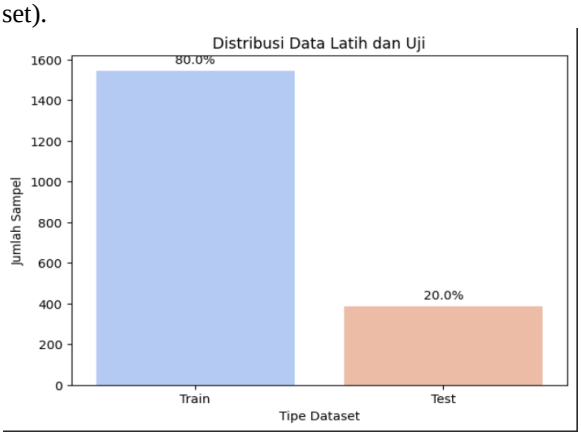
Tabel 6. Hasil stemming

Komen	Stemming
amazing membantu memakai upgrade makanancamilan kekenian merk dapan membatu semoga langsing kitakita yaakkkk	amazing bantu pakai upgrade makanancamilan nian merk dap batu moga langsing kitakita yaakkkk
kerenn bgtt pliss tp coba perbaiki pake kamera gitu biar tau kalori makan tuh jdi deteksi otomatis	kerenn bgtt pliss tp coba baik pake kamera gitu biar tau kalori makan tuh jdi deteksi otomatis
aplikasi yang detail dan tepat untuk orang yg ingin mengatur pola makan	aplikasi detail orang yg atur pola makan

Berdasarkan tabel diatas, tahapan stemming diimplementasikan menggunakan pustaka Sastrawi untuk mentransformasi kata-kata berimbuhan menjadi bentuk dasar (root word) guna mereduksi variasi morfologis data. Proses ini terbukti efektif menyederhanakan fitur leksikal, sebagaimana terlihat pada konversi kata "membantu" menjadi "bantu", "memakai" menjadi "pakai", serta "perbaiki" menjadi "baik". Meskipun demikian, algoritma ini juga menunjukkan keterbatasan pada kata tidak baku atau typo, seperti "dapan" yang terpotong menjadi "dap", namun secara umum langkah ini berhasil menyeragamkan kata-kata yang memiliki makna dasar sama untuk keperluan analisis selanjutnya.

4.8. Hasil Pembagian Data

Sesuai dengan skenario pengujian yang dirancang pada Bab 3, dataset dibagi menjadi dua bagian: data latih (training set) dan data uji (testing set).



Gambar 2. Hasil Pembagian Data Dalam Bentuk Diagram

Dari gambar diatas dapat dilihat total data bersih yang tersedia, komposisi pembagian data adalah sebagai berikut:

a. Data Latih (80%): Digunakan untuk melatih model algoritma agar dapat mengenali pola sentimen. Jumlah data latih adalah sebanyak 1544 data.
b. Data Uji (20%): Digunakan untuk mengukur kinerja model yang telah dilatih. Jumlah data uji adalah sebanyak 387 data.

4.9. Hasil Pembobotan TF-IDF

Data teks pada data latih dan data uji tidak dapat langsung diproses oleh algoritma machine learning. Oleh karena itu, dilakukan proses ekstraksi fitur untuk mengubah data teks menjadi vektor numerik. Metode yang digunakan dalam penelitian ini adalah TF-IDF (Term Frequency-Inverse Document Frequency)

Tabel 7. Sepuluh Kata dengan Bobot TF-IDF Tertinggi

No	Kata	*TF-IDF* Score
1	membantu	98.608
2	sangat	97.426
3	kalori	85.991
4	untuk	67.057
5	aplikasi	64.097
6	dan	63.945
7	yang	59.389
8	diet	58.267
9	bisa	55.416
10	makanan	55.398

Dari tabel diatas menampilkan sepuluh kata dengan bobot TF-IDF tertinggi dalam kumpulan ulasan pengguna aplikasi FatSecret.

4.10. Evaluasi Hasil Model Naïve Bayes

Pengujian pertama dilakukan menggunakan algoritma Multinomial Naïve Bayes. Algoritma ini dipilih karena cocok untuk klasifikasi teks dengan fitur diskrit (seperti frekuensi kata). Model dilatih menggunakan data latih yang telah dibobotkan, kemudian model tersebut digunakan untuk memprediksi label sentimen pada data uji.

```
MultinomialNB Accuracy: 0.9715762273901809
MultinomialNB Precision: 0.5
MultinomialNB Recall: 0.36363636363636365
MultinomialNB f1_score: 0.42105263157894735
confusion_matrix:
 [[  4   7]
 [  4 372]]
==================================================

              precision    recall  f1-score   support

     Negatif       0.50      0.36      0.42        11
     Positif       0.98      0.99      0.99       376

    accuracy                           0.97       387
   macro avg       0.74      0.68      0.70       387
weighted avg       0.97      0.97      0.97       387
```

Gambar 3. Hasil evaluasi metode naïve bayes

Pada gambar 3 diatas penerapan algoritma Naïve Bayes dalam penelitian saya ini dengan menggunakan data testing sebesar 20% dapay menghasilkan akurasi sebesar 97,16%.

4.11. Evaluasi Hasil Model SVM

Pengujian kedua dilakukan menggunakan algoritma Support Vector Machine (SVM). Sesuai dengan kode program, kernel yang digunakan adalah kernel linear (kernel='linear'), yang dikenal efektif untuk memisahkan data teks berdimensi tinggi secara linear.

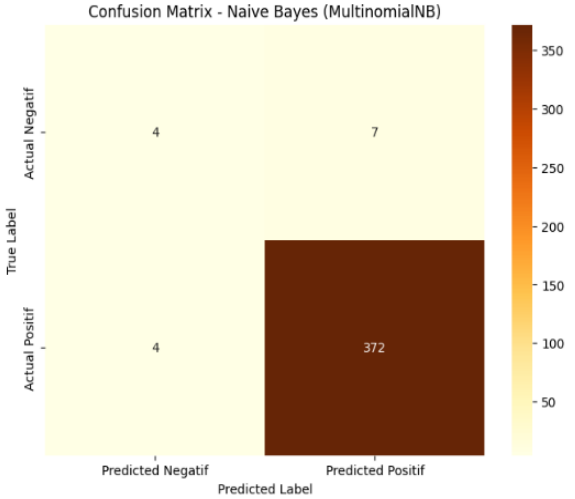
```
SVM Accuracy: 0.9664082687338501
SVM Precision: 0.4
SVM Recall: 0.36363636363636365
SVM f1_score: 0.38095238095238093
confusion_matrix:
 [[  4   7]
 [  6 370]]
==================================================

              precision    recall  f1-score   support

     Negatif       0.40      0.36      0.38        11
     Positif       0.98      0.98      0.98       376

    accuracy                           0.97       387
   macro avg       0.69      0.67      0.68       387
weighted avg       0.96      0.97      0.97       387
```

Gambar 4. Hasil evaluasi metode SVM

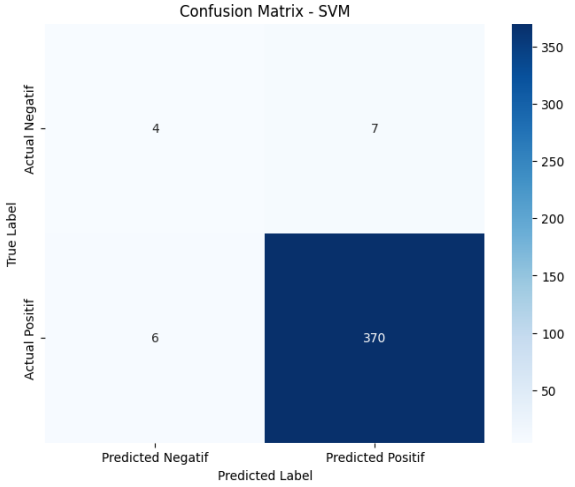
Pada gambar diatas penerapan algoritma Support Vector Machine (SVM) dalam penelitian saya ini dengan menggunakan data testing sebesar 20% dapay menghasilkan akurasi sebesar 96,64%.

4.12. Hasil Analisis Perbandingan

Berdasarkan hasil pengujian yang telah dilakukan, berikut adalah perbandingan kinerja antara algoritma Naïve Bayes dan SVM dalam melakukan analisis sentimen terhadap ulasan aplikasi FatSecret.



Gambar 5. Confusion matrix naïve bayes



Gambar 6. Confusion matrix SVM

Berdasarkan *confusion matrix* pada Gambar 5 dan 6, algoritma *Naïve Bayes* menunjukkan performa superior dengan akurasi 97,16% (mengidentifikasi 4 ulasan negatif dan 372 positif), mengungguli *Support Vector Machine* (SVM) yang mencapai 96,64% (4 negatif, 370 positif). Temuan empiris ini menegaskan bahwa dalam studi kasus aplikasi FatSecret, *Naïve Bayes* terbukti lebih optimal dan stabil dibandingkan SVM, meskipun hasil ini berbeda dengan tren pada beberapa literatur terdahulu.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengimplementasikan text mining dengan pembobotan TF-IDF untuk klasifikasi sentimen ulasan aplikasi FatSecret melalui serangkaian tahapan mulai dari scraping hingga preprocessing. Berdasarkan hasil pengujian, algoritma Naïve Bayes Classifier terbukti lebih superior dengan akurasi 97,16%, mengungguli Support Vector Machine (SVM) yang memperoleh 96,64%. Keunggulan selisih 0,52% tersebut mengindikasikan bahwa pendekatan probabilistik Naïve Bayes bekerja lebih efektif dalam menangani karakteristik independensi fitur pada data ulasan ini dibandingkan pendekatan hyperplane yang diterapkan SVM.

Terkait pengembangan di masa depan, peneliti selanjutnya disarankan untuk mengeksplorasi metode ekstraksi fitur semantik (seperti Word2Vec), menerapkan teknik penanganan imbalanced dataset, membandingkan kinerja dengan algoritma Deep Learning, serta menyempurnakan kamus normalisasi bahasa gaul. Di sisi lain, pengembang aplikasi dapat memanfaatkan hasil klasifikasi ini sebagai instrumen monitoring kepuasan pengguna secara real-time, di mana ulasan bersentimen negatif harus menjadi prioritas evaluasi utama demi perbaikan fitur dan kualitas teknis aplikasi.

DAFTAR PUSTAKA

[1] F. Nufairi, N. Pratiwi, and F. Herlando, “Analisis Sentimen Pada Ulasan Aplikasi Threads Di Google Play Store Menggunakan Algoritma Support Vector Machine,” *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 1, pp. 339–348, 2024, doi: 10.29100/jipi.v9i1.4929.

[2] R. P. Setiawan, B. Irawan, and W. P. Prihartono, “Analisis Sentimen Ulasan Growtopia Di Google Play Store Menggunakan Naïve Bayes Classifier Untuk Identifikasi Kebutuhan Pengguna,” *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 2, 2025, doi: 10.23960/jitet.v13i2.6415.

[3] R. R. Rismansyah, A. Sudiarjo, and T. Mufizar, “Analisis Sentimen Ulasan Shopee Pada Google Play Store Menggunakan Algoritma Naive Bayes,” *J. Elektro Inform. Swadharma*, vol. 5, no. 1, 2025, [Online]. Available: https://doi.org/10.57152/jeis.v5i1.661

[4] M. I. Fikri, T. S. Sabrila, and Y. Azhar, “Comparison of Naïve Bayes and Support Vector Machine Methods in Twitter Sentiment Analysis,” *Smatika J.*, vol. 10, no. 02, pp. 71–76, 2020.

[5] F. Matheos Sarimole and Kudrat, “Analisis Sentimen Terhadap Aplikasi Satu Sehat Pada Twitter Menggunakan Algoritma Naive Bayes Dan Support Vector Machine,” *J. Sains dan Teknol.*, vol. 5, no. 3, pp. 1–8, 2024, [Online]. Available: https://doi.org/10.55338/saintek.v5i1.2702

[6] B. A. Maulana, M. J. Fahmi, A. M. Imran, and N. Hidayati, “Sentiment Analysis of Pluang

Applications With Naive Bayes and Support Vector Machine (SVM)," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 2, pp. 375–384, 2024.

[7] U. Khaira, R. Aryani, and R. W. Hardian, "Komparasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM) Pada Analisis Sentimen Kebijakan Kemdikbudristek Mengenai Kuota Internet Selama Covid-19," *J. Process.*, vol. 18, no. 2, pp. 183–191, 2023, doi: 10.33998/processor.2023.18.2.897.

[8] I. P. Rahayu, A. Fauzi, and J. Indra, "Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Dan Support Vector Machine," *J. Sist. Komput. dan Inform.*, vol. 4, no. 2, p. 296, 2022, doi: 10.30865/json.v4i2.5381.

[9] Friska Aditia Indriyani, Ahmad Fauzi, and Sutan Faisal, "Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine," *TEKNOSAINS J. Sains, Teknol. dan Inform.*, vol. 10, no. 2, pp. 176–184, 2023, doi: 10.37373/tekno.v10i2.419.

[10] A. Supian, B. Tri Revaldo, N. Marhadi, L. Efrizoni, and R. Rahmaddeni, "Perbandingan Kinerja Naïve Bayes Dan Svm Pada Analisis Sentimen Twitter Ibukota Nusantara," *J. Ilm. Inform.*, vol. 12, no. 01, pp. 15–21, 2024, doi: 10.33884/jif.v12i01.8721.

[11] L. O. Alyandi, "PERBANDINGAN KINERJA NAÏVE BAYES DAN SVM DALAM ANALISIS SENTIMEN ULASAN FREE FIRE DI PLAY STORE," vol. 9, no. 3, pp. 5127–5135, 2025.

[12] M. B. Anggara, F. T. Informasi, and U. B. Bandung, "PERBANDINGAN NAÏVE BAYES DAN SVM DALAM ANALISIS SENTIMEN ULASAN APLIKASI RSUD AL IHSAN MOBILE," vol. 20, pp. 32–42, 2025.