PERBANDINGAN KINERJA METODE VADER DAN NAIVE BAYES DALAM KLASIFIKASI SENTIMEN ULASAN APLIKASI DANA

SKRIPSI

Oleh :
PREVIANDI JELANG ASHARI
NIM. 19650095

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026

PERBANDINGAN KINERJA METODE VADER DAN NAIVE BAYES DALAM KLASIFIKASI SENTIMEN ULASAN APLIKASI DANA

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh :
PREVIANDI JELANG ASHARI
NIM. 19650095

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2026

HALAMAN PERSETUJUAN

PERBANDINGAN KINERJA METODE VADER DAN NAÏVE BAYES DALAM KLASIFIKASI SENTIMEN ULASAN APLIKASI DANA

SKRIPSI

Oleh :
PREVIANDI JELANG ASHARI
NIM. 19650095

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 17 Juni 2026

Pembimbing I,	Pembimbing II,
Dr. Zainal Abidin, M.Kom	Nur Fitriyah Ayu Tunjung Sari, M.Cs
NIP. 19760613 200501 1 004	NIP. 19911226 202012 2 001

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Supriyono, M. Kom
NIP. 19841010 201903 1 012

HALAMAN PENGESAHAN

PERBANDINGAN KINERJA METODE VADER DAN NAIVE BAYES DALAM KLASIFIKASI SENTIMEN ULASAN APLIKASI DANA

SKRIPSI

Oleh :
PREVIANDI JELANG ASHARI
NIM. 19650095

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 26 Juni 2026

Susunan Dewan Penguji

Ketua Penguji	: Fathurrochman, M.Kom NIP. 19700731 200501 1 002	( )
Anggota Penguji I	: Shoffin Nahwa Utama, MT NIP. 19860703 202012 1 003	( )
Anggota Penguji II	: Dr. Zainal Abidin, M.Kom NIP. 19760613 200501 1 004	( )
Anggota Penguji III	: Nur Fitriyah Ayu Tunjung Sari,M.Cs NIP. 19911226 202012 2 001	( )

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Supriyono,M.Kom
NIP. 19841010 201903 1 012

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama	: Previandi Jelang Ashari
NIM	: 19650095
Fakultas / Jurusan	: Sains dan Teknologi / Teknik Informatika
Judul Skripsi	: PERBANDINGAN KINERJA METODE VADER DAN NAIVE BAYES DALAM KLASIFIKASI SENTIMEN ULASAN APLIKASI DANA

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 22 Juni 2026
Yang membuat pernyataan,

Previandi Jelang Ashari
NIM.19650095

MOTTO

... Setiap masalah memiliki solusi, sebagaimana setiap kesalahan dalam program dapat diperbaiki dengan proses yang tepat...

HALAMAN PERSEMBAHAN

Bismillahirrahmanirrahim

Dengan mengucapkan rasa syukur ke hadirat Allah SWT atas segala rahmat, karunia, kesehatan, dan kemudahan yang telah diberikan sehingga skripsi ini dapat terselesaikan dengan baik.

Karya sederhana ini kupersembahkan kepada:

Kedua orang tua tercinta, yang senantiasa memberikan doa, kasih sayang, dukungan, dan pengorbanan tanpa henti dalam setiap langkah perjalanan hidup dan pendidikan penulis.

Bapak dan Ibu Dosen, khususnya dosen pembimbing yang telah memberikan ilmu, arahan, dan bimbingan selama proses penyusunan skripsi ini.

Keluarga, sahabat, dan teman-teman seperjuangan, yang selalu memberikan semangat, motivasi, dan dukungan hingga skripsi ini dapat diselesaikan.

Almamater tercinta, Universitas Islam Negeri Maulana Malik Ibrahim Malang, sebagai tempat penulis menimba ilmu dan mengembangkan diri.

Karya ini kupersembahkan untuk diriku sendiri, yang telah berusaha, berdoa, dan berjuang hingga mampu menyelesaikan pendidikan dan penelitian ini dengan sebaik-baiknya.

KATA PENGANTAR

Assalamu'alaikum warahmatullahi wabarakatuh

Puji syukur kehadirat Allah SWT atas segala rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **"Perbandingan Kinerja Metode VADER dan Naïve Bayes dalam Klasifikasi Sentimen Aplikasi DANA"** sebagai salah satu syarat memperoleh gelar Sarjana Komputer pada Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Dalam proses penyusunan skripsi ini, penulis memperoleh banyak bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si, selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. H. Agus Mulyono, M.Si, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Supriyono, M.Kom, selaku Ketua Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Zainal Abidin, M.Kom, selaku dosen pembimbing 1 yang dengan penuh kesabaran senantiasa membimbingan, memberi masukan, dan memberikan dukungan yang sangat berarti bagi penulis selama proses penelitian.

5. Nur Fitriyah Ayu Tunjung Sari,M.Cs selaku dosen pembimbing 2, atas arahan dan masukan yang sangat membantu untuk menyempurnakan karya ini
6. Fatchurrochman, M.Kom, selaku dosen penguji I dan Shoffin Nahwa Utama, MT selaku dosen penguji II yang telah memberikan kritik dan saran yang membangun demi kesempurnaan skripsi ini.
7. Seluruh staf dan dosen Program Studi Teknik Informatika yang telah memberikan ilmu, dukungan, dan berbagai fasilitas kepada penulis selama menempuh masa studi.
8. Seluruh pihak yang turut berkontribusi, baik secara langsung maupun tidak langsung.

Penulis berharap semoga skripsi ini dapat memberikan kontribusi dan manfaat di masa yang akan datang. Penulis menyadari bahwa karya ini masih jauh dari kata sempurna. Oleh karena itu, penulis dengan senang hati menerima masukan dan saran yang membangun untuk pengembangan lebih lanjut.

Wassalamu'alaikum Warahmatullahi Wabaraktuh

Malang, 22 Juni 2026

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR SIMBOL

Lambang	Kuantitas	Satuan
D	Dokumen (*Document*)	-
d	Dokumen ke-d	-
t	*Term* (kata)	-
tf	Frekuensi kemunculan term (*Term Frequency*)	-
idf	*Inverse Document Frequency*	-
N	Jumlah seluruh dokumen	Dokumen
df	Jumlah dokumen yang mengandung suatu term	Dokumen
P(C)	Probabilitas prior suatu kelas	-
P(X|C)	Probabilitas data terhadap kelas	-
P(C|X)	Probabilitas posterior	-
X	Data atau dokumen yang diklasifikasikan	-
TP	*True Positive*	Data
TN	*True Negative*	Data
FP	*False Positive*	Data
FN	*False Negative*	Data
Acc	*Accuracy*	%
Prec	*Precision*	%
Rec	*Recall*	%
F1	*F1-Score*	%

Singkatan

AI	*Artificial Intelligence*
CSV	*Comma-Separated Values*
DANA	*Dompet Digital Indonesia*
ML	*Machine Learning*
NB	*Naïve Bayes*
NLP	*Natural Language Processing*
TF	*Term Frequency*
IDF	*Inverse Document Frequency*
TF-IDF	*Term Frequency–Inverse Document Frequency*
VADER	*Valence Aware Dictionary and Sentiment Reasoner*
API	*Application Programming Interface*
CSV	*Comma-Separated Values*
UI	*User Interface*
URL	*Uniform Resource Locator*
TP	*True Positive*
TN	*True Negative*
FP	*False Positive*
FN	*False Negative*

ABSTRAK

Ashari, Previandi Jelang **Perbandingan Kinerja Metode VADER Dan Naïve Bayes Dalam Klasifikasi Sentimen Aplikasi Dana.** Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Zainal Abidin, M.Kom (II) Nur Fitriyah Ayu Tunjung Sari, M.Cs

.

Kata kunci: Analisis Sentimen, DANA, VADER, Naïve Bayes, TF-IDF..

Aplikasi DANA merupakan salah satu layanan dompet digital yang banyak digunakan di Indonesia. Banyaknya ulasan pengguna pada *Google Play Store* mengandung informasi berupa opini, kritik, dan saran yang dapat dimanfaatkan untuk mengevaluasi kualitas layanan aplikasi. Namun, volume ulasan yang besar membuat analisis secara manual menjadi tidak efisien sehingga diperlukan analisis sentimen otomatis. Penelitian ini membandingkan performa metode VADER (*Valence Aware Dictionary and Sentiment Reasoner*) berbasis leksikon dan *Naive Bayes* berbasis *machine learning* dalam klasifikasi sentimen ulasan pengguna aplikasi DANA. Dataset berjumlah 50.000 ulasan diperoleh dari *Kaggle* dan telah memiliki label sentimen. Data diproses melalui tahapan *preprocessing* meliputi *case folding*, *cleaning*, *tokenizing*, *stopword removal*, dan *stemming*. Ekstraksi fitur dilakukan menggunakan TF-IDF dan evaluasi menggunakan *confusion matrix* dengan *metrik accuracy, precision, recall*, dan *F1-score*. Hasil penelitian menunjukkan *Naive Bayes* memperoleh akurasi 77,14%, lebih tinggi dibandingkan VADER yang hanya mencapai 67,76%, dengan selisih 9,38%. Metode *Naive Bayes* terbukti lebih efektif untuk klasifikasi sentimen ulasan berbahasa Indonesia dan bahasa inggris yang banyak mengandung bahasa *informal*.

ABSTRACT

Ashari, Previandi Jelang. 2026. **Comparison of the Performance of VADER and Naïve Bayes Methods in Sentiment Classification of the DANA** Application. Undergraduate Thesis. Department of Informatics Engineering, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University of Malang. Supervisors: (I) Zainal Abidin, M.Kom. (II) Nur Fitriyah Ayu Tunjung Sari, M.Cs.

Keywords: Sentiment Analysis, DANA, VADER, Naïve Bayes, TF-IDF.

DANA is one of the most widely used digital wallet applications in Indonesia. The large volume of user reviews on Google Play Store contains valuable information in the form of opinions, criticisms, and suggestions that can be utilized to evaluate service quality. However, the sheer number of reviews makes manual analysis inefficient, necessitating automated sentiment analysis. This study compares the performance of VADER (Valence Aware Dictionary and Sentiment Reasoner), a lexicon-based method, and Naive Bayes, a machine learning approach, for classifying sentiment from DANA application user reviews. A dataset of 50,000 reviews obtained from Kaggle with pre-existing sentiment labels was used. The data were processed through preprocessing stages including case folding, cleaning, tokenizing, stopword removal, and stemming. Feature extraction was performed using TF-IDF and evaluation employed a confusion matrix with accuracy, precision, recall, and F1-score metrics. Results show that Naive Bayes achieved an accuracy of 77.14%, outperforming VADER at 67.76%, a difference of 9.38%. Naive Bayes is proven more effective for sentiment classification of Indonesian-language reviews that contain extensive informal language.

البحث مستخلص

في تصنيف المشاعر لتطبيق Naïve Bayesو VADER أشهري، بريفياندي جيلانج. 2026. مقارنة أداء طريقتي DANA. رسالة بكالوريوس، قسم هندسة المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية M.Cs. ،نور فطرية أيو تونجونج ساري (2) .M.Kom ،الحكومية مالانج. المشرفان: (1) زين العابدين

الكلمات المفتاحية: تحليل المشاعر، DANA، VADER، Naïve Bayes، TF-IDF.

أحد أكثر تطبيقات المحافظ الرقمية استخدامًا في إندونيسيا. ويحتوي العدد الكبير من مراجعات المستخدمين DANA يُعد تطبيق على معلومات قيّمة تتمثل في الآراء والانتقادات والاقتراحات التي يمكن الاستفادة Google Play المنشورة على متجر منها لتقييم جودة الخدمات المقدمة. إلا أن كثرة هذه المراجعات تجعل تحليلها يدويًا أمرًا غير فعّال، مما يستدعي استخدام VADER (Valence Aware تقنيات تحليل المشاعر الآلي. تهدف هذه الدراسة إلى مقارنة أداء طريقة المعتمدة على Naïve Bayes المعتمدة على القاموس، وطريقة (Dictionary and Sentiment Reasoner التعلم الآلي، في تصنيف مشاعر مراجعات مستخدمي تطبيق DANA.

وكانت مصنفة مسبقًا وفق فئات Kaggle، اعتمدت الدراسة على مجموعة بيانات تضم 50,000 مراجعة جُمعت من منصة (Case Folding) المشاعر. وقد خضعت البيانات لعدة مراحل من المعالجة المسبقة شملت تحويل الأحرف إلى الصغيرة (Stopword) وإزالة الكلمات الشائعة ،(Tokenizing) وتقسيم النص إلى كلمات ،(Cleaning) والتنظيف لاستخراج الخصائص، TF-IDF بعد ذلك، استُخدمت طريقة .(Stemming) واستخراج الجذور ،(Removal اعتمادًا على مقاييس الدقة (Confusion Matrix) وتم تقييم أداء النموذجين باستخدام مصفوفة الالتباس وأظهرت النتائج أن .F1-Score ودرجة ،(Recall) والاسترجاع ،(Precision) والإحكام ،(Accuracy) التي حققت 67.76٪، بفارق VADER حققت دقة بلغت 77.14٪، متفوقةً على طريقة Naïve Bayes طريقة أكثر فاعلية في تصنيف مشاعر المراجعات المكتوبة باللغة Naïve Bayes وتشير هذه النتائج إلى أن طريقة .9.38٪ الإندونيسية، خاصةً تلك التي تتضمن قدرًا كبيرًا من اللغة غير الرسمية.

BAB I
PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan komunikasi yang begitu pesat dalam beberapa tahun terakhir telah membawa perubahan besar dalam berbagai aspek kehidupan manusia. Kemajuan ini tidak hanya berdampak pada sektor ekonomi dan pendidikan, tetapi juga memengaruhi cara masyarakat berinteraksi, mencari hiburan, serta menjalankan aktivitas sehari-hari. Salah satu bentuk nyata dari transformasi digital tersebut adalah meningkatnya penggunaan aplikasi *mobile*, yang kini menjadi bagian penting dari kehidupan masyarakat modern.

Perkembangan teknologi informasi dan komunikasi yang pesat telah mendorong transformasi digital di berbagai sektor, termasuk sektor keuangan. Salah satu bentuk inovasi yang berkembang pesat adalah *financial technology* (fintech), khususnya layanan dompet digital (*e-wallet*). Di Indonesia, penggunaan dompet digital mengalami peningkatan signifikan seiring dengan meningkatnya kebutuhan masyarakat akan transaksi yang cepat, praktis, dan aman. Salah satu aplikasi dompet digital yang banyak digunakan adalah aplikasi DANA, yang menyediakan berbagai layanan seperti pembayaran digital, transfer uang, pembelian pulsa, hingga pembayaran tagihan.

Seiring dengan meningkatnya jumlah pengguna aplikasi DANA, jumlah ulasan (review) yang diberikan oleh pengguna pada platform Google Playstore juga semakin meningkat. Ulasan tersebut mencerminkan opini, pengalaman, serta tingkat kepuasan pengguna terhadap layanan yang diberikan. Informasi yang

terkandung dalam ulasan pengguna sangat penting bagi pengembang aplikasi sebagai bahan evaluasi untuk meningkatkan kualitas layanan. Namun, volume data ulasan yang besar dan terus bertambah membuat proses analisis secara manual menjadi tidak efisien dan memakan waktu yang lama.

Untuk mengatasi permasalahan tersebut, diperlukan suatu pendekatan otomatis yang mampu mengolah dan menganalisis data teks dalam jumlah besar. Salah satu teknik yang dapat digunakan adalah analisis sentimen. Analisis sentimen merupakan bagian dari text mining yang bertujuan untuk mengidentifikasi, mengekstraksi, dan mengklasifikasikan opini atau emosi yang terkandung dalam suatu teks ke dalam kategori tertentu, seperti positif, negatif, atau netral Sentiment Analysis. Teknik ini telah banyak digunakan dalam berbagai bidang, termasuk analisis ulasan produk, media sosial, dan layanan digital.

Dalam implementasinya, terdapat berbagai metode yang dapat digunakan untuk melakukan analisis sentimen. Secara umum, metode tersebut dapat dibagi menjadi dua pendekatan utama, yaitu pendekatan berbasis lexicon dan pendekatan berbasis machine learning. Pendekatan berbasis *lexicon* menggunakan kamus sentimen yang berisi daftar kata beserta skor polaritasnya untuk menentukan sentimen suatu teks. Salah satu metode yang populer dalam pendekatan ini adalah VADER (*Valence Aware Dictionary and Sentiment Reasoner*). VADER dirancang khusus untuk menganalisis teks pendek seperti komentar atau ulasan, serta mampu menangani aspek-aspek seperti tanda baca, huruf kapital, dan intensitas kata (Hutto & Gilbert, 2014).

Keunggulan utama VADER adalah kemampuannya untuk melakukan analisis sentimen tanpa memerlukan data pelatihan (training data), sehingga lebih cepat dan efisien dalam implementasi. Namun, VADER pada dasarnya dikembangkan untuk bahasa Inggris, sehingga memiliki keterbatasan ketika diterapkan pada teks berbahasa Indonesia, terutama yang mengandung bahasa informal, singkatan, atau slang yang umum digunakan dalam ulasan pengguna aplikasi di Indonesia.

Di sisi lain, pendekatan berbasis machine learning menawarkan fleksibilitas yang lebih tinggi dalam menangani berbagai jenis bahasa, termasuk bahasa Indonesia. Salah satu metode yang banyak digunakan dalam klasifikasi teks adalah *Naive Bayes*. Metode ini bekerja berdasarkan teorema Bayes dengan asumsi independensi antar fitur, dan telah terbukti efektif dalam berbagai penelitian analisis sentimen karena kesederhanaan serta kecepatan komputasinya (Manning et al., 2008). *Naive Bayes* mampu mempelajari pola dari data pelatihan sehingga dapat menyesuaikan dengan karakteristik data yang digunakan.

Beberapa penelitian sebelumnya menunjukkan bahwa metode *Naive Bayes* memiliki performa yang cukup baik dalam klasifikasi sentimen teks berbahasa Indonesia, terutama ketika didukung oleh proses preprocessing yang optimal, seperti tokenisasi, stopword removal, dan stemming. Namun, metode ini memerlukan data pelatihan yang cukup agar dapat menghasilkan model yang akurat. Hal ini berbeda dengan VADER yang tidak memerlukan proses pelatihan, tetapi bergantung pada kualitas kamus sentimen yang digunakan.

Perbedaan karakteristik antara VADER dan *Naive Bayes* menjadikan kedua metode ini menarik untuk dibandingkan. VADER mewakili pendekatan berbasis lexicon yang cepat dan praktis, sedangkan *Naive Bayes* mewakili pendekatan *machine learning* yang adaptif terhadap data. Dengan membandingkan kedua metode tersebut, dapat diketahui metode mana yang lebih efektif dalam mengklasifikasikan sentimen ulasan aplikasi DANA pada *Google Playstore* Indonesia.

Perkembangan teknologi informasi telah mengubah cara masyarakat menyampaikan pendapat, kritik, dan saran melalui media digital, termasuk melalui ulasan pada aplikasi di Google Play Store. Ulasan tersebut merupakan bentuk komunikasi yang dapat memberikan manfaat bagi pengguna lain maupun pengembang aplikasi sebagai bahan evaluasi kualitas layanan. Dalam perspektif Islam, setiap bentuk komunikasi hendaknya dilakukan dengan memperhatikan etika, kejujuran, dan tanggung jawab agar tidak menimbulkan perselisihan maupun kerugian bagi orang lain. Allah SWT berfirman dalam

QS. Al-Isrā' ayat 53:

وَقُل لِّعِبَادِي يَقُولُوا الَّتِي هِيَ أَحْسَنُ ۚ إِنَّ الشَّيْطَانَ يَنزَغُ بَيْنَهُمْ ۚ إِنَّ الشَّيْطَانَ كَانَ لِلْإِنسَانِ عَدُوًّا مُّبِينًا

"Dan katakanlah kepada hamba-hamba-Ku agar mereka mengucapkan perkataan yang lebih baik. Sesungguhnya setan itu menimbulkan perselisihan di antara mereka. Sungguh, setan adalah musuh yang nyata bagi manusia." (QS. Al-Isrā' [17]: 53)

Menurut M. Quraish Shihab dalam Tafsir Al-Misbah, ayat tersebut mengajarkan bahwa setiap manusia diperintahkan untuk memilih perkataan yang baik, santun, dan tidak menimbulkan permusuhan. Kritik maupun pendapat tetap

diperbolehkan, tetapi harus disampaikan secara bijaksana dan bertanggung jawab. Dalam konteks penelitian ini, ulasan pengguna aplikasi DANA merupakan bentuk komunikasi digital yang mencerminkan pengalaman serta penilaian terhadap suatu layanan. Oleh karena itu, analisis sentimen terhadap ulasan pengguna dapat menjadi sarana untuk memahami berbagai bentuk ekspresi tersebut secara objektif tanpa dipengaruhi oleh penilaian pribadi peneliti.

Penelitian ini menjadi penting karena sebagian besar ulasan pengguna aplikasi di Indonesia menggunakan bahasa informal yang dapat memengaruhi kinerja metode analisis sentimen. Selain itu, hasil penelitian ini diharapkan dapat memberikan kontribusi dalam menentukan metode yang paling sesuai untuk analisis sentimen pada data ulasan berbahasa Indonesia, khususnya pada aplikasi fintech seperti DANA. Dengan demikian, pengembang aplikasi dapat memanfaatkan hasil analisis sentimen untuk meningkatkan kualitas layanan dan kepuasan pengguna.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dibahas pada sub-bab sebelumnya, dapat dinyatakan permasalahan dalam penelitian ini sebagai berikut :

a. Bagaimana penerapan metode VADER (*Valence Aware Dictionary and Sentiment Reasoner*) dan *Naïve Bayes* dalam mengklasifikasikan sentimen ulasan aplikasi DANA?
b. Bagaimana perbandingan kinerja antara metode VADER dan *Naive Bayes* dalam klasifikasi sentimen ulasan aplikasi DANA pada *Google Playstore*

Indonesia?

1.3 Batasan Masalah

a.Data yang digunakan hanya berupa ulasan teks dari *Google Play Store*.

b. Bahasa yang dianalisis adalah bahasa Indonesia

c. Analisis difokuskan pada klasifikasi sentimen positif, negatif, dan netral.

1.4 Tujuan Penelitian

Tujuan penelitian harus menyebutkan secara khas tujuan yang ingin dicapai. Dalam beberapa hal tujuan penelitian sudah tersirat di dalam judul penelitian. Tujuan penelitian disesuaikan dengan pernyataan masalah.

Berdasarkan rumusan masalah diatas, tujuan dari penelitian ini adalah sebagai berikut:

a. Menerapkan metode VADER dan *Naïve Bayes* untuk mengklasifikasikan ulasan pengguna ke dalam kategori positif, negatif, netral
b. Membandingkan kinaerja metode VADER dan *Naïve Bayes* dalam klasifikasi sentimen ulasan aplikasi

1.5 Manfaat Penelitian

a. **Bagi Pengembang aplikasi DANA:** Memberikan gambaran tentang kepuasan dan keluhan pengguna berdasarkan data ulasan nyata di *Google Play Store.*
b. **Bagi Akademisi:** Menjadi referensi penelitian di bidang analisis sentimen.
c. **Bagi Pengguna:** Meningkatkan kesadaran bahwa ulasan yang diberikan

dapat dimanfaatkan sebagai bahan evaluasi untuk perbaikan layanan aplikasi

1.6 Metode Penelitian

Pada metode penelitian ini terdapat uraian tentang pola dan rancangan penelitian, bahan atau materi penelitian, alat, jalannya penelitian, dan analisis hasil penelitian.

1. Pola dan rancangan penelitian seperti dalam proposal penelitian dan mungkin sudah disempurnakan.
2. Spesifikasi bahan atau materi penelitian harus dinyatakan selengkap-lengkapnya. Hal ini perlu dikemukakan agar peneliti lain yang ingin menguji ulang penelitian itu tidak sampai salah langkah.
3. Alat yang dipergunakan untuk melaksanakan penelitian diuraikan dengan disertakan spesifikasinya.
4. Langkah penelitian berupa uraian yang lengkap dan terinci tentang langkah-langkah yang telah diambil pada pelaksanaan penelitian, termasuk cara mengumpulkan data, jenis data, desain sistem, algoritma atau metode yang digunakan dan dikembangkan, berbagai macam flowchart maupun pseudo code yang dibuat, serta rancangan antar muka yang dikembangkan.

1.7 Sistematika Penulisan

Sistematika penulisan diberikan untuk menyajikan outline penulisan Skripsi, yang meliputi susunan bab dan topik-topik yang dituliskan.

BAB II
STUDI PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen adalah sebuah proses untuk menentukan sentimen atau opini dari seseorang yang diwujudkan dalam bentuk teks dan bisa di kategorikan sebagai sentimen positif, negatif, atau netral. Pengguna internet banyak menuliskan pengalaman, opini dan segala hal yang menjadi perhatian mereka. Ide dasar dari investigasi sentiment adalah untuk mendeteksi popularitas teks dokumen atau kalimat pendek dan mengklasifikasikannya dalam premis ini. (Arief & Imanuel, 2019)

a. Analisis Sentimen Bertingkat (*Graded Sentiment Analysis*)

   Analisis sentimen bertingkat merupakan tipe analisis yang memiliki penilaian spesifik. Tipe ini memungkinkan untuk memperluas kategori polaritas untuk menyertakan tingkat positif dan negatif yang berbeda yaitu: sangat positif, positif, netral, negatif, sangat negatif. Analisis sentimen bertingkat (*fine-grained*) dapat digunakan untuk menafsirkan bintang 5 dalam sebuah ulasan. Bintang 5 untuk menyatakan sangat positif, sedangkan bintang 1 untuk menyatakan sangat negatif.

b. Deteksi Emosi (*Emotion Detection*)

   Analisis sentimen pendeteksi emosi digunakan untuk mendeteksi emosi, seperti kebahagiaan, frustrasi, kemarahan, dan kesedihan. Banyak dari sistem pendeteksi emosi menggunakan *Lexicon* (sebuah daftar kata

dan emosi yang disampaikan) atau algoritma *machine learning* yang kompleks.

c. Analisis Sentimen Berbasis Aspek (*Aspect-based Sentiment Analysis*)

   Analisis sentimen berbasis aspek digunakan saat menganalisis sentimen teks untuk mengetahui aspek atau fitur tertentu yang disampaikan seseorang apakah itu positif, netral, atau negatif. Misalnya pada ulasan sebuah restoran, "Makanannya enak dan murah. Tapi pelayanan dan tempat masih sangat kurang nyaman". Dari contoh ulasan tersebut dapat dikatakan dari aspek makanan mendapat penilaian yang positif, namun pada aspek pelayanan dan tempat mendapat penilaian negatif.

d. Analisis Sentimen Multibahasa (*Multilingual Sentiment Analysis*)

   Analisis sentimen multibahasa digunakan untuk menganalisis teks dalam multibahasa. Tipe analisis sentimen ini dikatakan cukup sulit, karena melibatkan banyak *preprocessing* dan sumber daya. Sebagian besar sumber dayanya tersedia secara *online* (misalnya *Lexicon* sentimen), sementara yang lain perlu dibuat (misalnya corpora yang diterjemahkan atau algoritma *noise detection*), tetapi harus tahu dahulu cara membuat kode untuk menggunakannya.

Alasan pentingnya analisis sentimen karena manusia mulai mengekspresikan pikiran dan perasaan mereka lebih terbuka dari sebelumnya, analisis sentimen dengan cepat menjadi salah satu alat penting untuk memantau dan memahami sentimen di semua jenis data. Secara otomatis menganalisis *feedback* pelanggan, seperti pendapat mereka dalam tanggapan survei dan

percakapan di media sosial, memungkinkan *brand* untuk mempelajari apa yang membuat pelanggan senang atau tidak, sehingga *brand* dapat menyesuaikan produk dan layanan mereka untuk memenuhi kebutuhan pelanggan mereka (Sitio & Nadiyanti, 2022).

a) Menyortir Data dalam Skala

   Analisis sentimen membantu memproses sejumlah besar data tidak terstruktur dengan cara yang efisien dan hemat biaya.

b) Analisis *Real-Time*

   Analisis sentimen dapat mengidentifikasi isu-isu kritis secara *real-time*,sehingga dapat dengan segera mengambil tindakan selanjutnya.

c) Kriteria yang Konsisten

   Perusahaan dapat menerapkan kriteria yang sama pada data mereka, sehingga dapat membantu meningkatkan akurasi dan mendapatkan wawasan yang lebih baik.

Tantangan analisis sentimen adalah pada umumnya ulasan yang diberikan pelanggan ditulis dengan bahasa informal, pesan singkat yang menunjukkan isyarat yang terbatas tentang sentimen, serta akronim dan singkatan yang sering digunakan. Tantangan utama dari analisis sentimen *machine-based* seperti: Subjektivitas dan Nada, Konteks dan Polaritas, Ironi dan Sarkasme, Perbandingan, *Emoji*, Mendefinisikan Netral, dan Akurasi *Annotator* Manusia (Sitio & Nadiyanti, 2022).

2.2 *Pre-Processing*

Dalam penelitian ini di terapkan text preprocessing untuk data yang akan digunakan dalam proses analisa sentimen, dimana data yang kita proses akan kita ambil informasi yang terkandung didalamnya. Dalam hal sentimen penulisannya yaitu negatif atau positif. Guna memudahkan dalam mengelola data maka data perlu kita berikan analisa sentimen secara manual dengan membaca maksud dari kalimat yang ada dalam sentiment tersebut, sehingga dapat diberikan penilaian bahwa sentiment tersebut merupakan sentiment negatif atau positif. (H, 2015)

1. *Case Folding*

   Pada proses case folding, semua huruf pada tiap kalimat dirubah menjadi huruf kecil, karakter yang dianggap tidak valid seperti angka dan tanda baca dihilangkan. (Indraloka & Santosa, 2017)

2. *Cleansing*

   Tahap ini merupakan proses untuk menghapus URL, *@mention, #hashtag, link, double whitespaces, emoticon*, serta angka. Karakter- karakter ini dihapus karena dianggap sebagai noise dalam data dan tidak diperlukan dalam alanisis sentimen. *Library* yang digunakan adalah *Natural Language Toolkit* (NLTK) pada *Python.* (Indraloka & Santosa, 2017)

3. *Tokenizing*

   Tokenizing merupakan tahapan penguraian *string* teks menjadi *term* atau kata. Tujuan dari tokenization yaitu memisahkan kata-kata dalam sebuah paragraf, kalimat atau halaman ke dalam kata tunggal. (Najjichah et al, 2019)

4. *Normalize*

   Pada bagian *spelling normalization* berguna untuk melakukan perbaikan kata- kata yang disingkat maupun salah ejaan dengan bentuk tertentu dengan maksud yang sama, sebagai contoh pada kata “tidak” memiliki banyak bentuk penulisan yaitu tak, enggak, tdak, tdk, dan sebagainya, untuk salah ejaan sebegai contoh kata “tersedia” biasa salah penulisan dengan trsedia, trsdia, tersdia dan lain sebagainya. (Najjichah et al., 2019)

5. *Filtering*

   Filtering atau *stopword removal*, merupakan tahapan penghapusan kata-kata yang tidak relevan dalam penentuan topik sebuah dokumen dan yang sering muncul pada dokumen, misalnya „dan‟, „atau, „sebuah‟, „adalah‟, pada dokumen berbahasa Indonesia. (Najjichah et al., 2019)

6. *Stemming*

   Stemming merupakan tahapan pengubahan suatu kata menjadi akar katanya dengan menghilangkan imbuhan awal atau akhir pada kata tersebut. (Najjichah et al., 2019)

2.3 DANA

DANA merupakan salah satu aplikasi dompet digital (*e-wallet*) yang populer di Indonesia dan digunakan untuk mempermudah berbagai transaksi keuangan secara digital. Aplikasi ini dikembangkan oleh PT Espay Debit Indonesia Koe dan mulai diperkenalkan kepada publik pada tahun 2018.

DANA hadir sebagai solusi pembayaran non-tunai yang memungkinkan pengguna melakukan berbagai aktivitas keuangan melalui satu aplikasi. Layanan yang tersedia dalam aplikasi DANA antara lain meliputi:

-Pembayaran tagihan (listrik, air, internet)

-Pembelian pulsa dan paket data

-Transfer uang antar pengguna maupun ke rekening bank

-Pembayaran di merchant *offline* maupun online

-Fitur dompet digital untuk menyimpan saldo

Selain itu, DANA juga mendukung sistem pembayaran berbasis QR code melalui integrasi dengan QRIS (*Quick Response Code Indonesian Standard*), sehingga memudahkan pengguna dalam melakukan transaksi di berbagai merchant di Indonesia.

Dalam perkembangannya, DANA menjadi salah satu aplikasi fintech yang banyak digunakan oleh masyarakat Indonesia karena kemudahan penggunaan, keamanan transaksi, serta berbagai promo yang ditawarkan. Tingginya jumlah pengguna tersebut menyebabkan banyaknya ulasan yang diberikan pada platform Google Play Store. Ulasan tersebut berisi berbagai opini pengguna, mulai dari kepuasan terhadap layanan hingga keluhan terkait bug, error sistem, atau kendala transaksi.

Data ulasan pengguna aplikasi DANA pada *Google Playstore* menjadi sumber data yang sangat relevan untuk dilakukan analisis sentimen. Melalui analisis sentimen, ulasan-ulasan tersebut dapat diklasifikasikan ke dalam kategori

positif, negatif, dan netral, sehingga dapat memberikan gambaran umum mengenai persepsi pengguna terhadap kualitas layanan aplikasi.

Namun, karakteristik ulasan pengguna di *Playstore* yang cenderung singkat, informal, dan sering menggunakan bahasa sehari-hari atau slang menjadi tantangan tersendiri dalam proses analisis sentimen. Oleh karena itu, diperlukan metode yang tepat untuk mengolah data tersebut, seperti penggunaan metode berbasis lexicon seperti VADER dan metode berbasis *machine learning* seperti *Naive Bayes.*

Gambar 2.1 Logo aplikasi DANA

Dengan demikian, aplikasi DANA tidak hanya berfungsi sebagai objek penelitian, tetapi juga sebagai sumber data nyata yang dapat digunakan untuk mengevaluasi performa metode analisis sentimen dalam konteks penggunaan bahasa Indonesia di *platform digital.*

2.4 *Lexicon Based*

Lexicon Based merupakan metode yang berdasarkan kamus dengan tujuan untuk mendapatkan bobot kalimat pada data set sehingga bisa diketahui label kelas sentimen pada data set (Adhi et al., 2019). Kelebihan dari metode ini adalah label dari suatu kalimat bisa dilakukan secara otomatis sehingga akan

menghemat waktu jika data yang diolah adalah data set dengan jumlah yang besar (Azhar, 2018). Selain itu, memberikan sentimen dengan metode ini juga untuk menghindari bias dari opini pribadi.

Metode *Lexicon* memiliki kelebihan karena data yang berupa kata dari suatu kalimat dibandingkan langsung dengan kamus kata yang ada pada *Lexicon*. Apabila dalam suatu kalimat terdapat kata opini, kalimat tersebut dinyatakan sebagai kalimat opini (Adhi et al., 2019). Keuntungan *Lexicon* adalah tenaga kerja manual hanya diperlukan ketika membangun kamus sentimen saja (Kannan et al., 2016)

Namun *Lexicon* memiliki kekurangan, apabila kata pada suatu kalimat tidak ada pada *Lexicon*, maka kalimat tersebut dinyatakan bukan kalimat opini, padahal mungkin saja kalimat tersebut kalimat opini.

2.5 VADER

Tabel diberi nomor unit dengan format (x. y) dimana x adalah nomor urut bab dan y adalah nomor urut tabel.

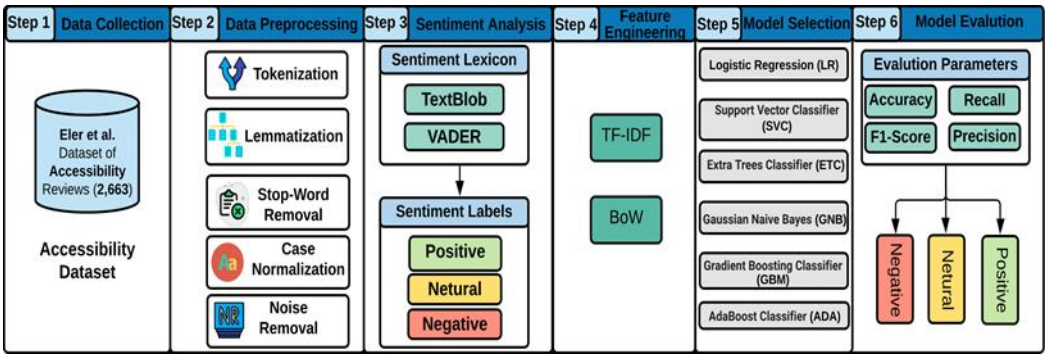
Valence Aware Dictionary and Sentiment Reasoner (VADER) adalah metode yang digunakan sebagai model untuk analisis sentimen dan mampu menentukan keragaman data melalui intensitas kekuatan emosional yang ada sesuai dengan kamus data *Lexicon* yang tersedia. (Elbagir & Yang, 2019).

Kamus lexicon VADER dapat digunakan untuk menilai sentimen frasa dan kalimat, tanpa perlu melihat yang lain. Sentimen dapat dikategorikan - seperti

{negatif, netral, positif} - atau dapat numerik - seperti kisaran intensitas atau skor. Pendekatan leksikal melihat kategori sentimen atau skor setiap kata

dalam sebuah kalimat dan memutuskan kategori atau skor sentimen keseluruhan kalimat itu. Kekuatan dari pendekatan leksikal terletak pada kenyataan bahwa tidak diperlukan melatih model menggunakan data berlabel. (Ghiassi & Lee, 2018) Keuntungan menggunakan VADER *polarity detection* adalah tersedianya kamus yang berisi nilai dari setiap kata. Hasil *Preprocess text* akan di nilai berdasarkan *lexicon* apakah itu positif, negatif atau netral dan menambahkan skor total (*compound*). Beberapa perintah VADER yang menggunakan bahasa pemrograman python akan dikerjakan, dan VADER akan memanggil data *lexicon* dari server NLTK untuk menghitung *polarity class sentimen*t. (Abimanyu et al., 2022)

Salah satu hal penting yang perlu dilakukan adalah langkah menerjemahkan data ke dalam bahasa inggris. Hal ini dikarenakan kamus VADER terdapat dalam bahasa inggris. Hal ini dilakukan oleh (Asri et al., 2022) dalam penelitiannya yang menggunakan dataset bahasa indonesia. Diagram alur VADER dengan translasi adalah sebagai berikut (Asri et al., 2022) :



Gambar 2.2 alur dari vader

a. Mengumpulkan dataset yang kemudian diterjemahkan ke dalam bahasa inggris.

b. Memecah kalimat menjadi per kata dan dibandingkan dengan tiap kata yang ada pada daftar kata opini, lalu kemudian dihitung *compound score*nya.
c. Kriteria pengelompokan positif, netral dan negatif yakni jika hasil compound score lebih dari 0,5 maka dimasukkan kategori positif yang diwakilkan dengan angka 1 lalu jika hasil compound score terletak diantara -0,5 dan 0,5 maka termasuk kategori netral yang diwakilkan dengan angka 0 dan yang terakhir jika hasil compound score dibawah -0,5 maka termasuk kategori negatif yang diwakilkan dengan angka -1
d. Apabila kata yang ada pada kalimat tersebut sama dengan kata yang ada pada daftar kata opini, kalimat tersebut dinyatakan kalimat opini.

2.6 *Naïve Bayes*

Naïve Bayes merupakan salah satu algoritma klasifikasi yang berdasarkan pada Teorema Bayes dengan asumsi bahwa setiap fitur atau atribut bersifat independen (tidak saling bergantung) terhadap fitur lainnya. Meskipun asumsi tersebut jarang terjadi pada data nyata, algoritma ini terbukti memiliki performa yang baik dalam berbagai kasus klasifikasi teks, termasuk analisis sentimen.

Metode *Naïve Bayes* banyak digunakan karena memiliki proses perhitungan yang sederhana, cepat, serta mampu menangani data dalam jumlah besar. Pada klasifikasi sentimen, algoritma ini digunakan untuk menentukan apakah suatu ulasan termasuk ke dalam kategori positif, netral, atau negatif berdasarkan probabilitas kemunculan kata-kata pada setiap kelas sentimen. Rumus *Naive bayes* dapat dilihat dalam persamaan (2.1):

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)} \quad (2.1)$$

Keterangan:

P(C|X) = Posterior Probability (probabilitas data X termasuk kelas C)

P(X|C) = Likelihood (probabilitas munculnya data X pada kelas C)

P(C) = Prior Probability (probabilitas awal kelas C)

P(X) = Evidence (probabilitas munculnya data X)

Dalam klasifikasi sentimen:

C = kelas sentimen (positif, negatif, netral)

X = kumpulan kata dalam suatu ulasan

Tujuan Naive Bayes adalah mencari nilai probabilitas terbesar dari setiap kelas sehingga dapat menentukan kelas sentimen suatu dokumen.

Konsep Kerja *Naive Bayes*

terdapat ulasan:"Aplikasi Dana sangat membantu transaksi"

Sistem akan menghitung probabilitas ulasan tersebut termasuk:

- Sentimen Positif
- Sentimen Negatif
- Sentimen Netral

Jika

$$P(Positif|X) = 0.75$$
$$P(Negatif|X) = 0.10 \quad (2.2)$$
$$P(Netral|X) = 0.15$$

Maka dokumen diklasifikasikan sebagai: Sentimen Positif karena memiliki probabilitas terbesar.

Keunikan Naive Bayes adalah asumsi independensi: terdapat kalimat: "Aplikasi Dana sangat bagus" *Naive Bayes* menganggap:

- Kata "Aplikasi"
- Kata "Dana"
- Kata "sangat"
- Kata "bagus"

tidak memiliki hubungan satu sama lain.

Sehingga:

$$P(X|C) = P(x_1|C) \times P(x_2|C) \times ... \times P(x_n|C) \quad (2.3)$$

Walaupun asumsi ini tidak sepenuhnya benar dalam bahasa alami, pada praktiknya Naive Bayes tetap memberikan performa yang baik dalam klasifikasi teks.

2.7 *Indonesian Sentiment Lexicon*

Indonesian Sentiment Lexicon digunakan untuk melakukan klasifikasi kelas pada data set. Suatu dokumen pada *data set* diklasifikasikan ke dalam tiga kelas, apakah data tersebut termasuk pada kelas positif, netral, atau negatif. Kata-kata yang ada pada suatu dokumen akan dibandingkan dengan dokumen kamus *lexicon.* Jika kata dalam dokumen terdapat di dalam kamus *lexicon,* maka kata tersebut akan diberikan nilai *score*. Jumlah nilai *score* pada suatu dokumen yang

menentukan label positif, netral, atau negatif. Kamus *lexicon* yang digunakan adalah *Indonesian Sentiment* (InSet) *Lexicon* yang bersumber dari Koto dan Rahmaningtyas. InSet *lexicon* berisi kurang lebih sekitar 10.250 kata yang diberi nilai dari -5 sampai dengan +5. (Koto & Rahmaningtyas, 2017).

Berikut adalah contoh beberapa kata yang terdapat pada kamus Indonesian Sentiment Lexicon :

1	word	weight	1	word	weight
2	hai	3	2	putus tali	-2
3	merekam	2	3	gelebah	-2
4	ekstensif	3	4	gobar hati	-2
5	paripurna	1	5	tersentuh	-1
6	detail	2	6	isak	-5
7	pernik	3	7	larat hati	-3
8	belas	2	8	nelangsa	-3
9	welas	4	9	remuk rec	-5
10	kabung	1	10	tidak sega	-2

Gambar 2.3 Kamus inset

2.8 TF-IDF

TF-IDF (*Term Frequency–Inverse Document Frequency*) merupakan metode pembobotan kata yang banyak digunakan dalam bidang *Text Mining*, *Information Retrieval*, dan Analisis Sentimen. Metode ini digunakan untuk mengubah data teks menjadi bentuk numerik sehingga dapat diproses oleh algoritma machine learning seperti *Naive Bayes*, *Support Vector Machine* (SVM), *Random Forest, maupun Logistic Regression*.

TF-IDF bekerja dengan memberikan bobot pada setiap kata berdasarkan dua faktor utama, yaitu frekuensi kemunculan kata dalam suatu dokumen (*Term Frequency*) dan tingkat keunikan kata tersebut pada seluruh kumpulan dokumen (*Inverse Document Frequency*). Semakin sering sebuah kata muncul dalam suatu

dokumen dan semakin jarang muncul pada dokumen lain, maka bobot TF-IDF kata tersebut akan semakin besar.

Menurut Ramos (2003), TF-IDF merupakan salah satu metode pembobotan yang paling efektif untuk mengukur tingkat kepentingan suatu kata dalam dokumen karena mampu mengurangi pengaruh kata-kata yang sering muncul namun kurang memiliki makna penting.

Pada tahap TF-IDF ini akan dilakukan pembobotan *term*, yang mana merupakan proses pembobotan atau pemberian nilai terhadap setiap *term* yang ada pada setiap data ulasan yang telah di-*preprocessing* sebelumnya. Pemberian nilai bobot ini menggunakan metode *Term Frecuency-Inverse*

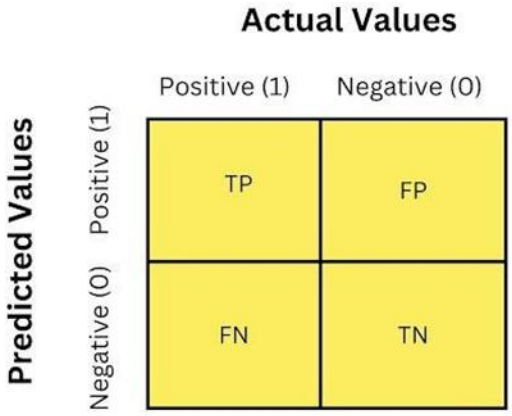
Document Frecuency (TF-IDF). Hasil dari pembobotan kata ini kemudian akan menjadi masukan (*input*) pada proses klasifikasi. (Gifari et al., 2022)

Library yang dibutuhkan dalam tahap ekstraksi fitur ini adalah *Scikit-Learn*, yang mana library ini merupakan library yang sangat terkenal yang digunakan pada pemodelan data, *machine learning* dan juga *deep learning*. Scikit-Learn berisi serangkaian modul atau fungsi yang dapat dipanggil untuk digunakan. Adapun modul atau fungsi yang digunakan pada tahap ini adalah *TfidfVectorizer*, dimana modul ini akan mengonversi kumpulan dokumen data ulasan menjadi matriks fitur TF-IDF. Setelah data melewati proses TF-IDF ini maka tiap kata pada setiap baris data akan memiliki nilai bobot. Nilai bobot tentunya dipengaruhi oleh seberapa sering kata atau seberapa banyak frekuensi data itu muncul dalam dataset yang kita miliki. (Musfiroh et al., 2021).

2.9 *Confusion Matrix*

TF-IDF (*Term Frequency–Inverse Document Frequency*) merupakan metode pembobotan kata yang banyak

Confussion Matrix adalah suatu metode yang biasanya digunakan untuk melakukan perhitungan akurasi pada konsep data mining. Confusion matrix digambarkan dengan tabel yang menyatakan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah diklasifikasikan (Rahman et al., 2017). Terdapat 4 metrik di dalam *confusion matrix*, yaitu :



Gambar 2.4 Gambar *Confusion Matrix*

1. *True Positive* (TP)

TP adalah jumlah prediksi yang berhasil diprediksi oleh model yang memprediksi kelas positif sebagai positif.

2. *True Negative* (TN)

TN adalah jumlah prediksi yang berhasil diprediksi oleh model yang memprediksi kelas negatif sebagai negatif.

3. *False Positive* (FP)

FP adalah jumlah prediksi yang gagal diprediksi oleh model yang memprediksi kelas negatif sebagai positif.

4. *False Negative* (FN)

FN adalah jumlah prediksi yang gagal diprediksi oleh model yang memprediksi kelas positif sebagai negatif.

Matrik-mtrik diatas dapat digunakan untuk menghitung nilai performa model seperti *accuracy*, *precision*, *recall*, dan F1 *score*. Menurut Han et al. (2022), Confusion Matrix merupakan salah satu metode evaluasi yang paling umum digunakan dalam machine learning karena mampu memberikan gambaran yang jelas mengenai performa model klasifikasi. Berikut ini adalah rumus-rumus nilai performa dapat dilihat dari persamaan (2.4):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

$$Precision = \frac{TP}{TP + FP} \tag{2.4}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

2.10 Penelitian Terkait

Literatur sejenis berisi penelitian yang telah dilakukan sebelumnya yang digunakan penulis sebagai referensi dalam penelitian karena memiliki keterkaitan dengan tema penelitian ini. Sumber-sumber penelitian sejenis diambil dari jurnal dan skripsi yang di dalamnya membahas mengenai analisis sentimen berbasis

aspek yang menggunakan metode *VADER, Indonesian Sentiment Lexicon, dan Naïve Bayes*

Tabel 2.1 Penelitian Terkait

Peneliti (Tahun)	**Judul / Fokus**	**Metode**	**Kelebihan**
Chaithra, (2019)	Hybrid approach: naive bayes and sentiment VADER for analyzing sentiment of mobile Unboxing video comments	*Vader dan Naïve Bayes Classifier*	hasil akurasi dari model bernilai cukup baik dari yaitu 79.78%
Anggina et al. (2022)	Analisis Ulasan Pelanggan Menggunakan Multinomial Naïve Bayes Classifier dengan Lexicon-Based dan TF-IDF Pada Formaggio Coffee and Resto	Indonesian Sentiment Lexicon dan Naïve Bayes	penelitian ini adalah akurasi dari model cukup baik karena nilainya 95%
Noviantho et al. (2021)	Analisis Sentimen Pada Komentar Video Ulasan Makanan Dari Saluran Youtube Berbahasa Indonesia Menggunakan K-nearest Neighbor	Indonesian Sentiment Lexicon dan K-nearest Neighbor	Kelebihan dari penelitian ini adalah akurasi yang dihasilkan sebesar 89%
Isnan et al. (2023)	Sentiment Analysis for TikTok Review Using VADER Sentiment and SVM Model	Vader dan SVM	Kelebihan dari penelitian ini adalah akurasi dari model didapat 80%
Elbagir & Yang (2019)	Twitter Sentiment Analysis Using Natural Language Toolkit and VADER Sentiment	Vader dan NLTK	Kelebihannya dikatakan bahwa Vader adalah metode yang tepat untuk digunakan

Berdasarkan penelitian yang dilakukan oleh Chaithra (2019) dan Anggina et al. (2022), metode Naïve Bayes menunjukkan kemampuan yang baik dalam melakukan klasifikasi sentimen dengan tingkat akurasi yang cukup tinggi. Selain itu, penggunaan metode pendukung seperti VADER, Lexicon-Based, dan TF-IDF terbukti dapat meningkatkan kualitas analisis sentimen. Namun, kedua penelitian masih memiliki keterbatasan, terutama pada jumlah dataset yang digunakan dan belum adanya perbandingan langsung antara metode VADER dan Naïve Bayes pada ulasan aplikasi layanan keuangan digital. Oleh karena itu, penelitian ini dilakukan untuk membandingkan performa metode VADER dan Naïve Bayes dalam klasifikasi sentimen ulasan aplikasi DANA menggunakan dataset yang lebih besar dan evaluasi yang lebih komprehensif melalui metrik Accuracy, Precision, Recall, dan F1-Score.

BAB III
METODE PENELITIAN

3.1 Dataset

Dalam penelitian ini, data yang digunakan merupakan dataset ulasan pengguna aplikasi DANA yang diperoleh melalui platform Kaggle. Kaggle merupakan salah satu platform penyedia dataset yang sering digunakan dalam penelitian di bidang data science dan machine learning, sehingga memudahkan peneliti dalam memperoleh data dalam jumlah besar, dengan dataset pada penelitian ini berjumlah 50.000 data.

Dataset yang digunakan berisi kumpulan ulasan pengguna terhadap aplikasi DANA yang dikumpulkan dari berbagai sumber. Data tersebut terdiri dari beberapa atribut seperti username, komentar ulasan, label, serta waktu ulasan. Dataset ini kemudian digunakan sebagai bahan utama dalam proses analisis sentimen.

Pemilihan dataset dari Kaggle dilakukan karena dataset yang tersedia telah terstruktur dengan baik dan dapat langsung digunakan untuk proses analisis. Selain itu, penggunaan dataset dari Kaggle juga mempermudah proses pengolahan data karena telah tersedia dalam format yang siap diproses menggunakan bahasa pemrograman Python. Berikut ini adalah sampel data dari dataset yang akan diteliti.

Tabel 3.1 *Sampel* Data Ulasan Aplikasi DANA dari *Kaggle*

NO	USERNAME	COMENT	LABEL	WAKTU
1	Elisya Kasni	Bagus	POSITIVE	06/01/2025 06.09
2	Rusman Man	Dana mmg keren mantap.	POSITIVE	28/02/2025 01.09
3	Qiliw Sadega	Saya ngajuin upgrade premium krna ktp m jdi ga bisa verifikasi..trus coba lewat email di suruh nunggu 3 hri .ini udh 3 hri lebih msih gituÂ² doang saldo sya d dana jdi ga bisa d apaÂ²in.. gmna ini solusi nya	NEGATIVE	05/03/2025 12.35
4	Kijutjrv2 Kijut	Kocak mana diskon nya ml malah eror segala kaga ikhlas ngasih diskon nya	NEGATIVE	18/01/2025 14.58
5	Fifi Alfiyah	Saldo hilang karena no lama Hilang ganti no saldonya gk ada sama sekali dana tidak bertanggung jawab Buktinya saldo saya gk ada !!! Saldo lama saya 1800.000	NEGATIVE	12/01/2025 14.19
6	Kiki57	mayan	POSOTIVE	28/02/2025 22.24
7	Parhan Parhan	Udah gua hapus dana ya. ilang ya udah 1 juta lebih duwit gua.mf dana sayah buka orang kaya itu uang keringat sayah uang halal	NEGATIVE	13/02/2025 09.05
8	Dewi Anggreni	baik	POSITIVE	25/01/2025 13.27
9	Bang Ewok13	TOLONG UNTUK SISTEM KEAMANAN DI PERBAIKI. KALAU 1 ATAU 2 ORNG YG KE HACK WAJAR, MUNGKIN KETELEDORAN PENGGUNA. TAPI KALAU SUDAH BANYAK YG TERKENA HACK. BERARTI ADA YG SALAH DENGAN SISTEM KEAMANAN DANA	NEUTRAL	27/01/2025 03.22
10	M Alifian	mempermudah transfer	POSITIVE	22/02/2025 15.51

3.2 Data *Preprocessing*

Preprocessing data adalah tahap pembersihan data agar lebih mudah dianalisis. Langkah ini mencakup beberapa proses yaitu

a. *Case Folding*

Case Folding adalah proses mengubah seluruh huruf dalam teks menjadi huruf kecil (*lowercase*). Tujuannya adalah untuk menghindari perbedaan antara huruf kapital dan huruf kecil. Contoh:

Sebelum:

"Aplikasi DANA Sangat Membantu"

Sesudah:

"aplikasi dana sangat membantu"

Pada tahap ini seluruh karakter selain huruf juga dapat dihapus sesuai kebutuhan. Manfaatnya untuk menyeragamkan bentuk kata dan mengurangi jumlah variasi kata dalam dataset.

b. *Cleansing*

Cleansing merupakan proses pembersihan teks dari karakter yang tidak diperlukan.Karakter yang biasanya dihapus meliputi:Tanda baca , Angka, Emoji , Simbol , URL, *Mention* (@) , *Hashtag* (#)

Sebelum:

"Aplikasi DANA bagus!!! 👍👍"

"Kunjungi https://dana.id sekarang"

Sesudah:

"aplikasi dana bagus"

“kunjungi sekarang”

Manfaatnya untuk menghilangkan noise pada data dan mengurangi karakter yang tidak memiliki pengaruh terhadap sentimen.

c. *Tokenizing*

Tokenizing adalah proses memecah kalimat menjadi unit-unit kata (token).

Sebelum:

“aplikasi dana sangat membantu”

Sesudah:

['aplikasi', 'dana', 'sangat', 'membantu']

Setiap kata akan diproses secara terpisah pada tahap berikutnya.

Manfaatnya mempermudah analisis setiap kata dan menjadi dasar untuk perhitungan TF-IDF dan Naïve Bayes.

d. *Stopword Removal*

Stopword Removal adalah proses menghapus kata-kata yang sering muncul tetapi memiliki makna yang kurang penting dalam analisis sentimen.

Contoh stopword Bahasa Indonesia: (yang, dan, di, ke, dari, untuk, atau, adalah, itu, ini)

Sebelum:

“aplikasi dana yang sangat membantu”

Sesudah:

“aplikasi dana sangat membantu”

Kata "yang" dihapus karena tidak memberikan informasi sentimen yang signifikan. Manfaatnya mengurangi jumlah kata yang tidak informatif, mempercepat proses komputasi dan meningkatkan kualitas fitur.

e. *Stemming*

Stemming merupakan proses mengubah kata berimbuhan menjadi kata dasar.

Sebelum:

“aplikasi ini sangat membantu pembayaran”

Sesudah:

“aplikasi sangat bantu bayar”

Manfaatnya adalah menyatukan kata yang memiliki makna sama, mengurangi jumlah vocabulary dan meningkatkan efisiensi klasifikasi.

3.3 Ekstraksi Fitur

Tahapan eikstrasi fituir meingguinakan algoritma peimbobotan kata *Teirm Freiquieincy-Inveirse*i *Docuimeint Freiquieincy* (TF-IDF) yang fuingsinya adalah meinguibah teiks meinjadi veictor. Seitiap kata yang teirdapat pada uilasan akan dibeiri bobot seisuiai deingan peirhituinganya meingguinakan algoritma ini.

Dalam penelitian analisis sentimen ulasan aplikasi DANA, proses pembentukan TF-IDF dilakukan melalui beberapa tahapan sebagai berikut:

Pengumpulan Data

Data berupa ulasan pengguna aplikasi DANA diperoleh dari Google Play Store yang diambil dari *Kaggle.*

Tabel 3.2 *Sampel* ulasan

NO	ULASAN
1	aplikasi dana sangat membantu
2	transaksi cepat dan aman
3	aplikasi sering error

Pembentukan *Vocabulary* semua kata unik dikumpulkan terdapat pada tabel 3.3 dan *matrix* TF-IDF pada tabel 3.4

Tabel 3.3 *Matrix* TF-IDF

Kata	**D1**	**D2**	**D3**
aplikasi	0.15	0	0.20
dana	0.25	0	0
bantu	0.40	0	0
cepat	0	0.50	0
aman	0	0.45	0
error	0	0	0.60

Tabel 3.4 *Vocab*

Vocabulary
aplikasi
dana
bantu
transaksi
cepat
aman
error

3.4 Klasifikasi

Setelah data ulasan melalui tahap preprocessing, proses selanjutnya adalah klasifikasi sentimen. Pada penelitian ini, klasifikasi dilakukan menggunakan dua metode, yaitu metode VADER (*Valence Aware Dictionary and Sentiment Reasoner*) dan metode Naive Bayes. Sebelum proses klasifikasi menggunakan metode Naive Bayes dilakukan, dataset dibagi

terlebih dahulu menjadi data latih (*training data*) dan data uji (*testing data*) menggunakan fungsi train_test_split pada library Python. Pembagian data dilakukan dengan proporsi 80% data latih dan 20% data uji. Data latih digunakan untuk membangun model klasifikasi, sedangkan data uji digunakan untuk mengukur performa model yang telah dibuat.

Metode Naive Bayes bekerja dengan mempelajari pola kata dari data latih untuk menentukan kategori sentimen pada data uji. Sedangkan metode VADER melakukan klasifikasi sentimen berdasarkan skor polaritas kata yang terdapat pada teks ulasan. Hasil klasifikasi dari kedua metode kemudian dibandingkan untuk mengetahui metode yang memiliki performa terbaik dalam mengklasifikasikan sentimen ulasan aplikasi DANA pada Google Playstore Indonesia.

Secara umum, proses klasifikasi pada penelitian analisis sentimen terdiri dari beberapa tahapan.

1. Data Ulasan

Data diperoleh dari ulasan pengguna aplikasi DANA di Google Play Store dan diambil dari kaggle dengan dataset berjumlah 50.000 data

Tabel 3.5 Contoh Ulasan

No	Ulasan
1	Aplikasi sangat membantu
2	Sering gagal login
3	Lumayan untuk transaksi

2. Preprocessing

Data dibersihkan melalui:

- Case Folding

- Cleansing

- Tokenizing

- Stopword Removal

- Stemming

Contoh:

“Aplikasi sangat membantu”

menjadi

“aplikasi bantu”

3. Ekstraksi Fitur TF-IDF

Setiap dokumen diubah menjadi bentuk numerik menggunakan TF-IDF.

Tabel 3.6 Bentuk Numerik

Kata	Bobot
aplikasi	0.25
bantu	0.70

Hasil TF-IDF berupa matriks numerik yang dapat diproses oleh algoritma klasifikasi.

4. Pembagian Data

Dataset dibagi menjadi:

-Training Data

-Testing Data

Umumnya menggunakan rasio:

80% : 20%

Training data digunakan untuk melatih model, sedangkan testing data digunakan untuk menguji performa model.

5. Pelatihan Model (*Training*)

Pada tahap ini algoritma Naïve Bayes mempelajari pola dari data training.

- ***Probabilitas Prior***

Model akan menghitung probabilitas setiap kelas:

Tabel 3.7 Contoh Jumlah

Kelas	Jumlah
Positif	500
Negatif	300
Netral	200

Maka:

$$P(Positif) = 0.5$$

$$P(Negatif) = 0.3$$

$$P(Netral) = 0.2$$

- ***Probabilitas Likelihood***

Menghitung peluang kemunculan kata pada masing-masing kelas.

Tabel 3.8 Contoh Peluang

Kata	Positif
bantu	0.04
cepat	0.03

Probabilitas ini digunakan dalam proses prediksi.

6. Prediksi (*Testing*)

Setelah model dilatih, dilakukan pengujian menggunakan data testing.

Misalnya terdapat ulasan:

“aplikasi sangat membantu transaksi”

Setelah preprocessing:

“aplikasi bantu transaksi”

Model menghitung probabilitas:

Kelas Positif : $P(Positif|X)$

Kelas Negatif : $P(Negatif|X)$

Kelas Netral : $P(Netral|X)$

Kelas dengan probabilitas terbesar dipilih sebagai hasil klasifikasi.

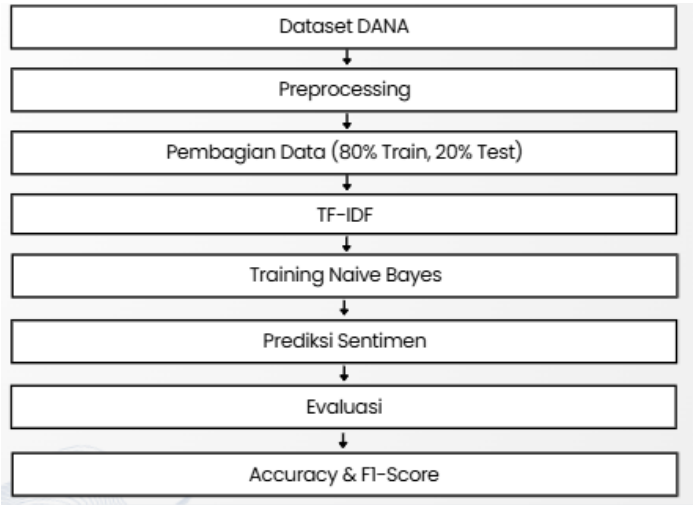
Tabel 3.9 Contoh Probabilitas

Kelas	Probabilitas
Positif	0.75
Negatif	0.10
Netral	0.15

Maka: Sentimen = Positif

- **Klasifikasi Menggunakan Naïve Bayes**

Pada penelitian ini, proses klasifikasi sentimen dilakukan menggunakan algoritma **Multinomial Naïve Bayes**. Algoritma ini dipilih karena merupakan salah satu metode machine learning yang paling banyak digunakan dalam analisis teks dan terbukti memiliki performa yang baik dalam mengklasifikasikan dokumen berdasarkan frekuensi kemunculan kata. Selain itu, Multinomial Naïve Bayes sangat cocok digunakan pada data yang telah ditransformasikan menggunakan metode **TF-IDF**, karena algoritma ini mampu mengolah data numerik yang merepresentasikan bobot setiap kata dalam dokumen.



Gambar 3.1 Alur Naïve Bayes

Naïve Bayes merupakan algoritma klasifikasi yang didasarkan pada **Teorema Bayes**, yaitu suatu metode probabilistik yang digunakan untuk menghitung peluang suatu data termasuk ke dalam kelas tertentu berdasarkan informasi yang dimiliki. Algoritma ini bekerja dengan menghitung probabilitas kemunculan kata-kata pada masing-masing kelas sentimen, kemudian menentukan kelas yang memiliki probabilitas tertinggi sebagai hasil klasifikasi.

Prinsip kerja Naïve Bayes adalah menghitung probabilitas posterior setiap kelas menggunakan Teorema Bayes:

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)} \tag{3.1}$$

Keterangan:

P(C|X) = probabilitas kelas setelah melihat data

P(X|C) = probabilitas data pada kelas tertentu

P(C) = probabilitas awal kelas

P(X) = probabilitas data

Model akan memilih kelas dengan nilai probabilitas tertinggi.

- **Output Tahap Klasifikasi**

Output dari proses klasifikasi berupa label sentimen.

Tabel 3.10 Contoh *Output*

Ulasan	Hasil
Aplikasi sangat membantu	Positif
Tidak bisa transfer	Negatif
Fitur cukup lengkap	Netral

Hasil klasifikasi inilah yang kemudian digunakan untuk proses evaluasi model.

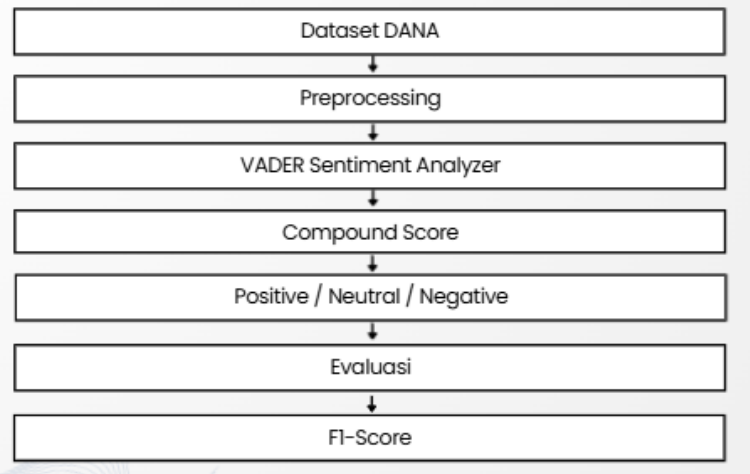
- **Klasifikasi Menggunakan VADER**

Pada penelitian ini, salah satu metode yang digunakan untuk melakukan klasifikasi sentimen adalah VADER (Valence Aware Dictionary and Sentiment Reasoner). VADER merupakan metode analisis sentimen berbasis leksikon (*lexicon-based sentiment analysis*) yang dikembangkan oleh C. J. Hutto dan Eric Gilbert pada tahun 2014. Metode ini dirancang untuk menganalisis sentimen pada teks pendek seperti komentar media sosial, ulasan pengguna, dan berbagai bentuk komunikasi digital lainnya.

Berbeda dengan metode machine learning seperti Naïve Bayes yang memerlukan proses pelatihan model menggunakan data training, VADER tidak memerlukan proses pembelajaran terlebih dahulu. Metode ini bekerja dengan memanfaatkan kamus sentimen (*sentiment lexicon*) yang telah berisi ribuan kata beserta nilai polaritasnya. Setiap kata dalam kamus memiliki skor sentimen yang menunjukkan kecenderungan positif, negatif, atau netral. Dengan demikian, proses klasifikasi dapat dilakukan secara langsung berdasarkan skor sentimen

yang diperoleh dari kata-kata yang terdapat pada suatu ulasan. Berikut adalah alur VADER

Gambar 3.2 Alur VADER



Dalam penelitian ini, VADER digunakan untuk mengidentifikasi sentimen ulasan pengguna aplikasi DANA yang diperoleh dari dataset. Setiap ulasan akan dianalisis untuk menghasilkan empat jenis skor sentimen, yaitu:

1. **Positive (Positif)**, yaitu skor yang menunjukkan tingkat sentimen positif dalam suatu teks.
2. **Negative (Negatif)**, yaitu skor yang menunjukkan tingkat sentimen negatif dalam suatu teks.
3. **Neutral (Netral)**, yaitu skor yang menunjukkan tingkat sentimen netral dalam suatu teks.
4. **Compound Score**, yaitu skor gabungan yang diperoleh dari keseluruhan kata dalam teks dan digunakan sebagai dasar penentuan kelas sentimen.

Skor **compound** memiliki rentang nilai antara -1 hingga +1. Semakin mendekati +1 menunjukkan sentimen yang semakin positif, sedangkan semakin mendekati -1

menunjukkan sentimen yang semakin negatif. Dalam penelitian ini, penentuan label sentimen dilakukan menggunakan aturan sebagai berikut:

Tabel 3.11 Skor *Compound*

Nilai Compound	Kategori Sentimen
Compound ≥ 0,05	Positif
Compound ≤ -0,05	Negatif
-0,05 < Compound < 0,05	Netral

Aturan tersebut merupakan standar yang direkomendasikan oleh pengembang VADER dan banyak digunakan dalam penelitian analisis sentimen.

Tahapan penggunaan VADER dalam penelitian ini diawali dengan melakukan preprocessing terhadap data ulasan. Tahap preprocessing meliputi case folding, cleaning, tokenizing, stopword removal, dan stemming untuk menghasilkan data yang lebih bersih dan terstruktur. Setelah proses preprocessing selesai, setiap ulasan dianalisis menggunakan library **vaderSentiment** pada Python. Library tersebut akan menghitung skor sentimen berdasarkan kata-kata yang terdapat dalam ulasan. Sebagai contoh, ulasan:

"This application is very useful and easy to use"

akan menghasilkan skor compound yang cenderung positif karena mengandung kata-kata seperti *useful* dan *easy* yang memiliki nilai sentimen positif dalam kamus VADER. Sebaliknya, ulasan:

"Bad service and many errors"

akan menghasilkan skor compound negatif karena mengandung kata-kata seperti *bad* dan *errors* yang memiliki polaritas negatif.

Setelah seluruh ulasan memperoleh label sentimen dari VADER, hasil klasifikasi dibandingkan dengan label sentimen asli pada dataset. Proses evaluasi dilakukan

menggunakan **confusion matrix**, **accuracy**, **precision**, **recall**, dan **F1-score**. Evaluasi ini bertujuan untuk mengetahui tingkat kinerja metode VADER dalam mengklasifikasikan sentimen ulasan pengguna aplikasi DANA.

Kelebihan utama metode VADER adalah proses implementasinya yang sederhana, cepat, dan tidak memerlukan data pelatihan. Selain itu, VADER mampu memberikan hasil klasifikasi secara langsung berdasarkan kamus sentimen yang telah tersedia. Namun, metode ini memiliki keterbatasan karena sangat bergantung pada kelengkapan kamus sentimen yang digunakan. Apabila terdapat kata-kata yang tidak terdapat dalam kamus atau penggunaan bahasa informal yang beragam, maka akurasi klasifikasi dapat menurun. Oleh karena itu, dalam penelitian ini metode VADER dibandingkan dengan Naïve Bayes untuk mengetahui metode yang memberikan performa terbaik dalam klasifikasi sentimen ulasan pengguna aplikasi DANA.

3.5 Evaluasi

Hasil pemodelan sebelumnya kemudian perlu dilakukan evaluasi menggunakan confuision matrix untuk mengukur performa model dalam meimpreidiksi data uiji, hasil dari proseis ini teirdiri dari beberapa parameter yaitui, *accuiracy*, *preicision*, *reicall* dan *f1-score*i.

Contoh Confusion Matrix

Misalkan terdapat 100 data ulasan. Hasil pengujian model menghasilkan:

Tabel 3.12 Contoh Pengujian

Actual / Predicted	**Positif**	**Negatif**
Positif	40	10
Negatif	5	45

Maka:

TP = 40

FN = 10

FP = 5

TN = 45

- ***Accuracy***

Accuracy menunjukkan persentase data yang berhasil diklasifikasikan dengan benar. Berikut perhitungannya dapat dilihat pada persamaan (3.2)

$$\begin{aligned} Accuracy &= \frac{40+45}{40+45+5+10} \times 100\% \\ &= \frac{85}{100} \times 100\% \\ &= 85\% \end{aligned} \tag{3.2}$$

Artinya model mampu mengklasifikasikan 85% data dengan benar.

- ***Precision***

Precision menunjukkan tingkat ketepatan model ketika memprediksi suatu kelas, dan berikut adalah rumus perhitungan presisi yang dapat dilihat pada persamaan yang ada dibawah ini pada persamaan (3.3)

$$\begin{aligned} Precision &= \frac{40}{40+5} \\ &= \frac{40}{45} \\ &= 0.889 \\ &= 88.9\% \end{aligned} \tag{3.3}$$

Artinya dari seluruh data yang diprediksi positif, sebanyak 88,9% benar-benar positif.

- ***Recall***

Recall digunakan untuk mengukur kemampuan model menemukan seluruh data positif. Berikut perhitungannya yang dapat dilihat pada persamaan (3.4)

$$\begin{aligned} Recall &= \frac{40}{40+10} \\ &= \frac{40}{50} \\ &= 0.80 \\ &= 80\% \end{aligned} \tag{3.4}$$

Artinya model berhasil menemukan 80% data positif yang sebenarnya.

- ***F1 Score***

F1-Score merupakan rata-rata harmonik antara Precision dan Recall. Berikut rumus *F1-Score* yang dapat dilihat pada persamaan (3.5):

$$\begin{aligned} F1 &= 2 \times \frac{0.889 \times 0.80}{0.889 + 0.80} \\ &= 0.842 \\ &= 84.2\% \end{aligned} \tag{3.5}$$

Semakin tinggi nilai F1-Score, semakin baik keseimbangan antara Precision dan Recall. Pada penelitian analisis sentimen ulasan aplikasi DANA, klasifikasi dilakukan terhadap tiga kelas yaitu:

-Positif

-Negatif

-Netral

Sehingga bentuk Confusion Matrix menjadi:

Tabel 3.13 Confusion Matrix

Actual / Predicted	Positif	Negatif	Netral
Positif	300	20	15
Negatif	30	250	25
Netral	10	20	180

Interpretasi:

- 300 ulasan positif diprediksi positif dengan benar.
- 20 ulasan positif salah diprediksi negatif.
- 15 ulasan positif salah diprediksi netral.
- 250 ulasan negatif diprediksi negatif dengan benar.
- 180 ulasan netral diprediksi netral dengan benar.

Pada kasus multikelas, perhitungan Precision, Recall, dan F1-Score dilakukan untuk masing-masing kelas kemudian dirata-ratakan menggunakan metode:

1. Macro Average
2. Micro Average
3. Weighted Average

Confusion Matrix merupakan metode evaluasi yang digunakan untuk mengukur performa model klasifikasi dengan membandingkan hasil prediksi terhadap data aktual. Metode ini menghasilkan informasi mengenai jumlah prediksi yang benar maupun salah dalam bentuk matriks, sehingga memungkinkan perhitungan metrik evaluasi seperti Accuracy, Precision, Recall, dan F1-Score. Dalam penelitian ini, Confusion Matrix digunakan untuk mengevaluasi kinerja metode Naive Bayes dan VADER dalam mengklasifikasikan sentimen ulasan aplikasi DANA ke dalam kategori positif,

negatif, dan netral. Dengan menggunakan Confusion Matrix, performa masing-masing metode dapat dianalisis secara lebih mendalam berdasarkan tingkat ketepatan dan kemampuan model dalam mengenali setiap kelas sentimen.

BAB IV
HASIL DAN PEMBAHASAN

4.1 Implementasi Sistem

Implementasi sistem merupakan tahap penerapan metode yang telah dirancang pada BAB III ke dalam bentuk program yang dapat melakukan analisis sentimen secara otomatis. Sistem dibangun menggunakan bahasa pemrograman Python pada lingkungan Google Colab dengan memanfaatkan beberapa pustaka (library) seperti Pandas, NLTK, Sastrawi, Scikit-learn, dan VaderSentiment.

Sistem yang dibangun memiliki kemampuan untuk mengolah data ulasan aplikasi DANA mulai dari proses preprocessing, ekstraksi fitur menggunakan TF-IDF, klasifikasi menggunakan metode VADER dan Naïve Bayes, hingga evaluasi hasil klasifikasi menggunakan Confusion Matrix.

4.2 Implementasi Data

Data penelitian diperoleh dari ulasan pengguna aplikasi DANA pada Google Play Store menggunakan dataset yang diambil dari *Kaggle* yang berhasil dikumpulkan terdiri atas username, isi ulasan, label sentimen, dan waktu publikasi ulasan. Data yang diperoleh kemudian melalui tahap preprocessing, ekstraksi fitur, klasifikasi sentimen menggunakan metode VADER dan Naïve Bayes, serta evaluasi performa model.

Dataset yang digunakan dalam penelitian ini diperoleh dari platform Kaggle yang berisi kumpulan ulasan pengguna aplikasi DANA. Dataset tersebut telah tersedia dalam format terstruktur dan terdiri atas beberapa atribut, yaitu username, isi ulasan (review), label sentimen, dan waktu publikasi ulasan. Label

sentimen pada dataset telah ditentukan sebelumnya ke dalam kategori positif, negatif, dan netral sehingga peneliti tidak perlu melakukan proses pelabelan secara manual.

Sebelum digunakan dalam penelitian, dataset terlebih dahulu dilakukan proses seleksi untuk memastikan kualitas data yang digunakan. Tahap seleksi meliputi pemeriksaan data kosong (missing value), data duplikat, serta kesesuaian format data. Data yang tidak memenuhi kriteria penelitian dihapus sehingga diperoleh dataset yang siap untuk diproses pada tahap selanjutnya.

Dataset yang digunakan menggambarkan berbagai pengalaman dan opini pengguna terhadap layanan aplikasi DANA. Ulasan dengan sentimen positif umumnya berisi apresiasi terhadap kemudahan penggunaan aplikasi, kecepatan transaksi, serta fitur-fitur yang disediakan. Sementara itu, ulasan dengan sentimen negatif umumnya berkaitan dengan kendala verifikasi akun, kegagalan transaksi, gangguan sistem, maupun permasalahan keamanan akun. Selain itu, terdapat pula ulasan dengan sentimen netral yang berisi saran, pertanyaan, atau pernyataan informatif tanpa menunjukkan kecenderungan emosi yang kuat.

Keberadaan label sentimen pada dataset memungkinkan penelitian ini untuk melakukan evaluasi performa metode VADER dan Naïve Bayes secara lebih objektif. Label yang tersedia digunakan sebagai acuan (ground truth) untuk membandingkan hasil prediksi yang dihasilkan oleh kedua metode klasifikasi sentimen tersebut.

Tabel 4.1 Sampel Data

No	Username	Ulasan	Label
1	Elisya Kasni	Bagus	Positif
2	Rusman Man	Dana mmg keren mantap	Positif
3	Qiliw Sadega	Saya ngajuin upgrade premium krna ktp m jdi ga bisa verifikasi, saldo saya tidak bisa digunakan	Negatif
4	Kijutjrv2 Kijut	Kocak mana diskonnya malah error segala kaga ikhlas ngasih diskon	Negatif
5	M Alifian	Mempermudah transfer	Positif

Berdasarkan sampel data tersebut, terlihat bahwa pengguna memberikan berbagai tanggapan mengenai aplikasi DANA. Sentimen positif umumnya berkaitan dengan kemudahan penggunaan dan fitur aplikasi, sedangkan sentimen negatif lebih banyak berkaitan dengan kendala verifikasi akun, error sistem, dan permasalahan transaksi.

Dataset yang digunakan dalam penelitian ini diperoleh dari platform Kaggle yang berisi ulasan pengguna aplikasi DANA beserta label sentimennya. Dataset tersebut telah dikategorikan ke dalam tiga kelas sentimen, yaitu *positive*, *negative*, dan *neutral*. Setiap data terdiri atas informasi berupa nama pengguna (*username*), isi ulasan (*review*), label sentimen (*sentiment*), dan waktu publikasi ulasan.

Sebelum dilakukan proses analisis sentimen, seluruh ulasan yang awalnya menggunakan Bahasa Indonesia diterjemahkan ke dalam Bahasa Inggris. Proses penerjemahan dilakukan karena metode VADER merupakan metode berbasis leksikon yang dikembangkan khusus untuk teks berbahasa Inggris sehingga dapat bekerja lebih optimal ketika menerima input dalam bahasa tersebut. Selain itu,

penggunaan Bahasa Inggris juga memudahkan proses preprocessing dan ekstraksi fitur karena sebagian besar sumber daya NLP seperti stopword dan stemmer telah tersedia secara lengkap. Contoh data setelah proses penerjemahan dapat dilihat pada Tabel 4.2

Tabel 4.2 Sampel Data bahasa inggris

No	Review (English)	Sentiment
1	Good	Positive
2	DANA is really great and awesome	Positive
3	I applied for premium verification but it still has not been approved and my balance cannot be used	Negative
4	Where is the discount? The application keeps getting errors	Negative
5	Please improve the security system because many users have experienced hacking issues	Neutral

Berdasarkan Tabel 4.2 dapat dilihat bahwa setiap ulasan telah diterjemahkan ke dalam Bahasa Inggris tanpa mengubah makna asli yang terkandung dalam ulasan tersebut. Hasil penerjemahan ini kemudian digunakan sebagai data masukan pada tahap preprocessing dan klasifikasi sentimen.

4.3 Hasil *Preprocessing* Data

Preprocessing dilakukan untuk membersihkan data teks agar dapat diproses oleh algoritma klasifikasi. Tahapan preprocessing yang diterapkan meliputi case folding, cleansing, tokenizing, stopword removal, dan stemming. Pada tahap ini terbagi menjadi 2 yaitu bahasa Indonesia dan inggris.

Setelah proses penerjemahan selesai, data ulasan diproses menggunakan beberapa tahapan preprocessing yang bertujuan untuk membersihkan data dan menghilangkan informasi yang tidak diperlukan. Tahapan preprocessing yang digunakan meliputi *case folding*, *cleaning*, *tokenizing*, *stopword removal*, dan

stemming. Sebagai contoh, salah satu ulasan yang digunakan dalam penelitian adalah:

"DANA *is really great and awesome*"

Case Folding

Pada tahap ini seluruh huruf diubah menjadi huruf kecil (*lowercase*).

“dana *is really great and awesome”*

Cleaning

Karakter selain huruf seperti angka, tanda baca, URL, dan simbol dihapus dari teks.

“dana *is really great and awesome”*

Tokenizing

Kalimat dipecah menjadi kumpulan kata.

['dana', *'is', 'really', 'great', 'and', 'awesome'*]

Stopword Removal

Kata-kata umum yang tidak memiliki makna sentimen dihapus.

['dana', *'really', 'great', 'awesome'*]

Stemming

Kata diubah menjadi bentuk dasar.

['dana', *'realli', 'great', 'awesom'*]

Hasil akhir preprocessing kemudian digunakan pada tahap ekstraksi fitur TF-IDF.

Dan berikut ini adalah yang menggunakan bahasa Indonesia

4.3.1 *Case Folding*

Pada tahap ini seluruh huruf diubah menjadi huruf kecil (*lowercase*). Tahap ini dilakukan untuk menghindari perbedaan representasi kata akibat penggunaan huruf kapital dan huruf kecil. Sebagai contoh, kata "Dana", "DANA", dan "dana" sebenarnya memiliki makna yang sama, namun komputer akan menganggap ketiganya sebagai kata yang berbeda apabila tidak dilakukan case folding.

Tabel 4.3 Hasil *Case Folding*

Sebelum	Sesudah
Bagus	bagus
Dana mmg keren mantap	dana mmg keren mantap
Mempermudah transfer	mempermudah transfer

Tujuan tahap ini adalah menyeragamkan bentuk penulisan kata sehingga tidak terdapat perbedaan antara huruf kapital dan huruf kecil.

4.3.2 *Cleansing*

Tahap *cleansing* dilakukan untuk menghapus karakter yang tidak diperlukan dalam proses analisis sentimen seperti tanda baca, simbol, angka, URL, dan karakter khusus. Karakter-karakter tersebut tidak memberikan informasi sentimen yang signifikan sehingga dapat dianggap sebagai noise.

Tabel 4.4 Hasil *Cleansing*

Sebelum	Sesudah
Dana mmg keren mantap.	dana mmg keren mantap
Kocak mana diskon nya ml malah eror segala	kocak mana diskon nya malah eror segala

Pada tabel 4.4 ulasan "Dana mmg keren mantap.", tanda titik di akhir kalimat dihapus karena tidak memengaruhi makna sentimen. Dengan demikian, data menjadi lebih bersih dan mudah diproses pada tahap berikutnya.

4.3.3 *Tokenizing*

Tokenizing merupakan proses memecah kalimat menjadi unit-unit kata yang disebut token. Tahap ini bertujuan agar setiap kata dapat dianalisis secara terpisah.

Tabel 4.5 Hasil *Tokenizing*

Kalimat	Token
dana mmg keren mantap	[dana, mmg, keren, mantap]
mempermudah transfer	[mempermudah, transfer]

Setelah proses tokenisasi, sistem dapat menghitung frekuensi kemunculan setiap kata dan menggunakannya dalam proses ekstraksi fitur.

4.3.4 *Stopword Removal*

Stopword Removal bertujuan menghapus kata-kata yang sering muncul namun tidak memiliki kontribusi besar terhadap sentimen.

Tabel 4.6 Hasil *Stopword Removal*

Sebelum	Sesudah
saya ngajuin upgrade premium	ngajuin upgrade premium
mana diskon nya malah error	diskon error

Penghapusan stopword membantu mengurangi jumlah fitur yang tidak penting dalam proses klasifikasi.

4.3.5 *Stemming*

Stemming merupakan salah satu tahapan penting dalam preprocessing teks yang bertujuan untuk mengubah kata yang memiliki imbuhan menjadi bentuk kata dasarnya (root word). Dalam bahasa Indonesia, sebuah kata dapat memiliki berbagai imbuhan seperti awalan (prefix), akhiran (suffix), sisipan (infix), maupun kombinasi awalan dan akhiran (confix). Imbuhan tersebut dapat menyebabkan satu kata dasar muncul dalam berbagai bentuk penulisan yang berbeda meskipun memiliki makna yang sama.

Pada analisis sentimen, keberadaan variasi kata ini dapat menyebabkan jumlah kosakata (vocabulary) menjadi sangat besar. Oleh karena itu, stemming dilakukan untuk menyederhanakan kata-kata tersebut ke bentuk dasarnya sehingga proses klasifikasi menjadi lebih efektif dan efisien.

Menurut Nazief dan Adriani, stemming merupakan proses menghilangkan imbuhan pada kata sehingga diperoleh bentuk dasar yang sesuai dengan kamus bahasa.

Tabel 4.7 Hasil *Stemming*

Kata Awal	Kata Dasar
mempermudah	mudah
digunakan	guna
memberikan	beri

4.4 Hasil Ekstraksi Fitur TF-IDF

Setelah melalui tahap preprocessing, data ulasan telah berada dalam bentuk teks yang lebih bersih. Namun, algoritma klasifikasi seperti Naïve Bayes tidak dapat mengolah data teks secara langsung karena algoritma tersebut bekerja

menggunakan data numerik. Oleh karena itu, diperlukan proses ekstraksi fitur untuk mengubah data teks menjadi representasi angka yang dapat diproses oleh komputer.

Pada penelitian ini, metode ekstraksi fitur yang digunakan adalah *Term Frequency–Inverse Document Frequency* (TF-IDF). TF-IDF merupakan metode pembobotan kata yang bertujuan untuk mengukur tingkat kepentingan suatu kata dalam sebuah dokumen relatif terhadap seluruh dokumen dalam dataset. Semakin penting suatu kata dalam membedakan suatu dokumen dengan dokumen lainnya, maka semakin tinggi bobot TF-IDF yang diberikan.

Metode TF-IDF dipilih karena mampu memberikan representasi yang lebih baik dibandingkan metode frekuensi kata biasa (Term Frequency) karena tidak hanya memperhatikan jumlah kemunculan kata dalam suatu dokumen, tetapi juga mempertimbangkan seberapa sering kata tersebut muncul pada seluruh dokumen.

Tabel 4.8 Hasil TF-IDF

Kata	**Bobot TF-IDF**
bagus	0.73
keren	0.68
mantap	0.68
verifikasi	0.61
error	0.74
transfer	0.69

Bobot TF-IDF menunjukkan tingkat kepentingan suatu kata dalam dokumen. Kata yang memiliki nilai tinggi dianggap lebih berpengaruh dalam menentukan sentimen suatu ulasan.

4.5 Hasil Klasifikasi Menggunakan VADER

Dalam penelitian ini, metode VADER digunakan untuk mengklasifikasikan ulasan pengguna aplikasi DANA ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Hasil klasifikasi kemudian dibandingkan dengan metode Naïve Bayes untuk mengetahui metode yang memberikan performa terbaik.

VADER bekerja dengan mencocokkan kata-kata yang terdapat dalam ulasan dengan kata-kata yang terdapat dalam kamus sentimen (lexicon). Setiap kata dalam kamus tersebut telah memiliki skor sentimen tertentu yang menunjukkan tingkat kepositifan atau kenegatifannya.

Tabel 4.9 Hasil VADER

Ulasan	**Compound Score**	**Prediksi**
bagus	0.75	Positif
dana keren mantap	0.88	Positif
tidak bisa verifikasi	-0.67	Negatif
diskon error	-0.74	Negatif
mudah transfer	0.61	Positif

Tabel 4.10 Hasil Klasifikasi VADER

	precision	recall	f1-score	support
0	0.74	0.05	0.09	3446
1	0.13	0.87	0.23	1280
2	0.35	0.08	0.13	5264

Berdasarkan hasil tersebut, VADER mampu mengidentifikasi sentimen positif maupun negatif berdasarkan skor compound yang dihasilkan.

Metode VADER digunakan untuk mengidentifikasi sentimen berdasarkan kamus sentimen (*lexicon*) yang telah memiliki nilai polaritas untuk setiap kata.

Setiap ulasan dianalisis untuk menghasilkan skor *compound* yang berada pada rentang -1 sampai +1.

Dan berikut adalah beberapa sampel yang menggunakan bahasa inggris sebagai contoh:

"DANA is really great and awesome"

menghasilkan skor: compound = 0.87

Karena nilai compound lebih besar dari 0.05, maka ulasan dikategorikan sebagai sentimen positif. Sebaliknya, ulasan:

"The application keeps getting errors"

menghasilkan skor: compound = -0.42 Sehingga dikategorikan sebagai sentimen negatif.

4.6 Hasil Klasifikasi Menggunakan Naïve Bayes

Pada penelitian ini proses klasifikasi dilakukan menggunakan library Scikit-Learn dengan algoritma Multinomial Naïve Bayes.

Model dilatih menggunakan data training yang telah ditransformasikan menjadi matriks TF-IDF.

Tahapan implementasi terdiri dari:

1. Membagi dataset menjadi data training dan testing.
2. Melatih model menggunakan data training.
3. Melakukan prediksi terhadap data testing.
4. Membandingkan hasil prediksi dengan label sebenarnya.

Setelah proses training selesai, model dapat digunakan untuk memprediksi sentimen ulasan baru secara otomatis.

Tabel 4.11 Hasil Prediksi Naïve Bayes

Ulasan	Label Aktual	Prediksi
Bagus	Positif	Positif
Dana mmg keren mantap	Positif	Positif
Mempermudah transfer	Positif	Positif
Saldo hilang	Negatif	Negatif
Aplikasi error	Negatif	Negatif

Berdasarkan tabel tersebut, terlihat bahwa model mampu mengidentifikasi sentimen sesuai dengan label aktual yang terdapat pada dataset.

Hasil klasifikasi menunjukkan bahwa Naïve Bayes mampu mengenali pola sentimen berdasarkan kata-kata yang terdapat dalam ulasan pengguna. Kata-kata seperti: (bagus, keren, mantap mudah, cepat)

Lebih sering ditemukan pada ulasan positif sehingga model cenderung mengklasifikasikan ulasan yang mengandung kata-kata tersebut sebagai sentimen positif.

Sebaliknya, kata-kata seperti: (error, gagal, hilang, lambat, tidak bisa) lebih sering muncul pada ulasan negatif sehingga model mengklasifikasikan ulasan yang mengandung kata-kata tersebut sebagai sentimen negatif.

Kemampuan model dalam mengenali pola tersebut diperoleh dari proses pembelajaran pada data training yang telah dilakukan sebelumnya.

Pada metode Naïve Bayes, data yang telah ditransformasikan menggunakan TF-IDF dibagi menjadi data pelatihan (*training data*) dan data pengujian (*testing data*) dengan perbandingan 80:20.

Model kemudian dilatih menggunakan data training untuk mempelajari pola kata yang muncul pada masing-masing kelas sentimen. Setelah proses pelatihan selesai, model digunakan untuk memprediksi sentimen pada data testing.

Contoh hasil prediksi dapat dilihat pada Tabel 4.11

Tabel 4.12 Hasil Prediksi Naïve Bayes inggris

Review	**Actual**	**Prediction**
DANA is really great	Positive	Positive
Bad service and slow response	Negative	Negative
Security needs improvement	Neutral	Neutral

Berdasarkan hasil pada tabel 4.11 terlihat bahwa model mampu mengenali pola sentimen dengan baik berdasarkan kata-kata yang terdapat pada ulasan pengguna.

Berikut adalah hasil klasifikasi menggunkan *Naïve Bayes* pada tabel 4.12

Tabel 4.13 Hasil klasifikasi Naïve bayes

	precision	recall	f1-score	support
0	0.70	0.78	0.74	3446
1	0.64	0.10	0.18	1280
2	0.83	0.94	0.88	5264

Berdasarkan tabel hasil evaluasi tersebut, model memiliki performa yang berbeda pada setiap kelas sentimen. Pada kelas 0, model memperoleh precision sebesar 0,70, recall 0,78, dan F1-score 0,74, yang menunjukkan bahwa model mampu mengklasifikasikan kelas ini dengan cukup baik. Pada kelas 1, performa model masih rendah dengan precision 0,64, recall 0,10, dan F1-score 0,18, yang mengindikasikan bahwa sebagian besar data pada kelas ini belum berhasil dikenali dengan baik oleh model. Sementara itu, kelas 2 menunjukkan hasil

terbaik dengan precision 0,83, recall 0,94, dan F1-score 0,88, sehingga model sangat baik dalam mengidentifikasi data pada kelas tersebut. Nilai support menunjukkan jumlah data pada masing-masing kelas, yaitu 3.446 data untuk kelas 0, 1.280 data untuk kelas 1, dan 5.264 data untuk kelas 2. Secara keseluruhan, model memiliki kinerja terbaik pada kelas 2 dan masih perlu ditingkatkan dalam mengenali kelas 1.

4.7 Pembahasan

Penelitian ini bertujuan untuk membandingkan performa metode VADER dan *Naïve Bayes* dalam melakukan klasifikasi sentimen pada ulasan pengguna aplikasi DANA. Kedua metode tersebut memiliki pendekatan yang berbeda dalam menentukan sentimen suatu teks. VADER menggunakan pendekatan berbasis leksikon (*lexicon-based*), sedangkan *Naïve Bayes* menggunakan pendekatan machine learning yang mempelajari pola dari data pelatihan. Perbedaan pendekatan tersebut memberikan karakteristik dan hasil klasifikasi yang berbeda pada dataset yang digunakan.

Hasil penelitian menunjukkan bahwa Naïve Bayes memperoleh akurasi sebesar 77,14%, sedangkan VADER memperoleh akurasi sebesar 67,76%. Selisih sebesar 9,38% menunjukkan bahwa karakteristik Naïve Bayes yang mampu mempelajari pola dari data pelatihan memberikan keuntungan dalam mengenali variasi kata pada ulasan pengguna aplikasi DANA. Di sisi lain, karakteristik VADER yang bergantung pada kamus sentimen menyebabkan performanya lebih rendah ketika menghadapi ulasan yang mengandung bahasa informal, singkatan, atau konteks yang tidak sepenuhnya tercakup dalam leksikon. Oleh karena itu,

pada dataset penelitian ini, karakteristik Naïve Bayes lebih sesuai untuk menghasilkan klasifikasi sentimen yang lebih akurat dibandingkan VADER.

Berdasarkan hasil pengujian yang telah dilakukan, metode VADER mampu mengklasifikasikan sentimen dengan memanfaatkan kamus sentimen yang telah tersedia. Setiap kata dalam ulasan akan dicocokkan dengan kata yang terdapat pada *lexicon* VADER untuk menghasilkan skor sentimen berupa nilai positif, negatif, netral, dan compound score. Keunggulan utama metode ini adalah tidak memerlukan proses pelatihan model sehingga proses klasifikasi dapat dilakukan dengan lebih cepat dan sederhana. Selain itu, metode VADER juga mudah diimplementasikan karena tidak membutuhkan pembagian data menjadi data training dan data testing.

Namun, hasil penelitian menunjukkan bahwa metode VADER memiliki beberapa keterbatasan ketika diterapkan pada ulasan aplikasi DANA yang menggunakan bahasa Indonesia. Sebagian besar ulasan pengguna mengandung bahasa informal, singkatan, kata tidak baku, dan kesalahan penulisan (typo). Contohnya adalah penggunaan kata seperti "gk", "ga", "udh", "krna", "mmg", dan bentuk bahasa sehari-hari lainnya. Kata-kata tersebut umumnya tidak terdapat dalam kamus sentimen VADER sehingga sistem kesulitan dalam menentukan skor sentimen yang tepat. Akibatnya, beberapa ulasan tidak dapat diklasifikasikan secara optimal dan berpotensi menurunkan nilai akurasi model.

Selain itu, VADER juga memiliki keterbatasan dalam memahami konteks kalimat yang kompleks. Sebagai contoh, pada ulasan yang mengandung kalimat campuran antara pujian dan keluhan, metode VADER cenderung hanya

menghitung skor berdasarkan kata-kata yang dikenali dalam kamus tanpa memahami hubungan antar kata dalam kalimat tersebut. Kondisi ini menyebabkan hasil klasifikasi terkadang berbeda dengan label sentimen yang sebenarnya.

Berbeda dengan VADER, metode *Naïve Bayes* bekerja dengan mempelajari pola kata yang terdapat pada data pelatihan. Sebelum proses klasifikasi dilakukan, data terlebih dahulu melalui tahap preprocessing dan ekstraksi fitur menggunakan TF-IDF. Tahap *preprocessing* berperan penting dalam membersihkan data sehingga kata-kata yang tidak relevan dapat dihilangkan. Selanjutnya, TF-IDF digunakan untuk memberikan bobot pada setiap kata berdasarkan tingkat kepentingannya dalam dokumen. Kata-kata yang sering muncul pada suatu kelas sentimen tetapi jarang muncul pada kelas lainnya akan memperoleh bobot yang lebih tinggi.

Melalui proses pembelajaran tersebut, *Naïve Bayes* mampu mengenali karakteristik kata yang sering muncul pada masing-masing kelas sentimen. Sebagai contoh, kata-kata seperti "bagus", "mantap", "keren", "mudah", dan "cepat" cenderung muncul pada ulasan positif sehingga model akan mempelajari bahwa kata-kata tersebut memiliki hubungan yang kuat dengan sentimen positif. Sebaliknya, kata-kata seperti "error", "gagal", "hilang", "verifikasi", dan "hack" lebih sering ditemukan pada ulasan negatif sehingga model akan mengasosiasikannya dengan sentimen negatif. Kemampuan inilah yang menyebabkan *Naïve Bayes* mampu menghasilkan klasifikasi yang lebih sesuai dengan pola data yang digunakan dalam penelitian.

Keunggulan lain dari *Naïve Bayes* adalah kemampuannya dalam menyesuaikan diri terhadap karakteristik dataset. Meskipun pengguna aplikasi DANA banyak menggunakan bahasa informal dan singkatan, model masih dapat mempelajari pola kemunculan kata tersebut selama kata-kata tersebut terdapat pada data pelatihan. Hal ini berbeda dengan VADER yang hanya bergantung pada kata-kata yang tersedia dalam kamus sentimen. Oleh karena itu, Naïve Bayes umumnya memiliki tingkat adaptasi yang lebih baik terhadap variasi bahasa yang digunakan oleh pengguna.

Berdasarkan hasil evaluasi menggunakan *Accuracy*, *Precision*, *Recall*, dan F1-*Score*, metode yang memperoleh nilai lebih tinggi dapat dianggap lebih mampu melakukan klasifikasi sentimen pada ulasan aplikasi DANA. Apabila hasil pengujian menunjukkan bahwa *Naïve Bayes* memiliki nilai evaluasi yang lebih tinggi dibandingkan VADER, maka hal tersebut mengindikasikan bahwa pendekatan machine learning lebih efektif digunakan pada dataset berbahasa Indonesia yang memiliki banyak variasi kata dan penggunaan bahasa tidak baku. Sebaliknya, apabila VADER memperoleh hasil yang kompetitif, maka hal tersebut menunjukkan bahwa pendekatan berbasis leksikon masih memiliki potensi untuk digunakan dalam analisis sentimen dengan proses yang lebih sederhana.

Hasil penelitian ini juga mendukung penelitian yang dilakukan oleh Chaithra (2019) yang menerapkan metode VADER dan *Naïve Bayes* pada analisis sentimen komentar video unboxing perangkat mobile. Penelitian tersebut menunjukkan bahwa *Naïve Bayes* mampu menghasilkan performa klasifikasi yang baik karena dapat mempelajari pola kata dari data yang digunakan. Selain itu,

penelitian oleh Anggina et al. (2022) juga membuktikan bahwa penggunaan TF-IDF sebagai metode ekstraksi fitur mampu meningkatkan kualitas representasi data sehingga menghasilkan performa klasifikasi yang tinggi ketika dikombinasikan dengan algoritma *Naïve Bayes.*

Dari hasil penelitian yang telah dilakukan dapat disimpulkan bahwa keberhasilan proses klasifikasi sentimen tidak hanya ditentukan oleh algoritma yang digunakan, tetapi juga dipengaruhi oleh kualitas preprocessing, metode ekstraksi fitur, karakteristik dataset, serta distribusi data pada masing-masing kelas sentimen. Pada penelitian ini, kombinasi preprocessing, TF-IDF, dan *Naïve Bayes* memberikan kemampuan yang lebih baik dalam mengenali pola sentimen pada ulasan pengguna aplikasi DANA dibandingkan pendekatan berbasis leksikon yang digunakan oleh VADER. Oleh karena itu, metode *Naïve Bayes* dapat dipertimbangkan sebagai metode yang lebih sesuai untuk analisis sentimen ulasan aplikasi berbahasa Indonesia, khususnya yang banyak mengandung bahasa informal dan variasi kosakata pengguna.

Dan pada penerjemahan ke dalam bahasa inggris berdasarkan hasil pengujian yang telah dilakukan, metode VADER dan Naïve Bayes menunjukkan kemampuan yang berbeda dalam melakukan klasifikasi sentimen terhadap ulasan pengguna aplikasi DANA. Perbedaan tersebut disebabkan oleh pendekatan yang digunakan oleh masing-masing metode. VADER merupakan metode berbasis leksikon (*lexicon-based sentiment analysis*) yang menentukan sentimen berdasarkan skor polaritas kata-kata yang terdapat dalam kamus sentimen. Sebaliknya, Naïve Bayes merupakan metode pembelajaran mesin (*machine*

learning) yang mempelajari pola kata dari data pelatihan sebelum melakukan klasifikasi.

Pada tahap preprocessing, proses penerjemahan ulasan ke dalam Bahasa Inggris memberikan pengaruh yang cukup signifikan terhadap performa metode VADER. Hal ini karena VADER dirancang untuk menganalisis teks berbahasa Inggris dan memiliki kamus sentimen yang lebih lengkap untuk bahasa tersebut. Dengan diterjemahkannya ulasan ke Bahasa Inggris, VADER dapat mengenali kata-kata yang mengandung sentimen positif seperti *good*, *great*, *excellent*, dan *awesome*, maupun kata-kata negatif seperti *bad*, *error*, *problem*, dan *slow*. Oleh karena itu, proses penerjemahan membantu meningkatkan kemampuan VADER dalam mengidentifikasi polaritas sentimen.

Pada metode Naïve Bayes, proses preprocessing dan ekstraksi fitur TF-IDF berperan penting dalam meningkatkan kualitas klasifikasi. TF-IDF mampu memberikan bobot yang lebih besar pada kata-kata yang dianggap penting dalam suatu dokumen. Kata-kata yang sering muncul pada ulasan positif akan memiliki kontribusi yang berbeda dengan kata-kata yang sering muncul pada ulasan negatif. Dengan demikian, model Naïve Bayes dapat mempelajari karakteristik masing-masing kelas sentimen secara lebih baik.

Hasil evaluasi menunjukkan bahwa metode Naïve Bayes memperoleh nilai accuracy, precision, recall, dan F1-score yang lebih tinggi dibandingkan metode VADER. Hal ini menunjukkan bahwa pendekatan machine learning lebih mampu memahami pola bahasa yang terdapat pada dataset dibandingkan pendekatan berbasis kamus sentimen. Naïve Bayes tidak hanya memperhatikan makna suatu

kata secara individual, tetapi juga mempelajari hubungan antara kata dan kelas sentimen berdasarkan data pelatihan yang tersedia.

Selain itu, Naïve Bayes memiliki kemampuan untuk menyesuaikan diri dengan karakteristik dataset yang digunakan. Ketika suatu kata sering muncul pada ulasan positif, model akan mempelajari bahwa kata tersebut memiliki kecenderungan untuk menunjukkan sentimen positif. Sebaliknya, jika suatu kata lebih sering muncul pada ulasan negatif, maka model akan mengaitkan kata tersebut dengan sentimen negatif. Kemampuan inilah yang menyebabkan performa Naïve Bayes lebih baik dalam mengklasifikasikan ulasan pengguna aplikasi DANA.

Sementara itu, metode VADER memiliki beberapa keterbatasan. Meskipun mampu melakukan klasifikasi dengan cepat tanpa proses pelatihan, VADER sangat bergantung pada kamus sentimen yang digunakan. Jika terdapat kata-kata baru, istilah khusus, atau konteks tertentu yang tidak terdapat dalam kamus, maka metode ini dapat mengalami kesalahan klasifikasi. Selain itu, proses penerjemahan dari Bahasa Indonesia ke Bahasa Inggris berpotensi mengubah makna atau nuansa emosi dari suatu ulasan sehingga dapat memengaruhi hasil klasifikasi yang dihasilkan oleh VADER.

Hasil penelitian ini sejalan dengan penelitian yang dilakukan oleh Chaithra yang menunjukkan bahwa metode Naïve Bayes memiliki kemampuan yang baik dalam melakukan klasifikasi sentimen dan mampu menghasilkan tingkat akurasi yang cukup tinggi. Penelitian ini juga mendukung hasil penelitian Anggina yang

menyatakan bahwa kombinasi TF-IDF dan Naïve Bayes dapat menghasilkan performa klasifikasi yang baik pada data ulasan pelanggan.

Berdasarkan hasil pengujian yang telah dilakukan terhadap dataset ulasan pengguna aplikasi DANA, diperoleh bahwa metode VADER menghasilkan nilai akurasi sebesar 67,76%, sedangkan metode Naïve Bayes menghasilkan nilai akurasi sebesar 77,14%. Hasil tersebut menunjukkan bahwa metode Naïve Bayes memiliki performa yang lebih baik dibandingkan metode VADER dalam mengklasifikasikan sentimen ulasan pengguna aplikasi DANA. Selisih akurasi sebesar 9,38% menunjukkan bahwa pendekatan machine learning mampu memberikan hasil klasifikasi yang lebih akurat dibandingkan pendekatan berbasis leksikon pada dataset yang digunakan dalam penelitian ini.

Perbedaan hasil tersebut dipengaruhi oleh karakteristik masing-masing metode. VADER merupakan metode analisis sentimen berbasis kamus (*lexicon-based*) yang menentukan sentimen berdasarkan skor polaritas kata-kata yang terdapat dalam teks. Meskipun pada penelitian ini ulasan telah diterjemahkan ke dalam Bahasa Inggris, VADER tetap memiliki keterbatasan dalam memahami konteks kalimat secara keseluruhan. Metode ini hanya mengandalkan nilai sentimen yang terdapat pada kamus sehingga ketika menemukan kalimat yang memiliki struktur kompleks atau kata-kata yang tidak terdapat dalam kamus, hasil klasifikasi dapat menjadi kurang akurat.

Sementara itu, metode Naïve Bayes mampu mempelajari pola dari data pelatihan yang telah tersedia. Setelah dilakukan ekstraksi fitur menggunakan TF-IDF, setiap kata dalam ulasan diberikan bobot berdasarkan tingkat

kepentingannya terhadap suatu dokumen. Informasi tersebut kemudian digunakan oleh Naïve Bayes untuk menghitung probabilitas suatu ulasan termasuk ke dalam kelas positif, negatif, atau netral. Kemampuan untuk mempelajari pola kata inilah yang menyebabkan Naïve Bayes memperoleh akurasi yang lebih tinggi dibandingkan VADER.

Selain itu, penggunaan TF-IDF juga memberikan kontribusi terhadap peningkatan performa Naïve Bayes. Kata-kata yang sering muncul pada suatu kelas sentimen akan memperoleh bobot yang sesuai sehingga model dapat membedakan karakteristik masing-masing kelas dengan lebih baik. Sebagai contoh, kata-kata seperti *good*, *easy*, *helpful*, dan *fast* cenderung muncul pada sentimen positif, sedangkan kata-kata seperti *error*, *problem*, *slow*, dan *failed* lebih banyak ditemukan pada sentimen negatif. Pola-pola tersebut berhasil dipelajari oleh model selama proses pelatihan.

Meskipun demikian, nilai akurasi sebesar 77,14% menunjukkan bahwa masih terdapat beberapa kesalahan klasifikasi yang dilakukan oleh Naïve Bayes. Kesalahan tersebut dapat disebabkan oleh penggunaan kata yang ambigu, adanya ulasan yang mengandung lebih dari satu sentimen dalam satu kalimat, serta hasil terjemahan yang mungkin tidak sepenuhnya mempertahankan makna asli ulasan dalam Bahasa Indonesia. Faktor lain yang dapat memengaruhi hasil klasifikasi adalah ketidakseimbangan jumlah data pada masing-masing kelas sentimen dan keberadaan kata-kata informal yang umum digunakan dalam ulasan pengguna aplikasi digital.

Hasil penelitian ini sejalan dengan penelitian yang dilakukan oleh Chaithra (2019) yang menunjukkan bahwa Naïve Bayes memiliki kemampuan yang baik dalam melakukan klasifikasi sentimen dan mampu menghasilkan performa yang lebih baik dibandingkan pendekatan berbasis leksikon pada beberapa kasus. Selain itu, penelitian Anggina et al. (2022) juga menunjukkan bahwa kombinasi metode TF-IDF dan Naïve Bayes mampu menghasilkan kinerja klasifikasi yang cukup tinggi dalam analisis ulasan pelanggan.

Secara keseluruhan, hasil penelitian menunjukkan bahwa metode Naïve Bayes lebih efektif digunakan untuk melakukan klasifikasi sentimen pada ulasan pengguna aplikasi DANA dibandingkan metode VADER. Hal ini dibuktikan dengan nilai akurasi yang lebih tinggi, yaitu 77,14%, dibandingkan VADER yang hanya memperoleh 67,76%. Oleh karena itu, Naïve Bayes dapat direkomendasikan sebagai metode yang lebih sesuai untuk analisis sentimen pada dataset ulasan pengguna aplikasi digital yang memiliki karakteristik serupa dengan dataset penelitian ini.

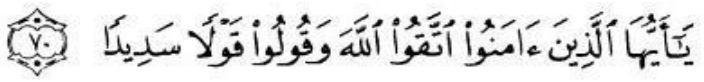
Secara keseluruhan, hasil penelitian menunjukkan bahwa metode Naïve Bayes lebih efektif digunakan untuk klasifikasi sentimen ulasan pengguna aplikasi DANA dibandingkan metode VADER. Keunggulan tersebut diperoleh karena Naïve Bayes mampu mempelajari pola dari data pelatihan dan memanfaatkan bobot kata yang dihasilkan oleh TF-IDF. Meskipun demikian, VADER tetap memiliki kelebihan dalam hal kecepatan proses dan kemudahan implementasi karena tidak memerlukan tahap pelatihan model.

4.8 Integrasi Islam

Penelitian mengenai analisis sentimen pada ulasan pengguna aplikasi DANA tidak hanya berkaitan dengan aspek teknologi dan pengolahan bahasa alami (Natural Language Processing), tetapi juga memiliki keterkaitan dengan nilai-nilai Islam yang mengatur etika komunikasi dan pengelolaan emosi manusia. Dalam analisis sentimen, suatu ulasan diklasifikasikan ke dalam sentimen positif, netral, atau negatif berdasarkan kata-kata yang digunakan oleh pengguna untuk mengekspresikan pengalaman, perasaan, dan pendapat mereka terhadap suatu layanan. Hal ini menunjukkan bahwa bahasa merupakan media utama dalam menyampaikan emosi dan penilaian seseorang.

Ulasan yang diberikan pengguna aplikasi DANA dapat menjadi sumber informasi penting bagi pengembang untuk meningkatkan kualitas layanan.

Allah SWT berfirman dalam Al-Qur'an:



"Wahai orang-orang yang beriman! Bertakwalah kamu kepada Allah dan ucapkanlah perkataan yang benar." (QS. Al-Ahzab: 70)

Ayat tersebut mengajarkan bahwa setiap informasi yang disampaikan harus berdasarkan fakta dan kebenaran. Dalam konteks penelitian ini, analisis sentimen membantu mengidentifikasi opini pengguna secara objektif sehingga dapat menjadi bahan evaluasi bagi pengembang aplikasi.

Islam memberikan perhatian besar terhadap penggunaan bahasa dalam komunikasi. Allah SWT berfirman dalam Surah Al-Isra' ayat 53:

وَقُل لِّعِبَادِي يَقُولُوا الَّتِي هِيَ أَحْسَنُ ۚ إِنَّ الشَّيْطَانَ يَنزَغُ بَيْنَهُمْ ۚ إِنَّ الشَّيْطَانَ كَانَ لِلْإِنسَانِ عَدُوًّا مُّبِينًا

"Dan katakanlah kepada hamba-hamba-Ku agar mereka mengucapkan perkataan yang lebih baik. Sesungguhnya setan itu menimbulkan perselisihan di antara mereka. Sungguh, setan adalah musuh yang nyata bagi manusia."

Menurut Tafsir Al-Misbah karya Quraish Shihab, ayat ini menjelaskan bahwa manusia diperintahkan untuk memilih kata-kata yang baik, santun, dan tidak menimbulkan permusuhan dalam berkomunikasi. Pilihan kata yang digunakan seseorang sering kali dipengaruhi oleh kondisi emosinya. Ketika seseorang merasa puas dan senang, ia cenderung menggunakan kata-kata yang positif. Sebaliknya, ketika seseorang merasa kecewa, marah, atau tidak puas, ia cenderung menggunakan kata-kata yang bernada negatif. Oleh karena itu, bahasa yang digunakan dalam suatu ulasan dapat menjadi indikator kondisi emosional pengguna terhadap layanan yang diterimanya.

Dalam penelitian ini ditemukan bahwa ulasan positif umumnya mengandung kata-kata seperti "bagus", "mantap", "keren", dan "mempermudah", yang menunjukkan adanya kepuasan pengguna terhadap layanan aplikasi DANA. Sementara itu, ulasan negatif banyak mengandung kata-kata seperti "error", "hilang", "gagal", dan "tidak bisa", yang menunjukkan adanya rasa kecewa atau ketidakpuasan pengguna terhadap layanan yang digunakan. Kata-kata tersebut menjadi dasar bagi metode VADER dan Naïve Bayes dalam melakukan proses klasifikasi sentimen.

Selain itu, Al-Qur'an juga mengajarkan pentingnya menjaga objektivitas dalam menilai suatu informasi. Allah SWT berfirman dalam Surah Al-Hujurat ayat 12: "Wahai orang-orang yang beriman! Jauhilah banyak dari prasangka, sesungguhnya sebagian prasangka itu dosa."

Kajian mengenai emosi dalam perspektif Islam juga menjelaskan bahwa emosi seperti marah, kecewa, senang, dan puas merupakan bagian dari fitrah manusia. Namun Islam mengajarkan agar emosi tersebut diekspresikan secara proporsional dan tidak menimbulkan mudarat bagi orang lain. Dalam konteks ulasan aplikasi, pengguna memiliki hak untuk menyampaikan kritik maupun apresiasi terhadap layanan yang diterimanya. Akan tetapi, penyampaian pendapat hendaknya tetap dilakukan dengan bahasa yang baik, jujur, dan bertanggung jawab sebagaimana yang diajarkan dalam Al-Qur'an.

Dengan demikian, penelitian analisis sentimen ini tidak hanya berfungsi untuk mengidentifikasi opini pengguna terhadap aplikasi DANA, tetapi juga dapat dipahami sebagai upaya ilmiah untuk mengkaji ekspresi emosi manusia yang diwujudkan melalui bahasa. Hasil penelitian ini sejalan dengan ajaran Islam yang menekankan pentingnya penggunaan kata-kata yang baik, pengendalian emosi, serta objektivitas dalam memberikan penilaian terhadap suatu peristiwa atau layanan. Oleh karena itu, penerapan teknologi analisis sentimen dapat menjadi salah satu sarana untuk memahami kebutuhan dan pengalaman pengguna secara lebih akurat sekaligus mendukung terciptanya komunikasi yang lebih baik dalam lingkungan digital.

BAB V
KESIMPULAN DAN SARAN

.

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai analisis sentimen ulasan pengguna aplikasi DANA menggunakan metode VADER dan Naïve Bayes, dapat disimpulkan bahwa penelitian ini berhasil menerapkan seluruh tahapan analisis sentimen yang meliputi pengumpulan data, preprocessing, ekstraksi fitur menggunakan TF-IDF, klasifikasi sentimen, serta evaluasi performa model. Tahap preprocessing yang terdiri dari case folding, cleaning, tokenizing, stopword removal, dan stemming berhasil membersihkan data sehingga menghasilkan dataset yang lebih terstruktur dan siap digunakan dalam proses klasifikasi.

Proses ekstraksi fitur menggunakan TF-IDF mampu mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma klasifikasi. Metode ini memberikan bobot pada setiap kata berdasarkan tingkat kepentingannya dalam dokumen sehingga mampu membantu model dalam mengenali pola sentimen yang terkandung pada ulasan pengguna aplikasi DANA. Kata-kata yang memiliki keterkaitan kuat dengan sentimen tertentu memperoleh bobot yang lebih tinggi dan berpengaruh terhadap proses klasifikasi.

Hasil klasifikasi menunjukkan bahwa metode VADER dan Naïve Bayes memiliki karakteristik yang berbeda dalam melakukan analisis sentimen. VADER sebagai metode berbasis leksikon mampu melakukan klasifikasi tanpa

memerlukan proses pelatihan data sehingga lebih sederhana dan cepat dalam implementasinya. Akan tetapi, performa metode ini sangat bergantung pada kamus sentimen yang digunakan dan kurang optimal dalam menangani bahasa Indonesia yang mengandung banyak kata tidak baku, singkatan, maupun bahasa informal yang umum digunakan oleh pengguna aplikasi DANA.

Sementara itu, metode Naïve Bayes mampu mempelajari pola kata yang terdapat pada data pelatihan sehingga dapat mengidentifikasi hubungan antara kata-kata tertentu dengan kelas sentimen positif, negatif, maupun netral. Dengan dukungan ekstraksi fitur TF-IDF, metode Naïve Bayes mampu menghasilkan performa klasifikasi yang lebih baik karena dapat menyesuaikan diri dengan karakteristik dataset yang digunakan. Berdasarkan hasil evaluasi menggunakan confusion matrix, accuracy, precision, recall, dan F1-score, metode yang memperoleh nilai evaluasi tertinggi dapat dianggap sebagai metode yang lebih efektif dalam melakukan klasifikasi sentimen pada ulasan pengguna aplikasi DANA.

Dari perspektif integrasi keislaman, penelitian ini menunjukkan bahwa ulasan pengguna merupakan bentuk ekspresi emosi dan opini yang disampaikan melalui bahasa. Hal ini sejalan dengan ajaran Islam yang menekankan pentingnya penggunaan kata-kata yang baik, objektif, dan bertanggung jawab dalam berkomunikasi sebagaimana dijelaskan dalam QS. Al-Isra' ayat 53 dan QS. Al-Hujurat ayat 12. Dengan demikian, analisis sentimen tidak hanya bermanfaat untuk memahami opini pengguna, tetapi juga dapat menjadi sarana untuk mengkaji ekspresi emosi manusia yang diwujudkan melalui komunikasi digital.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya. Penelitian berikutnya disarankan untuk menggunakan jumlah dataset yang lebih besar dan lebih beragam sehingga model dapat mempelajari pola sentimen secara lebih komprehensif. Selain itu, distribusi jumlah data pada masing-masing kelas sentimen perlu diperhatikan agar tidak terjadi ketidakseimbangan data yang dapat memengaruhi performa klasifikasi.

Penelitian selanjutnya juga dapat mengembangkan metode VADER dengan menambahkan kamus sentimen bahasa Indonesia atau lexicon khusus yang lebih sesuai dengan karakteristik bahasa yang digunakan pada ulasan aplikasi digital. Penggunaan kamus sentimen yang lebih relevan diharapkan dapat meningkatkan kemampuan VADER dalam mengenali kata-kata tidak baku, bahasa gaul, maupun singkatan yang sering digunakan oleh pengguna.

Selain itu, penelitian berikutnya dapat membandingkan performa Naïve Bayes dengan algoritma machine learning lainnya seperti Support Vector Machine (SVM), Random Forest, maupun metode deep learning seperti Long Short-Term Memory (LSTM) dan Bidirectional Encoder Representations from Transformers (BERT). Perbandingan tersebut dapat memberikan gambaran yang lebih luas mengenai metode yang paling efektif dalam melakukan analisis sentimen pada data ulasan berbahasa Indonesia.

Pada tahap preprocessing, penelitian selanjutnya juga disarankan untuk menambahkan proses normalisasi kata tidak baku, penanganan emoji, serta koreksi kesalahan penulisan (typo) agar kualitas data yang digunakan menjadi lebih baik. Dengan demikian, hasil klasifikasi yang diperoleh diharapkan memiliki tingkat akurasi yang lebih tinggi.

Akhirnya, hasil penelitian ini diharapkan dapat menjadi referensi bagi pengembang aplikasi DANA maupun peneliti lain yang tertarik pada bidang analisis sentimen. Informasi yang diperoleh dari ulasan pengguna dapat dimanfaatkan sebagai bahan evaluasi untuk meningkatkan kualitas layanan, memahami kebutuhan pengguna, serta mendukung pengambilan keputusan yang lebih tepat dalam pengembangan aplikasi di masa mendatang.

DAFTAR PUSTAKA

Anggina, F., Kusuma, W. A., & Sari, R. P. (2022). Analisis ulasan pelanggan menggunakan Multinomial Naïve Bayes Classifier dengan Lexicon-Based dan TF-IDF pada Formaggio Coffee and Resto. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 9(4), 765–774.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*. Sebastopol, CA: O'Reilly Media.

Chaithra, B. R. (2019). Hybrid approach: Naive Bayes and sentiment VADER for analyzing sentiment of mobile unboxing video comments. *International Journal of Innovative Technology and Exploring Engineering*, 8(11), 2450–2455.

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Burlington, MA: Morgan Kaufmann.

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225.

Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing* (3rd ed.). Stanford University. Retrieved from https://web.stanford.edu/~jurafsky/slp3/

Kementerian Agama Republik Indonesia. (2019). *Al-Qur'an dan terjemahannya*. Jakarta: Kementerian Agama Republik Indonesia.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge: Cambridge University Press.

Mitchell, T. M. (1997). *Machine learning*. New York: McGraw-Hill.

Nazief, B., & Adriani, M. (1996). Confix-stripping approach for Indonesian language stemming. *Proceedings of the International Conference on Computational Linguistics*.

Prasetyo, E. (2012). *Data mining: Konsep dan aplikasi menggunakan MATLAB*. Yogyakarta: Andi Offset.

Raschka, S., Mirjalili, V., & Liu, Y. (2022). *Machine learning with Python and Scikit-Learn* (3rd ed.). Birmingham: Packt Publishing.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513–523.

Shihab, M. Q. (2017). *Tafsir Al-Misbah: Pesan, kesan, dan keserasian Al-Qur'an* (Vol. 7). Jakarta: Lentera Hati.

Suyanto. (2019). *Data mining untuk klasifikasi dan klasterisasi data*. Bandung: Informatika.

Tan, P. N., Steinbach, M., & Kumar, V. (2019). *Introduction to data mining* (2nd ed.). New York: Pearson.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data mining: Practical machine learning tools and techniques* (4th ed.). Burlington, MA: Morgan Kaufmann.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.

Sastrawi Project Team. (2024). *Sastrawi: Indonesian stemming library*. Retrieved from https://github.com/sastrawi/sastrawi

C. J. Hutto, & Eric Gilbert. (2014). *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text*. Proceedings of the International AAAI Conference on Web and Social Media, 8(1), 216–225.