

Sentiment Analysis of User Reviews of the AdaKami Online Loan App from the App Store Using SVM and Naive Bayes

Wava Lativa Azzahra ¹, Jamaludin Indra ², Rahmat ³, Sutan Faisal ⁴

* Teknik Informatika, Universitas Buana Perjuangan Karawang

if21wavaazzahra@ubpkarawang.ac.id ¹, jamaludin.indra@ubpkarawang.ac.id ², rahmat@ubpkarawang.ac.id ³,
sutan.faisal@ubpkarawang.ac.id ⁴

Article Info

Article history:

Received 2025-04-29

Revised 2025-06-02

Accepted 2025-06-17

Keyword:

*AdaKami,
Apple App Store,
Naïve Bayes,
Online Loans Sentiment
Analysis,
Support Vector Machine,
TF-IDF.*

ABSTRACT

This study aims to classify sentiments on user reviews of the AdaKami online loan application, which are obtained through web scraping techniques from the Apple App Store platform. A total of 2000 reviews were collected, then selected and 1000 reviews were selected to be manually labeled by two linguistic experts, to ensure the validity of the classification. Sentiments are divided into three categories, namely negative, neutral, and positive. The classification model was built using two machine learning algorithms, namely Support Vector Machine (SVM) and Naïve Bayes (NB). The evaluation was carried out by measuring accuracy, precision, recall, F1-score, as well as through confusion matrix and cross-validation. The results showed that SVM performed better, with an accuracy of 97.5%, an F1-score of 0.97, and an average cross-validation accuracy of 84.69%. In contrast, Naïve Bayes recorded an accuracy of 81.4% and an F1-score of 0.77. The results of the paired t-test showed that the difference in performance between the two models was statistically significant ($p < 0.05$). The SVM model was then applied to predict 971 unlabeled reviews, and the results showed a dominance of negative sentiment. Wordcloud visualizations reinforced this finding, with words such as “bilih”, “bunganya”, and “teror” as the most frequently occurring words. These findings prove that SVM is more effective in classifying online loan review sentiments, as well as providing important insights for developers in understanding user perceptions and experiences.



This is an open access article under the [CC-BY-SA](#) license.

I. PENDAHULUAN

Kemajuan teknologi informasi serta komunikasi (TIK) telah mendorong terciptanya inovasi dalam layanan keuangan digital, termasuk pinjaman online. Salah satu platform yang mengalami pertumbuhan pesat adalah AdaKami, yang menyediakan pinjaman tanpa jaminan melalui aplikasi mobile serta diawasi oleh Otoritas Jasa Keuangan (OJK) [1]. Layanan ini semakin diminati setelah pandemi COVID-19 berkat kemudahan akses dan kecepatan pencairan dana, yang lebih unggul dibandingkan lembaga keuangan konvensional.

Ulasan pengguna di *Apple App Store* menjadi suatu informasi yang krusial pada menilai persepsi dan tingkat kepuasan terhadap layanan tersebut. Menurut penelitian Febriana et al [2] calon pengguna sering mengandalkan

pengalaman dari pengguna lain sebelum mengunduh aplikasi keuangan digital. Oleh sebab itu, analisis sentimen terhadap ulasan pemakai sangat krusial guna mengidentifikasi kekuatan dan kelemahan layanan, serta membantu pengambilan keputusan oleh pengembang aplikasi.

Berbagai algoritma telah diterapkan dalam klasifikasi sentimen, termasuk *Support Vector Machine (SVM)* serta *Naïve Bayes (NB)*. Studi sebelumnya mengindikasikan bahwasanya *SVM* lebih unggul pada mengklasifikasikan data teks yang lebih kompleks, sementara *NB* lebih efisien meskipun memiliki keterbatasan dalam kasus yang melibatkan ketergantungan fitur. Meskipun banyak penelitian yang telah membandingkan kedua algoritma ini, sebagian besar menggunakan data dari *Google Play*, sementara penggunaan data dari *Apple App Store* yang

memiliki karakteristik pengguna yang berbeda masih jarang dilakukan.

Selain itu, sebagian besar penelitian sebelumnya hanya mengklasifikasikan ulasan dalam dua kategori (positif-negatif) dan tidak memperhitungkan kategori netral, padahal sentimen netral juga penting untuk memahami opini pengguna. Penelitian ini juga mempertimbangkan tantangan Tantangan ketidakseimbangan data antar kelas sentimen yang dapat mempengaruhi performa model klasifikasi secara signifikan.

Menurut Wahyuni dan Turisno, pinjaman online memberikan syarat yang fleksibel dibandingi dengan lembaga keuangan tradisional seperti bank, menjadikannya solusi utama bagi masyarakat yang membutuhkan dana darurat [3]. Kemudahan akses dan kecepatan pencairan dana menjadi faktor utama yang meningkatkan popularitas layanan pinjaman online di Indonesia [4]. Data dari Kementerian Keuangan menunjukkan adanya pertumbuhan signifikan jumlah platform pinjaman *online*, dari 24 menjadi 147 dalam lima tahun terakhir [5].

AdaKami merupakan salah satu platform pinjaman online terkemuka yang diawasi oleh OJK. Aplikasi ini menawarkan pinjaman tanpa jaminan dengan proses pencairan dana yang cepat [6]. Irawan et al. menjelaskan bahwa AdaKami telah membantu banyak pengguna mendapatkan akses keuangan dengan mudah tanpa memerlukan jaminan [7].

Analisis sentimen bertujuan guna mengelompokkan opini pengguna menjadi kategori netral, negatif, maupun positif. Hidayat et al. menjelaskan bahwa pendekatan ini membantu pengembang memahami kelemahan serta kekuatan aplikasi berdasarkan umpan balik pengguna [8]. Yulianto dan Kurniawan juga menambahkan bahwa dengan analisis sentimen, pengembang dapat meningkatkan layanan secara lebih terarah [9].

Penelitian sebelumnya menunjukkan bahwa sebagian besar ulasan terhadap aplikasi pinjaman *online* cenderung negatif. Sasmitha dan Harto menggunakan *Naïve Bayes* pada aplikasi Kredivo dan menemukan dominasi sentimen negatif [10]. Setiawan et al. menggunakan *SVM* untuk lima aplikasi pinjaman *online*, termasuk AdaKami, dan hasilnya juga menunjukkan tren yang serupa [11]. Putra dan Rahmawati menemukan bahwa 68% cuitan mengenai pinjaman *online* di X mengandung sentimen negatif menggunakan algoritma *Random Forest* [12].

Berbagai penelitian lainnya juga telah membandingkan kinerja algoritma *Naïve Bayes* serta *Support Vector Machine (SVM)* pada analisis sentimen di berbagai platform. Misalnya, dalam studi yang dilaksanakan oleh Indriyani et al. [13] yang menganalisis sentimen ulasan aplikasi TikTok, ditemukan bahwasanya *SVM* menghasilkan akurasi 84%, lebih besar dibandingi *Naïve Bayes* yang hanya menyentuh 79%. Temuan oleh Rahayu et al. [14] terhadap ulasan Program Kampus Merdeka menunjukkan bahwa *SVM* dengan kernel linear dapat mencapai akurasi 93%, sedangkan *Naïve Bayes* hanya mencatatkan 86%.

Di sisi lain, penelitian yang menggunakan *Naïve Bayes* di berbagai domain, seperti analisis sentimen terhadap pembelajaran tatap muka terbatas [15], klasifikasi spam SMS [16], dan analisis sentimen terkait bakal calon presiden Indonesia 2024 [17], menunjukkan bahwa performa algoritma ini bervariasi tergantung pada karakteristik data yang dianalisis.

Temuan-temuan ini menegaskan bahwa pemilihan algoritma untuk analisis sentimen sangat dipengaruhi oleh sifat dan kompleksitas data yang digunakan. Namun, penelitian terkait ulasan aplikasi pinjaman online, khususnya dari platform *Apple App Store*, masih tergolong sedikit. Oleh karena itu, studi ini bermaksud guna menganalisis sentimen pengguna pada aplikasi AdaKami melalui ulasan di *Apple App Store*, serta membandingi dengan kinerja algoritma *SVM* serta *Naïve Bayes* dalam klasifikasi sentimen tersebut.

Namun, sebagian besar penelitian sebelumnya hanya berfokus pada satu algoritma saja, seperti *SVM* atau *Naïve Bayes*. Sedikit penelitian yang membandingkan performa kedua algoritma tersebut dalam konteks aplikasi pinjaman online tertentu, khususnya AdaKami. Selain itu, sebagian besar studi mengambil data dari platform *Google Play*, sehingga belum banyak yang mengeksplorasi data dari *App Store* yang memiliki karakteristik pengguna iOS yang berbeda.

Proses validasi manual oleh pakar dilakukan untuk menjamin bahwa pelabelan sentimen benar-benar mencerminkan makna yang tersirat dalam ulasan pengguna. Metode ini dipilih karena pelabelan otomatis berdasarkan rating sering kali bersifat bias dan tidak selalu menggambarkan konteks secara tepat. Dengan memanfaatkan data yang telah divalidasi secara manual, penelitian ini diharapkan dapat membangun model klasifikasi yang lebih akurat secara semantik dan lebih merepresentasikan opini pengguna secara nyata.

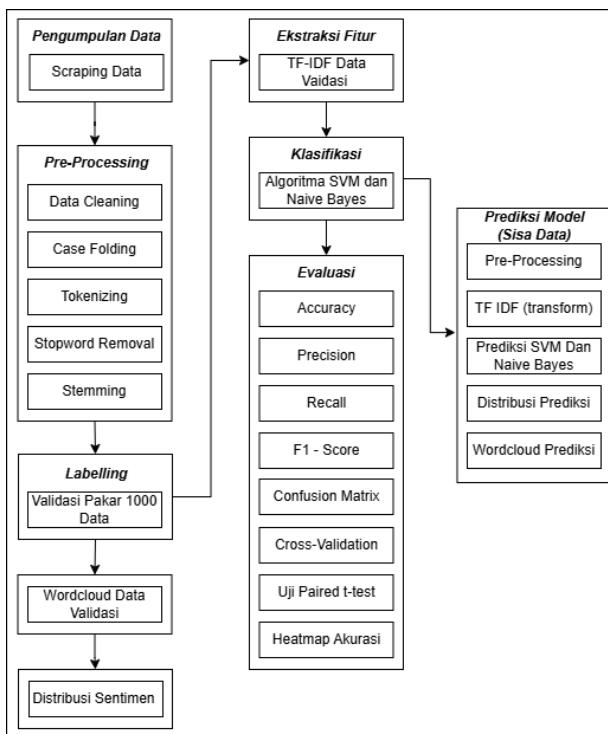
Penelitian ini bermaksud guna menganalisis sentimen pemakaian pada aplikasi AdaKami berdasarkan ulasan di *Apple App Store* menggunakan *Support Vector Machine (SVM)* serta teknik *Naïve Bayes*, serta membandingkan kinerja kedua pendekatan tersebut. Diharapkan bahwa temuan studi ini akan menyampaikan kontribusi yang konstruktif bagi pengembang aplikasi dan menjadi kontribusi dalam literatur penelitian analisis sentimen.

Studi ini mengelompokkan sentimen dalam ulasan pemakai aplikasi AdaKami ke tiga kelas, yakni negatif, netral, serta positif. Dari total 2000 ulasan yang berhasil dikumpulkan, sebanyak 1971 digunakan setelah melalui tahap pembersihan data. Dari jumlah tersebut, 1000 ulasan dipilih serta diberi label secara manual oleh pakar bahasa untuk keperluan pelatihan dan evaluasi model klasifikasi. Adapun 971 data lainnya dianalisis menggunakan model yang telah dilatih, guna mengidentifikasi kecenderungan opini pengguna secara keseluruhan. Besar harapan pendekatan ini dapat menyampaikan gambaran yang lebih utuh terkait persepsi pengguna, sekaligus melengkapi

penelitian terdahulu yang umumnya terbatas pada klasifikasi dua kelas dan tidak melibatkan proses validasi manual.

II. METODE

Penelitian ini menggunakan pendekatan analisis sentimen untuk mengelompokkan ulasan pengguna aplikasi AdaKami yang diperoleh dari Apple App Store. Rangkaian proses mencakup pengumpulan data, tahapan *pre-processing*, pelabelan manual sebagian data oleh pakar, transformasi data menjadi bentuk numerik menggunakan *TF-IDF*, pelatihan dan evaluasi model klasifikasi dengan algoritma *Support Vector Machine (SVM)* dan *Naïve Bayes (NB)*, hingga penerapan model terhadap data yang belum divalidasi. Alur lengkap proses penelitian ditampilkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Pada Gambar 1 menyajikan alur lengkap tahapan studi, dimulai dari pengumpulan 2000 ulasan, *pre-processing*, dan validasi manual terhadap 1000 data sebagai data latih dan uji. Setelah itu dilakukan visualisasi awal, ekstraksi fitur menerapkan *TF-IDF*, dan klasifikasi menerapkan algoritma *SVM* dan *Naïve Bayes*. Evaluasi model mencakup akurasi, *F1-score*, *confusion matrix*, dan uji *paired t-test*. Model terbaik kemudian digunakan untuk memprediksi 971 data sisa, yang hasilnya divisualisasikan melalui distribusi sentimen dan *wordcloud*.

A. Pengumpulan Data

Data dikumpulkan melalui teknik *web scraping* terhadap ulasan pengguna aplikasi AdaKami di Apple App Store. Sebanyak 2000 ulasan berhasil dikumpulkan dalam format CSV, kemudian dilakukan seleksi berdasarkan kategori

relevan, menggunakan bahasa Indonesia, dan tidak mengandung spam atau konten tidak relevan. Proses scraping dilakukan menggunakan bahasa pemrograman Python.

B. Pre-Processing

Tahap *pre-processing* dilakukan untuk membersihkan dan menstandarkan format data teks agar siap dianalisis. Sebanyak 2000 data ulasan yang diperoleh dari proses scraping diproses melalui lima langkah utama, yaitu *stopword removal*, *tokenizing*, *case folding*, *data cleaning*, serta *stemming*. Setelah tahap ini, tersisa 1971 data yang dianggap layak untuk dianalisis karena 29 lainnya terdeteksi sebagai duplikat, kosong, atau tidak relevan dan dihapus saat pembersihan. Untuk kebutuhan validasi manual oleh pakar, dipilih 1000 data yang hanya diproses hingga tahap *case folding* agar struktur semantik tetap terjaga dalam proses pelabelan. Data yang telah divalidasi ini digunakan dalam pelatihan dan evaluasi model klasifikasi. Sementara itu, 971 data lainnya digunakan pada tahap akhir untuk proses prediksi sentimen menggunakan model terbaik yang telah dilatih sebelumnya. Adapun tahapan *pre-processing* yang diterapkan pada studi ini dijelaskan seperti berikut:

1) Data Cleaning :

Data cleaning ialah tahapan guna memastikan kualitas data dengan menghapus duplikat, ulasan kosong, *missing values*, dan normalisasi teks seperti *case folding*, *tokenizing*, *stopword removal*, serta *stemming* agar data lebih bersih serta akurat.

2) Case Folding :

Proses penyamaan bentuk huruf dari kapital ke huruf kecil secara menyeluruh dikenal sebagai *case folding*. Langkah ini jadi penting karena dalam kenyataannya, penulisan huruf besar sering kali tidak seragam di berbagai dokumen, sehingga penyamaan ini membantu sistem pencarian bekerja lebih efisien.

3) Tokenizing :

Tokenizing adalah suatu tahapan guna memecah kalimat menjadi kata-kata untuk memberikan makna yang lebih jelas.

4) Stopword Removal :

Stopword Removal ialah proses menghapus kata-kata yang tidak penting atau tidak bermakna dalam analisis.

5) Stemming :

Stemming ialah untuk penyamaan kata yang mempunyai akhiran atau awalan kembali ke bentuk dasarnya untuk mengurangi variasi kata.

C. Labelling Data

Dari 2000 ulasan yang telah melalui *data cleaning* dan *case folding*, sebanyak 1000 data dipilih untuk divalidasi manual oleh dua pakar bahasa. Validasi dilakukan pada teks

yang belum diproses penuh agar makna tetap terjaga. Perbedaan penilaian diselesaikan melalui diskusi, dan hasil pelabelan digunakan sebagai dasar pelatihan dan pengujian model.

D. Wordcloud Data Validasi

Setelah data diberi label, dilakukan visualisasi *wordcloud* untuk masing-masing kategori sentimen, yaitu negatif, netral, serta positif. Visualisasi ini bertujuan memberikan kata yang sering muncul di setiap kelas sentimen, sehingga memberikan gambaran awal mengenai kecenderungan opini pengguna.

E. Distribusi Sentimen

Analisis distribusi sentimen dilakukan untuk mengetahui persebaran data pada tiga kategori utama, yaitu negatif, netral, serta positif. Visualisasi disajikan dalam bentuk diagram batang guna menampilkan perbandingan jumlah antar kategori secara jelas, serta diagram lingkaran untuk memperlihatkan proporsi keseluruhan sentimen. Tujuan dari analisis ini adalah untuk mendeteksi adanya ketidakseimbangan data antar kelas yang dapat berdampak pada performa model klasifikasi.

F. Ekstraksi Fitur (TF-IDF)

Langkah selanjutnya setelah menyelesaikan pre-processing adalah menggunakan metode TF-IDF untuk menentukan bobot setiap istilah tergantung pada seberapa sering istilah tersebut muncul dalam dokumen. Metrik numerik yang disebut TF-IDF dapat menampilkan kata kunci yang terkait dengan kata-kata tertentu. Lebih jauh, TF-IDF terkenal karena keakuratan, kesederhanaan, dan efisiensinya.

Rumus yang digunakan untuk menghitung bobot *TF-IDF* adalah :

$$W_{t,d} = tf_{t,d} \times idf_t = tf_{t,d} \times \log\left(\frac{N}{df_t}\right) \quad (1)$$

Keterangan:

- $W_{t,d}$: Bobot TF-IDF dari kata t pada dokumen d
- $tf_{t,d}$: Frekuensi kemunculan kata t dalam dokumen d
- idf_t : Merupakan nilai inverse dokumen frekuensi dari kata t
- df_t : Banyaknya dokumen yang memuat kata t
- N : Total jumlah dokumen yang ada dalam kumpulan data

G. Klasifikasi SVM dan Naive Bayes

Dalam tahapan ini, analisis dilakukan dengan menerapkan metode *Naïve Bayes* serta *SVM* guna mengklasifikasikan sentimen ulasan aplikasi AdaKami di Apple App Store. Tujuan utama dari tahap ini ialah guna memisahkan label sentimen positif serta negatif dengan tingkat akurasi yang

terbaik [14]. *Support Vector Machine (SVM)* adalah algoritma yang efisien serta efektif guna klasifikasi data, yang bekerja dengan membangun garis pemisah (*hyperplane*) guna membedakan kelas sentimen negatif serta positif [18].

Persamaan rumus *SVM* dapat dijelaskan seperti berikut :

$$(w \times xi) + b = 0 \quad (2)$$

Apabila data xi termasuk dalam kelas -1 maka rumusnya adalah :

$$(w \times xi + b) \leq 1, yi = -1 \quad (3)$$

Sementara itu, jika data xi termasuk dalam kelas +1, rumusnya menjadi :

$$(w \times xi + b) \geq 1, yi = +1 \quad (4)$$

Suatu teknik yang diterapkan pada data *mining* dengan pendekatan supervised learning adalah *Naïve Bayes*. Berdasarkan teorema Bayes, pendekatan ini dapat mengkategorikan data seperti halnya neural network dan decision tree [19]. Selain itu, khususnya dalam analisis sentimen, pendekatan *Naïve Bayes* mempunyai durasi pemrosesan klasifikasi yang relatif cepat [18].

Teorema *Naïve Bayes* umumnya memiliki rumus seperti berikut :

$$P(C | X) = \frac{P(X | C) \cdot P(C)}{P(X)} \quad (5)$$

Keterangan :

- $P(C | X)$: Probabilitas kelas C untuk data X (posterior)
- $P(X | C)$: Probabilitas data X muncul dari kelas C (likelihood).
- $P(C)$: Probabilitas awal kelas C (prior).
- $P(X)$: Probabilitas data X (evidence)

H. Evaluasi

Model dievaluasi memakai metrik performa seperti *recall*, *precision*, *accuracy*, serta *F1-score*, serta dianalisis melalui *confusion matrix* untuk melihat distribusi prediksi. Untuk menguji konsistensi kinerja model, diterapkan teknik *5-fold cross-validation*. Perbandingan hasil antara algoritma *SVM* dan *Naïve Bayes* dianalisis menggunakan uji *paired t-test*. Seluruh hasil evaluasi divisualisasikan dalam bentuk *heatmap* guna memberikan gambaran perbedaan kinerja masing-masing model secara lebih jelas.

I) Accuracy :

Mengukur ketepatan model dalam meramal hasil dilakukan dengan cara membandingkan berapa banyak tebakan yang sesuai dengan kenyataan terhadap keseluruhan jumlah data yang tersedia.

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+TN} \quad (6)$$

2) Precision :

Menghitung proporsi prediksi positif yang akurat (*True Positives*) dibandingkan dengan seluruh prediksi positif yang dihasilkan oleh model.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (7)$$

3) Recall :

Menghitung kemampuan model untuk mengidentifikasi seluruh data positif yang sesungguhnya (*True Positives*) dalam dataset.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (8)$$

4) F1 – Score :

Adalah rata-rata harmonis yang menggabungkan nilai *precision* serta *recall* secara seimbang. Pengukuran ini berguna untuk mengatasi ketimpangan distribusi kelas dalam dataset, dengan memberikan perhatian yang setara pada kedua indikator tersebut.

$$\text{F1 – Score} = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

5) Confusion Matrix :

Confusion matrix ialah tabel yang dipakai guna menilai kinerja model klasifikasi dengan membandingkan hasil prediksi dengan data sebenarnya, di mana sumbu vertikal menggambarkan label aktual dan sumbu horizontal menggambarkan label prediksi.

6) Cross-Validation :

Untuk menguji konsistensi kinerja model, digunakan teknik *k-fold cross-validation* dengan perolehan *k* senilai 5. Dalam metode ini, data dipisahkan menjadi lima bagian yang digunakan secara bergiliran sebagai data uji serta data latih. Tahapan pengujian serta pelatihan dilaksanakan sebanyak lima kali, dan hasil evaluasi diperoleh dari rata-rata performa seluruh pengujian tersebut.

7) Uji Paired t-test :

Selain metrik performa dasar, dilakukan pula pengujian paired t-test guna mengetahui apakah terdapat perbedaan performa yang signifikan secara statistik antara dua algoritma klasifikasi yang digunakan (*SVM* dan *Naïve Bayes*). Uji ini memperkuat validitas pemilihan model terbaik berdasarkan perbedaan hasil evaluasi secara matematis.

8) Heatmap Akurasi :

Sebagai tambahan, hasil evaluasi divisualisasikan menggunakan *heatmap* untuk membandingkan performa antar model berdasarkan nilai akurasi. Visualisasi ini membantu memperjelas perbedaan kinerja secara visual dan memudahkan interpretasinya.

I. Prediksi Model (Sisa Data)

Setelah proses pelatihan dan evaluasi model selesai, algoritma dengan performa terbaik—berdasarkan hasil pengukuran metrik dan uji statistik—digunakan untuk

melakukan prediksi sentimen pada 971 ulasan yang tidak melalui validasi pakar. Data ini sebelumnya telah menjalani tahapan *pre-processing* yang sama seperti data validasi, mencakup *stemming*, *stopword removal*, *tokenizing*, *case folding*, serta *data cleaning*.

Karena data ini tidak digunakan dalam proses pelatihan, konversi teks menjadi representasi numerik dilakukan memakai teknik *Term Frequency-Inverse Document Frequency (TF-IDF)* dengan fungsi *transform()*, bukan *fit_transform()*, agar tidak terjadi *data leakage*. Proses ini memastikan bahwa model hanya menerapkan bobot kata dari hasil pelatihan sebelumnya.

Selanjutnya, model klasifikasi (*SVM* dan *Naïve Bayes*) diterapkan guna mengelompokkan ulasan ke dalam kategori positif, netral, atau negatif. Hasil prediksi ini tidak hanya digunakan untuk mengetahui distribusi opini pada data tanpa label, tetapi juga menjadi dasar dalam tahap visualisasi dan analisis kecenderungan opini pengguna secara umum.

J. Visualisasi Hasil Prediksi

Setelah klasifikasi dilakukan terhadap 971 data yang tidak divalidasi, hasil prediksi divisualisasikan untuk memberikan gambaran umum kecenderungan opini pengguna. Visualisasi terdiri dari dua bentuk:

1) Distribusi Sentimen :

Menampilkan dalam bentuk diagram batang, yang menunjukkan jumlah prediksi untuk masing-masing kategori sentimen (positif, netral, negatif).

2) Wordcloud Prediksi :

Menampilkan kata-kata dominan dari tiap kategori sentimen untuk menggambarkan tema umum dalam hasil klasifikasi.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Studi ini memperoleh data melalui metode *web scraping* terhadap 2000 ulasan pengguna aplikasi AdaKami, yang secara khusus diambil dari Apple App Store menggunakan bahasa pemrograman Python. Setiap ulasan memuat tiga informasi utama, yaitu tanggal ulasan, isi review, dan rating bintang dari pengguna. Seluruh data kemudian disimpan dalam format CSV guna kebutuhan tahapan analisis sentimen lebih lanjut. Data hasil *scraping* ditampilkan secara ringkas pada Tabel I.

Langkah awal yang dilakukan adalah menyeleksi data dengan cara menghapus ulasan yang tidak ditulis dalam Bahasa Indonesia, kosong, duplikat, atau mengandung spam serta konten yang tidak relevan. Proses ini dilaksanakan secara otomatis memakai fungsi pembersihan berbasis Python. Setelah tahapan ini, sebanyak 1971 data yang dinyatakan valid untuk dianalisis.

Dari total data tersebut, 1000 ulasan dipilih secara acak dan dilakukan proses pelabelan sentimen secara manual oleh dua orang pakar linguistik guna memastikan akurasi dan validitas klasifikasi. Dataset hasil validasi ini kemudian dipakai menjadi data latih serta data uji dalam proses pelatihan model. Adapun 971 data sisanya digunakan untuk pengujian akhir, dengan sentimen diprediksi menggunakan model terbaik yang telah dilatih sebelumnya.

TABEL I.
HASIL SCRAPING DATA

	Date	review	Rating
0	2020-08-31 01:32:47	Udah bunga gede, belom jatuh tempo udah dikata...	1
1	2020-02-07 07:52:55	Jatuh tempo tgl 26 belum ada telat sama sekali...	1
....
1978	2023-04-15 00:10:03	aplikasi gak bermutu banget dah,, sukur udh tu...	1
1979	2023-04-14 17:49:41	aplikasi ga jelas	1
....

Tabel I menampilkan hasil *web scraping* dari ulasan pengguna aplikasi AdaKami di App Store, yang terdiri dari tiga kolom utama: date, review, dan rating. Kolom date merekam waktu ketika ulasan diberikan, review memuat teks yang berisi pengalaman pengguna, dan rating merepresentasikan skor yang diberikan dalam rentang 1 hingga 5. Berdasarkan sampel data yang ditampilkan, sebagian besar ulasan memiliki rating 1, yang mencerminkan sentimen negatif terhadap aplikasi.

B. Pre-processing

Pre-processing merupakan tahapan penting untuk membersihkan dan menstandarkan format teks sebelum dilakukan ekstraksi fitur dan klasifikasi sentimen. Proses ini dilakukan untuk mengurangi *noise*, menyederhanakan struktur kalimat, serta meningkatkan efisiensi dan akurasi model pembelajaran mesin.

Seluruh proses pre-processing dilakukan menggunakan bahasa pemrograman Python dengan bantuan pustaka seperti *Sastrawi* dan *NLTK*. *Pre-processing* diterapkan secara penuh pada seluruh data, kecuali untuk 1000 data validasi manual, yang hanya diproses sampai tahap *case folding* agar tidak mengubah struktur semantik saat proses pelabelan manual.

Tahapan *pre-processing* yang diterapkan terdiri dari lima langkah utama seperti berikut:

1) Data Cleaning :

Pada Tabel II menunjukkan hasil data *cleaning*, di mana karakter khusus seperti tanda titik, tanda hubung, dan angka telah dihapus. Selain itu, beberapa kata yang mengalami pemisahan tidak alami, seperti "seumur-umur" telah

disesuaikan agar lebih seragam. Proses ini memastikan teks lebih bersih dan siap untuk tahap *pre-processing* selanjutnya.

TABEL II
HASIL DATA CLEANING

Review	Data Cleaning
sudah tidak lagi pake aplikasi ini.. barusan seumur-umur saya pake aplikasi pinjaman online yang kayak gini ...	sudah tidak lagi pake aplikasi ini barusan seumurumur saya pake aplikasi pinjaman online yang kayak gini

2) Case Folding :

Pada Tabel III di atas menampilkan hasil *case folding*, di mana semua huruf diubah menjadi kecil guna menyamakan format teks dan mencegah perbedaan makna akibat penggunaan huruf kapital.

TABEL III
HASIL CASE FOLDING

Review	Case Folding
Ini aplikasi baru aja daftar udah banyak telfon dan SMS ga jelas, padahal baru daftar doang loh ga melakukan peminjaman ...	ini aplikasi baru aja daftar udah banyak telfon dan sms ga jelas, padahal baru daftar doang loh ga melakukan doang loh ga melakukan peminjaman ...

3) Tokenizing :

Pada Tabel IV di bawah yaitu pemecahan kalimat keluhan menjadi unit-unit kata ("aplikasi", "ngancem", "data", dsb.) yang memungkinkan model untuk mengidentifikasi elemen penting seperti ancaman atau keluhan, mendukung klasifikasi sentimen, ekstraksi topik, dan penanganan keluhan.

TABEL IV
HASIL TOKENIZING

Review	Tokenizing
Aplikasi apa ini bgni penagihannya kurang sopan ngancem ngancem sebar data	['aplikasi', 'apa', 'ini', 'bgni', 'penagihannya', 'kurang', 'sopan', 'ngancem', 'ngancem', 'sebar', 'data',]

4) Stopword Removal :

Tabel V menunjukkan hasil *stopword removal*, di mana kata-kata umum seperti 'saya', 'disini', dan 'terus' dihapus dari teks review. Setelah penghapusan, hanya kata-kata yang lebih bermakna, seperti 'pinjem', 'bayar', dan 'telat', yang dipertahankan.

TABEL V
HASIL STOPWORD REMOVAL

Review	Stopword Removal
[‘jangon’, ‘pinjem’, ‘disini’, ‘saya’, ‘selalu’, ‘bayar’, ‘tepat’, ‘waktu’, ‘gajihan’ ...]	[‘pinjem’, ‘bayar’, ‘gara’, ‘gara’, ‘gajihan’, ‘mundurin’, ‘telat’, ...]

5) Stemming :

Tabel VI di bawah ini menunjukkan hasil *stemming* yang diterapkan pada teks review. Proses *stemming* mengubah kata-kata menjadi bentuk dasarnya, seperti mengubah 'katain' menjadi 'kat', serta 'TAGIHAN' menjadi 'tagih',

TABEL VI HASIL STEMMING

Review	Stemming
['kecewa', 'telat', 'aja', 'udah', 'katain', 'maling', 'duit', 'perusahaan', 'udah', 'kena', 'biaya', 'tagihan', 'yg', 'nguras', 'duit', 'nasabah' ...]	kecewa telat aja udah katain maling duit usaha udah kena biaya tagih yg nguras duit nasabah aplikasi tim tagih kaya ...

C. Labelling Data

Proses pelabelan dilakukan untuk mengklasifikasikan setiap ulasan pengguna ke dalam salah satu dari tiga kategori sentimen, yaitu negatif, netral, serta positif. Dari total 1971 data yang telah melalui tahapan *pre-processing*, sebanyak 1000 ulasan dipilih secara acak dan diberi label sentimen secara manual oleh dua pakar linguistik. Data ini disimpan dalam variabel df_validasi pada kode program.

Pelabelan dilakukan terhadap data yang telah melalui proses *data cleaning* dan *case folding*, namun belum melalui tahap *tokenizing*, *stopword removal* dan *stemming* agar struktur makna kalimat tetap terjaga selama proses penilaian. Kedua pakar melakukan pelabelan secara terpisah, kemudian hasilnya direkonsiliasi melalui diskusi jika terdapat perbedaan penilaian. Dataset hasil pelabelan ini digunakan sebagai data latih dan uji awal dalam pelatihan model klasifikasi sentimen.

Hasil pelabelan manual dari beberapa ulasan yang mewakili bagian awal dan akhir data ditampilkan pada Tabel VII berikut.

TABEL VII
HASIL *LABELLING* DATA

	TEXT	SENTIMENT
1	udah bunga gede belom jatuh tempo udah dikatain ga mao bayar ...	Negative
739	selalu bayar tepat waktu kok tiba scor saya tidak memenuhi	Negative
....
989	okepersis seperti iklannya	Positive
994	solusi cepat untuk keperluan tepat	Netral
1000	apk apa ini belom jatuh tempo aj sudah ngancam teror terus parah ini	Negative

Setelah proses pelabelan selesai dilakukan, jumlah data pada setiap kategori sentimen bisa diamati pada Tabel VIII berikut :

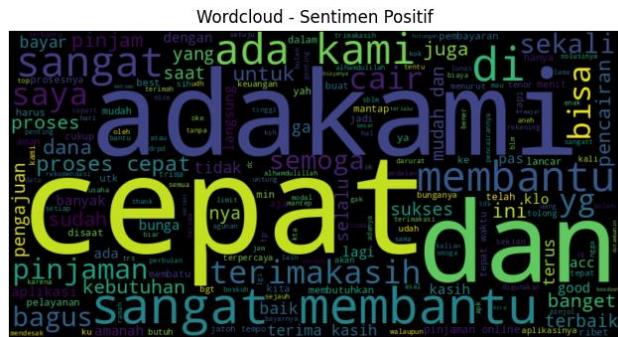
TABEL VIII
JUMLAH DATA PERLABEL

Kelas/Label	Jumlah
Positive	171
Netral	120
Negative	709

D. Wordcloud Data Validasi

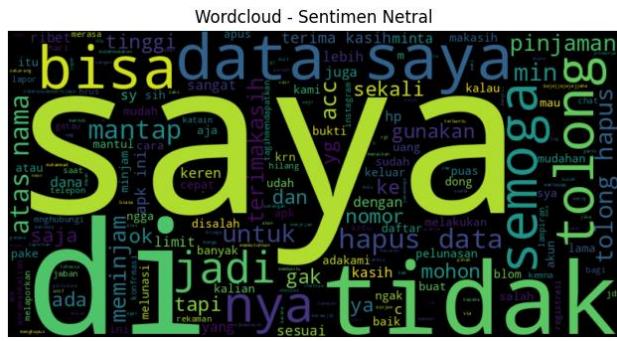
Setelah proses pelabelan selesai dilakukan, dilakukan visualisasi kata menggunakan wordcloud untuk masing-masing kategori sentimen. Tujuan dari visualisasi ini adalah guna menunjukkan kata yang sering muncul dalam setiap kelas sentimen berdasarkan data hasil validasi manual.

Wordcloud berguna untuk menggambarkan pola umum dalam ulasan pengguna, sekaligus mengidentifikasi kata-kata utama yang mencerminkan sentimen mereka. Dalam visualisasi, ukuran setiap kata merepresentasikan frekuensi kemunculannya semakin besar ukuran sebuah kata, semakin tinggi tingkat kemunculannya dalam ulasan yang termasuk dalam kategori tersebut. Hal ini memungkinkan pembaca untuk dengan cepat mengidentifikasi kata-kata yang paling dominan atau sering digunakan dalam kumpulan data yang dianalisis.



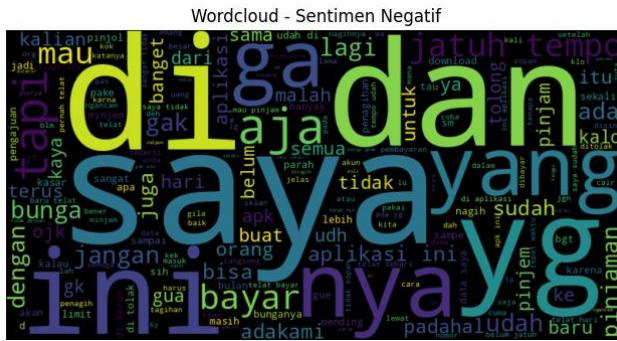
Gambar 2 Wordcloud Sentimen Positif

Gambar 2 Wordcloud ini menunjukkan bahwa kata-kata seperti “cepat”, “adakami”, “mudah”, “cair”, “membantu”, dan “proses” mendominasi ulasan dengan sentimen positif. Hal ini mencerminkan bahwa pengguna merasa puas terhadap kemudahan proses pencairan dana dan kecepatan layanan yang diberikan oleh aplikasi.



Gambar 3 Wordcloud Sentimen Netral

Gambar 3 kategori netral, kata yang paling menonjol adalah “saya”, “data”, “tidak”, “jadi”, dan “tolong”. Kata-kata ini umumnya bersifat deskriptif, tidak mengandung emosi yang kuat, dan sering digunakan dalam konteks pertanyaan, permintaan bantuan, atau penjelasan singkat, sehingga mencerminkan opini yang bersifat netral atau ambivalen.



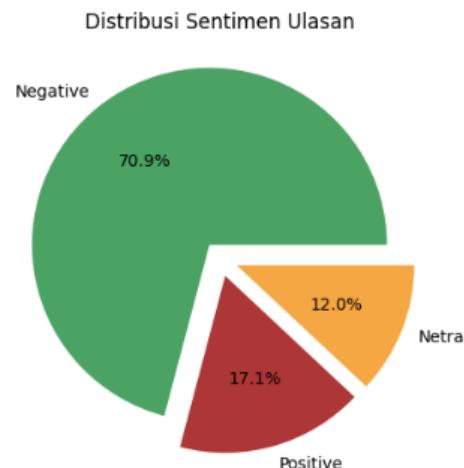
Gambar 4 Wordcloud Sentimen Negatif

Gambar 4 untuk ulasan negatif memperlihatkan kata-kata yang dominan seperti “saya”, “ga”, “bayar”, “yg”, “jangan”, “jatuh tempo”, dan “tagih”. Ini mengindikasikan adanya keluhan terhadap sistem penagihan, keterlambatan, bunga tinggi, dan interaksi pengguna yang buruk terhadap fitur atau pihak penagih dari aplikasi.

E. Distribusi Sentimen Ulasan

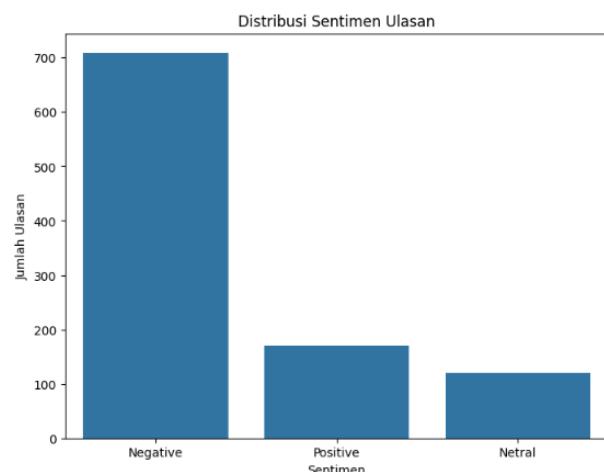
Setelah seluruh data ulasan yang telah melewati tahapan *pre-processing* selesai diberi label, dilakukan analisis distribusi sentimen untuk mengidentifikasi kecenderungan opini pengguna terhadap aplikasi AdaKami.

Analisis ini bertujuan untuk memberikan pemahaman mengenai distribusi data dalam setiap kategori sentimen. Informasi ini penting untuk memastikan proporsi data yang seimbang atau mengetahui adanya ketimpangan, yang selanjutnya akan berpengaruh pada proses pelatihan dan evaluasi model klasifikasi agar hasilnya lebih akurat dan representatif.



Gambar 5 Distribusi Ulasan per Sentimen

Gambar 5, visualisasi ini menampilkan jumlah ulasan yang termasuk dalam masing-masing kategori sentimen. Dapat diamati bahwa kelas negatif memiliki jumlah ulasan terbanyak, melebihi 700 data, diikuti oleh kategori positif sekitar 170 ulasan, dan kategori netral sekitar 120 ulasan.



Gambar 6 Distribusi Presentase Sentimen

Gambar 6 menggambarkan proporsi masing-masing kategori sentimen terhadap keseluruhan data ulasan. Kategori sentimen negatif mendominasi dengan nilai sebesar 70,9%, diikuti oleh positif sebesar 17,1%, dan netral hanya 12,0%. Visualisasi ini memperkuat temuan sebelumnya bahwa opini pengguna terhadap aplikasi cenderung bernada negatif, sementara ulasan netral jumlahnya sangat terbatas. Ketimpangan proporsi ini menegaskan pentingnya penerapan strategi khusus untuk menangani data tidak seimbang dalam proses pelatihan model klasifikasi.

F. Ekstraksi Fitur (TF-IDF)

Untuk mengonversi data ulasan menjadikan bentuk numerik yang dapat diproses oleh algoritma klasifikasi, digunakan pendekatan *Term Frequency-Inverse Document Frequency (TF-IDF)*. Metode ini memberikan bobot pada setiap kata sebagaimana tingkat kepentingannya dalam suatu ulasan dibandingkan dengan keseluruhan dokumen dalam kumpulan data.

Proses pembobotan dilakukan menggunakan fungsi `TfidfVectorizer()` dari pustaka scikit-learn, yang diterapkan pada 1000 data hasil validasi manual. Hasil transformasi berupa matriks *TF-IDF* berukuran (1000, 3619), dengan 18.030 nilai yang tidak bernilai nol, dan disimpan dalam format *Compressed Sparse Row (CSR)*. Representasi ini sangat efisien untuk pengolahan data teks, karena mayoritas sel dalam matriks berisi nol.

TABEL IX
HASIL TF-IDF

TF-IDF Matrix Shape	(1000, 3619)
<Compressed Sparse Row sparse matrix of dtype 'float64'>	
with 18030 stored elements and shape (1000, 3619)>	
Coords	Values
(0, 3475)	0.1510
(0, 599)	0.1556
(0, 1160)	0.1000
(0, 1410)	0.0803
(0, 2467)	0.3018
(0, 897)	0.1661
.....
(999, 3166)	0.2110
(999, 57)	0.2250

Pada Tabel IX menunjukkan hasil transformasi data teks menjadi bobot *TF-IDF*. Setiap baris merepresentasikan satu dokumen atau ulasan pengguna, sementara setiap kolom menunjukkan kata unik yang teridentifikasi dalam dataset. Angka-angka dalam tabel mencerminkan bobot sebuah kata pada dokumen tertentu, di mana nilai bobot yang lebih tinggi mengindikasikan bahwasanya kata tersebut mempunyai peran yang lebih signifikan dalam membentuk makna ulasan tersebut.

G. Klasifikasi

Setelah data ulasan pengguna dikonversi menjadi bentuk numerik melalui metode *Term Frequency-Inverse Document Frequency (TF-IDF)*, tahap klasifikasi dilakukan dengan menerapkan dua algoritma pembelajaran mesin, yaitu *Support Vector Machine (SVM)* dan *Naïve Bayes (NB)*. Pemilihan kedua algoritma ini didasarkan pada efektivitasnya dalam menyelesaikan tugas analisis sentimen serta performanya yang baik dalam menangani data teks, sebagaimana telah diuraikan pada Bab II.

1) Algoritma *Support Vector Machine* :

Support Vector Machine (SVM) adalah algoritma klasifikasi yang bertujuan menemukan *hyperplane* terbaik

yang mampu memisahkan data ke dalam kelas-kelas yang berbeda secara optimal. Dalam konteks penelitian ini, SVM digunakan untuk mengklasifikasikan data ke dalam tiga kategori sentimen, yaitu positif, netral, serta negatif. Algoritma ini bekerja dengan memaksimalkan margin antar kelas, sehingga dapat meningkatkan akurasi dalam membedakan sentimen berdasarkan karakteristik data yang dianalisis. Proses pelatihan dilakukan terhadap 1000 data yang telah divalidasi oleh pakar, dengan fitur masukan berupa vektor hasil transformasi *TF-IDF*. Model ini dibangun menggunakan pustaka scikit-learn dan dievaluasi dengan teknik *5-fold cross-validation* untuk meningkatkan keandalan dan menghindari overfitting.

2) Algoritma Naïve Bayes :

Naïve Bayes ialah algoritma klasifikasi probabilistik yang menerapkan Teorema *Bayes* dengan asumsi bahwasanya setiap fitur saling independen. Walaupun sederhana, algoritma ini memiliki efisiensi tinggi dan banyak digunakan dalam klasifikasi teks, termasuk pada tugas analisis sentimen. Pada penelitian ini, model *NB* dilatih menggunakan dataset yang sama seperti *SVM*, yaitu 1000 data hasil validasi pakar, serta dievaluasi dengan pendekatan dan skema validasi yang serupa.

H. Evaluasi

Evaluasi dalam studi ini bermaksud guna mengukur performa model *SVM* dan *Naïve Bayes* dalam mengklasifikasikan ulasan pengguna ke dalam tiga kategori sentimen: negatif, netral, serta positif. Evaluasi dilakukan secara bertahap, mulai dari pengujian awal terhadap data validasi manual hingga tahapan lanjutan untuk menguji akurasi, kestabilan, dan konsistensi model terhadap data yang tidak seimbang.

1) Evaluasi Awal :

Tahapan awal evaluasi dilakukan untuk menilai performa dasar model terhadap 1000 data ulasan yang sudah dilabeli secara manual oleh pakar. Pengukuran kinerja menggunakan metrik *accuracy*, *precision*, *recall*, serta *F1-score*, yang dihitung sebagaimana dari rata-rata berbobot (*weighted average*). Evaluasi ini dilakukan sebelum model melanjutkan ke tahap pengujian lanjutan yang lebih mendalam.

TABEL X
HASIL EVALUASI WAL

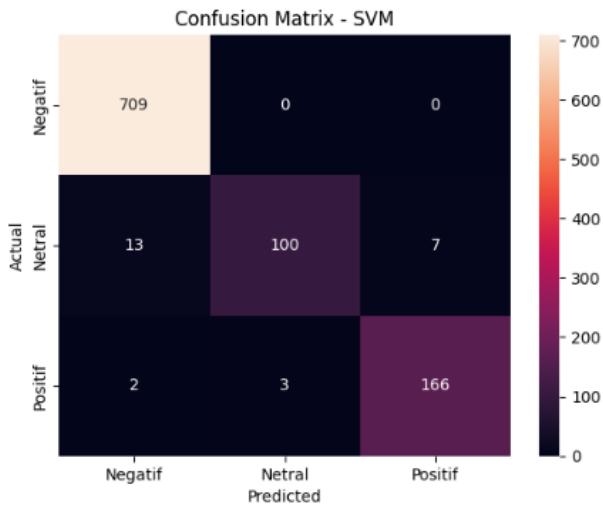
Model	Accuracy	Precision	Recall	F1-Score
SVM	0.975	0.97	0.97	0.97
Naïve Bayes	0.814	0.85	0.81	0.77

Tabel menunjukkan bahwa model *SVM* unggul dibandingkan *Naïve Bayes*, dengan akurasi tertinggi senilai 97,5%, serta *precision*, *recall*, serta *F1-score* yang seimbang

(0.97). Sebaliknya, *Naïve Bayes* hanya mencapai akurasi 81,4%, dengan *F1-score* lebih rendah (0.77), yang menunjukkan kinerjanya kurang optimal, terutama dalam mengenali kelas dengan jumlah data yang lebih sedikit.

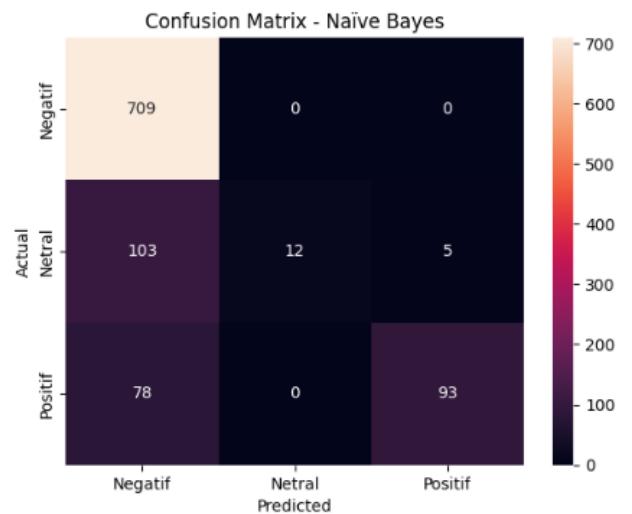
2) Confusion Matrix :

Confusion matrix dipakai guna mengevaluasi kesalahan klasifikasi model terhadap masing-masing kelas sentimen. Melalui visualisasi ini, dapat diketahui seberapa sering model menghasilkan prediksi yang benar maupun salah pada setiap kategori. Analisis terhadap *confusion matrix* juga membantu mengidentifikasi potensi bias model, khususnya pada kelas yang memiliki jumlah data lebih sedikit seperti sentimen netral. Hasil *confusion matrix* dari kedua model tersebut yaitu sebagai berikut.



Gambar 7 Confusion Matrix SVM

Gambar 7 menunjukkan *confusion matrix* dari model *Support Vector Machine* (SVM). Model ini mampu mengklasifikasikan seluruh data pada kelas negatif dengan tepat (709 data), dan menunjukkan performa yang cukup baik pada dua kelas lainnya. Untuk kelas netral, sebanyak 100 dari 120 data berhasil diprediksi dengan benar, sementara sisanya terkласifikasi sebagai positif atau negatif. Sementara itu, pada kelas positif, sebanyak 166 dari 171 data berhasil dikenali dengan tepat. Perolehan ini memperlihatkan bahwasanya *SVM* mempunyai kinerja yang lebih stabil dan seimbang dalam mengklasifikasikan seluruh kategori sentimen, terutama dalam kondisi data yang tidak seimbang, sehingga mendukung hasil evaluasi metrik sebelumnya.



Gambar 8 Confusion Matrix Naïve Bayes

Gambar 8 mengindikasikan bahwasanya model *Naïve Bayes* mampu mengklasifikasikan seluruh data sentimen negatif (709 data) dengan akurat. Namun, performanya menurun pada kelas netral dan positif. Dari total 120 data netral, hanya 12 yang diprediksi dengan benar, sementara 103 data salah diklasifikasikan sebagai negatif. Begitu pula untuk kelas positif, hanya 93 dari 171 data yang berhasil dikenali dengan tepat, dan 78 data lainnya juga keliru masuk ke kelas negatif. Hasil ini mengindikasikan bahwa model cenderung bias terhadap kategori negatif, dan mengalami kesulitan dalam membedakan kelas lain, khususnya kelas netral yang jumlahnya lebih sedikit.

3) Cross-Validation :

Setelah evaluasi awal dan analisis *confusion matrix*, dilakukan pengujian 5-fold cross-validation untuk mengukur konsistensi dan kemampuan generalisasi model. Sebanyak 1000 data validasi manual dibagi menjadi lima bagian, di mana setiap bagian secara bergantian digunakan sebagai data uji. Proses ini diulang lima kali agar seluruh data mendapatkan kesempatan yang sama sebagai data uji.

Hasil pengujian akurasi untuk masing-masing fold ditampilkan pada Tabel XI :

TABEL XI
HASIL CROSS-VALIDATION

Fold	SVM	Naive Bayes
Fold 1	0.845	0.77
Fold 2	0.87	0.755
Fold 3	0.815	0.75
Fold 4	0.84	0.765
Fold 5	0.865	0.77
Rata - Rata	0.8469	0.762

Berdasarkan Tabel XI, dapat dilihat bahwa model *SVM* menunjukkan performa yang lebih unggul dan konsisten dibandingkan *Naïve Bayes*. *SVM* mencatatkan rata-rata akurasi sebesar 84,69%, sedangkan *Naïve Bayes* hanya memperoleh 76,20%. Selain itu, performa *SVM* pada masing-masing fold relatif stabil, yang mengindikasikan bahwasanya model ini mempunyai daya generalisasi yang lebih baik saat dihadapkan pada variasi data.

Implementasi *cross-validation* ini turut memperkuat keandalan model karena seluruh dataset diuji secara menyeluruh. Oleh karena itu, dapat disimpulkan bahwa *SVM* tidak hanya lebih akurat, tetapi juga lebih stabil dan tahan terhadap risiko overfitting dibandingkan model *Naïve Bayes*.

4) Uji Paired *t-test* :

Untuk menguji apakah perbedaan performa antara model *SVM* dan *Naïve Bayes* signifikan secara statistik, dilakukan *paired t-test* terhadap nilai akurasi masing-masing model yang diperoleh dari *5-fold cross-validation*.

Hasil uji statistik tersebut disajikan pada Tabel XII berikut:

TABEL XII
HASIL UJI PAIRED *T-TEST*

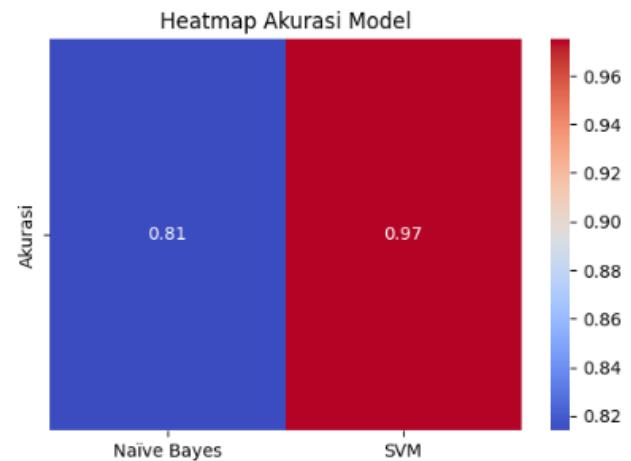
Parameter	Nilai
t-statistic	-11.636791865555802
p-value	8.229530353537046e-05
Kesimpulan	signifikan ($p < 0.05$)

Nilai *p-value* sebesar 0.0000823 menunjukkan bahwa hasil uji berada jauh di bawah ambang batas signifikansi 0,05. Hal ini menandakan bahwa perbedaan akurasi antara kedua model tidak terjadi secara acak, melainkan merupakan perbedaan yang signifikan secara statistik. Dengan demikian, dapat disimpulkan bahwa model *SVM* secara konsisten mempunyai kinerja yang lebih baik dibandingi dengan *Naïve Bayes*.

Temuan ini sejalan dengan hasil evaluasi sebelumnya, yang mengindikasikan bahwasanya *SVM* tidak hanya unggul dari segi akurasi, tetapi juga lebih stabil dan andal dalam mengklasifikasikan data sentimen pada ulasan pengguna aplikasi pinjaman *online*.

5) Heatmap Akurasi :

Sebagai bagian dari pelengkap evaluasi numerik dan statistik, dilakukan visualisasi hasil akurasi model dalam bentuk heatmap. Visualisasi ini bertujuan untuk memudahkan pembaca dalam membandingkan performa model secara cepat dan intuitif, tanpa harus menguraikan angka secara rinci.



Gambar 9 Heatmap Akurasi Model

Berdasarkan Gambar 9, terlihat bahwa model *SVM* memperoleh akurasi sebesar 0.97, sementara *Naïve Bayes* hanya mencapai 0.81. Gradiasi warna dalam heatmap dari biru ke merah memberikan penekanan visual terhadap selisih performa tersebut, di mana warna yang lebih terang menunjukkan tingkat akurasi yang lebih tinggi.

Visualisasi ini memperkuat rangkaian hasil evaluasi sebelumnya, mulai dari evaluasi awal, *confusion matrix*, hingga uji statistik, yang secara konsisten menunjukkan bahwa *SVM* memiliki performa yang lebih unggul dan stabil dibandingkan *Naïve Bayes* dalam tugas klasifikasi sentimen pada data ulasan aplikasi AdaKami.

I. Prediksi Model (Sisa Data)

Setelah melalui proses pelatihan dan evaluasi, kedua algoritma dalam penelitian ini *Support Vector Machine (SVM)* sebagai model unggulan dan *Naïve Bayes (NB)* sebagai pembandingdigunakan untuk melakukan prediksi sentimen terhadap 971 data ulasan yang tidak dilabeli secara manual. Seluruh data telah melewati tahapan pre-processing secara menyeluruh, kemudian dikonversi ke bentuk numerik menggunakan metode *TF-IDF* melalui fungsi *transform()*. Langkah ini dilakukan untuk mencegah data leakage serta menjaga konsistensi bobot kata sesuai dengan hasil pelatihan sebelumnya.

Prediksi dilakukan terhadap tiga kelas sentimen: positif, netral, dan negatif. Sebagian hasil klasifikasi dari kedua model ditampilkan pada Tabel XIII sebagai ilustrasi perbedaan prediksi.

TABEL XIII
HASIL PREDIKSI MODEL

	Preprocessed	pred_svm	pred_nb
1	biaya layan gede banget bunga gede banget udah limit coba ...	Negative	Negative
44	tolong hapus data an muhammad alvin aju pinjam bunga...	Netral	Negative
....
419	bagus pakai aplikasi	Positive	Negative
914	bismillah moga amanah karna proses pencairanya cepat ...	Positive	Negative
970	bayar ontime kali aju dana tolak alas gak sesuai no ktp ...	Negative	Negative

Tabel XIII memperlihatkan bahwa hasil prediksi antara kedua model tidak selalu sama. Dalam sejumlah kasus, *Naive Bayes* menunjukkan kecenderungan mengarahkan ulasan ke kategori sentimen negatif, meskipun isi ulasannya bersifat netral atau positif. Sebaliknya, *SVM* menghasilkan prediksi yang lebih proporsional, dengan kemampuan yang lebih baik dalam mengenali perbedaan nuansa sentimen dalam teks ulasan.

1) Distribusi Prediksi:

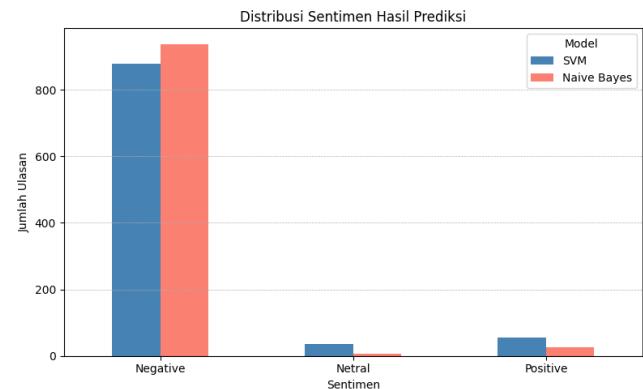
Setelah kedua model diterapkan pada 971 data ulasan yang belum dilabeli secara manual, hasil prediksi diklasifikasikan ke dalam tiga kategori sentimen, yaitu negatif, positif, dan netral. Rincian jumlah prediksi dari masing-masing model Sditampilkan pada Tabel XIV berikut :

TABEL XIV
JUMLAH PREDIKSI SENTIMEN

Sentimen	SVM	Naïve Bayes
Negative	878	938
Positive	56	26
Netral	37	7
Total	971	971

Tabel XIV menunjukkan bahwa sebagian besar data diklasifikasikan sebagai sentimen negatif oleh kedua model. Namun, model *Naïve Bayes* menghasilkan prediksi negatif yang jauh lebih dominan (938 data) dibandingkan *SVM* (878 data). Selisih yang cukup mencolok juga terlihat pada kategori positif dan netral, di mana *SVM* memprediksi lebih banyak ulasan sebagai positif (56 data) dan netral (37 data) dibandingkan prediksi *Naïve Bayes* yang hanya menghasilkan 26 dan 7 data secara berturut-turut.

Untuk memperjelas perbandingan antar model, hasil prediksi juga divisualisasikan dalam bentuk diagram batang yang ditampilkan pada Gambar 10.



Gambar 10 Distribusi Prediksi Model

Visualisasi ini memperkuat temuan bahwa meskipun kedua model mendekripsi dominasi opini negatif, *SVM* memberikan hasil klasifikasi yang lebih seimbang dan proporsional. Hal ini sejalan dengan hasil evaluasi sebelumnya yang menunjukkan kemampuan *SVM* dalam mengenali keragaman ekspresi sentimen, serta memperkuat alasan pemilihan *SVM* sebagai model utama dalam analisis sentimen ulasan pengguna aplikasi AdaKami.

2) Wordcloud Prediksi:

Sebagai bagian dari analisis lanjutan, dilakukan visualisasi *wordcloud* berdasarkan hasil klasifikasi sentimen model SVM terhadap 971 ulasan tanpa label. Tujuan dari visualisasi ini adalah untuk menampilkan kata-kata yang paling sering muncul dalam data, sehingga dapat memberikan gambaran umum mengenai topik yang banyak dibahas oleh pengguna.



Gambar 11 Wordcloud Prediksi Semua Sentimen

Gambar 11 menunjukkan bahwa kata seperti “aplikasi”, “pinjam”, “bayar”, “bunganya”, dan “gak” mendominasi. Kehadiran kata “tagih”, “teror”, “ojk”, dan “jatuh tempo” mencerminkan keluhan pengguna terhadap penagihan, bunga, dan pengalaman negatif lainnya.

Secara umum, visualisasi ini mendukung distribusi sentimen sebelumnya yang menunjukkan dominasi opini negatif. *Wordcloud* juga memberikan insight tambahan bagi pengembang aplikasi untuk memahami aspek layanan yang sering dikritik oleh pengguna.

IV. KESIMPULAN

Penelitian ini bertujuan untuk melakukan klasifikasi sentimen ulasan pengguna terhadap aplikasi pinjaman online AdaKami, yang datanya diperoleh dari platform App Store. Tahapan penelitian mencakup proses *pre-processing* teks, pelabelan manual terhadap 1000 data oleh pakar, serta pelatihan dan pengujian dua algoritma pembelajaran mesin, yaitu *Support Vector Machine (SVM)* dan *Naïve Bayes (NB)*.

Hasil evaluasi menunjukkan bahwa *SVM* memiliki performa yang lebih unggul dibandingkan *Naïve Bayes* dari berbagai aspek, termasuk akurasi, *F1-score*, hasil *cross-validation*, dan pengujian statistik (*paired t-test*). Model *SVM* mencatatkan akurasi sebesar 97,5% pada data validasi manual dan 84,69% sebagai rata-rata dari *5-fold cross-validation*. Sementara itu, *Naïve Bayes* hanya mencapai akurasi 81,4% dan 76,20% pada pengujian yang sama. Uji statistik menghasilkan nilai *p-value* di bawah 0,05, menandakan bahwa perbedaan performa kedua model bersifat signifikan secara statistik.

Setelah ditetapkan sebagai model terbaik, SVM digunakan untuk memprediksi 971 data ulasan tanpa label, yang hasilnya menunjukkan dominasi sentimen negatif, diikuti oleh positif dan netral. Visualisasi *wordcloud* mendukung temuan tersebut, dengan kata-kata dominan yang merepresentasikan keluhan terhadap bunga, proses penagihan, dan pengalaman pengguna yang kurang menyenangkan.

Penelitian ini juga mengidentifikasi bahwa kategori sentimen netral merupakan tantangan tersendiri karena jumlah datanya yang relatif kecil dan bentuk ekspresi opini yang cenderung tidak eksplisit. Meski demikian, model *SVM* mampu mengatasi tantangan ini dengan memberikan hasil klasifikasi yang lebih proporsional.

Secara keseluruhan, dapat disimpulkan bahwa *Support Vector Machine (SVM)* adalah algoritma yang paling sesuai digunakan untuk klasifikasi sentimen pada aplikasi pinjaman online, terutama ketika dihadapkan pada data teks yang tidak seimbang. Temuan ini diharapkan dapat menjadi bahan pertimbangan bagi pengembang aplikasi dalam memahami persepsi pengguna serta meningkatkan kualitas layanan yang ditawarkan.

Keterbatasan dalam penelitian ini adalah belum diterapkannya teknik balancing data seperti *SMOTE* atau *class weighting*. Hal ini menyebabkan ketimpangan distribusi data antar kelas tetap ada, dan berkontribusi pada rendahnya performa model dalam mengenali sentimen netral. Penelitian lanjutan disarankan untuk menerapkan teknik

balancing tersebut guna meningkatkan keandalan dan keakuratan model klasifikasi, khususnya dalam mengenali sentimen netral.

DAFTAR PUSTAKA

- [1] M. Hafil, "AdaKami dorong perekonomian masyarakat Indonesia," Republika Online, 2022. [Online]. Available:
 - [2] D. Febriana, A. N. Putri, dan R. A. Santoso, "Perbandingan algoritma dalam analisis sentimen aplikasi," Jurnal Informatika dan Komputasi, vol. 10, no. 3, 2023.
 - [3] R. A. E. Wahyuni dan B. E. Turisno, "Praktik finansial teknologi ilegal dalam bentuk pinjaman online ditinjau dari etika bisnis," Jurnal Pembangunan Hukum Indonesia, vol. 1, no. 3, pp. 379–391, 2019.
 - [4] D. N. Sastradinata, "Aspek hukum lembaga pinjaman online ilegal di Indonesia," Jurnal Independent, vol. 8, no. 1, p. 293, 2020.
 - [5] Kementerian Keuangan, "Laporan fintech lending di Indonesia," 2021.
 - [6] S. R. Irawan, L. Humaira, dan S. A. Sjarif, "Peran AdaKami dalam akses pinjaman tanpa jaminan," Lex Patrimonium, vol. 3, no. 1, pp. 1–17, 2024.
 - [7] M. Raharja dan B. Sulistyo, "Analisis sentimen ulasan aplikasi media berita online," Jurnal Informatika dan Sistem Informasi, vol. 8, no. 4, pp. 45–53, 2021.
 - [8] A. Yulianto dan D. Kurniawan, "Analisis sentimen ulasan aplikasi dengan algoritma Naïve Bayes: Studi kasus pada platform E-Commerce," Jurnal Sistem Informasi, vol. 16, no. 1, pp. 45–56, 2024.
 - [9] A. Sasmita dan B. Harto, "Analisa perhitungan suku bunga pinjaman harian pada aplikasi pinjaman online legal menggunakan metode simple interest," ATRABIS, vol. 7, no. 2, pp. 132–139, 2021.
 - [10] F. H. Setiawan, I. P. Radjimin, dan M. Ariani, "Pinjaman online: Perilaku konsumtif mahasiswa Surabaya dalam menunjang status sosial," MSEJ, vol. 5, no. 1, pp. 413–425, 2023.
 - [11] H. Putra dan L. Rahmawati, "Analisis sentimen pada tweet tentang pinjaman online menggunakan metode Random Forest," Jurnal Informatika dan Komputasi, vol. 14, no. 3, pp. 213–220, 2022.
 - [12] B. Pang dan L. Lee, "Opinion mining and sentiment analysis," Foundations and Trends in Information Retrieval, vol. 2, no. 1–2, pp. 1–135, 2008.] C. Cortes dan V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.
 - [13] N. Cristianini dan J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, 2000.
 - [14] F. A. Indriyani, A. Fauzi, dan S. Faisal, "Analisis sentimen aplikasi TikTok menggunakan algoritma Naïve Bayes dan Support Vector Machine," TEKNOSAINS, vol. 10, no. 2, pp. 176–184, 2024.
 - [15] Rahayu, I. P., Fauzi, A., & Indra, J. (2022). Analisis sentimen terhadap program Kampus Merdeka menggunakan Naïve Bayes dan Support Vector Machine. Jurnal Sistem Komputer dan Informatika (JSON), 4(2), 296–301.
 - [16] Y. Cahyana and A. M. Siregar, "Analisis Sentimen Pembelajaran Tatap Muka Terbatas (PTMT) Selama Pandemik COVID-19 Menggunakan Algoritma Naïve Bayes," Petir: Jurnal Teknik dan Rekayasa, vol. 16, no. 2, pp. 200–211, 2023.
 - [17] F. N. Azzahra, T. Rohana, R. Rahmat, and A. R. Juwita, "Penerapan Metode Naïve Bayes Dalam Klasifikasi Spam SMS Menggunakan Fitur Teks Untuk Mengatasi Ancaman Pada Pengguna," Journal of Information System Research (JOSH), vol. 5, no. 3, pp. 873–880, 2024.
 - [18] M. F. Haikal, J. Indra, dan R. Rahmat, "Analisis sentimen bakal calon presiden Indonesia 2024 dengan algoritma Naïve Bayes," JUTISI, vol. 13, no. 1, pp. 43–51, 2024.
 - [19] A. Adela, F. Nurjanah, dan R. Setiawan, "Analisis sentimen ulasan aplikasi pinjaman online," Jurnal Teknologi Informasi dan Komunikasi, vol. 15, no. 2, pp. 112–125, 2023.