

# ANALISIS SENTIMEN TERHADAP APLIKASI PINJAMAN ONLINE DAN PAYLATER PADA TWITTER DENGAN MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

**Muhammad Iqbal Wibowo, Fridy Mandita**

Teknik Informatika, Fakultas Teknik, Universitas 17 Agustus 1945, Surabaya, Indonesia,

Email: [muhiqbalwibowo@gmail.com](mailto:muhiqbalwibowo@gmail.com)

Teknik Informatika, Fakultas Teknik, Universitas 17 Agustus 1945, Surabaya, Indonesia,

Email: [fridymandita@untag-sby.ac.id](mailto:fridymandita@untag-sby.ac.id)

## ABSTRACT

Sentiment analysis is the computational study of expressed opinions, sentiments, and emotional expressions. The goal of sentiment analysis is to classify sentences or opinions into specific categories, such as Positive, Negative, or Neutral. Social media, as an interactive platform on the internet, allows users to communicate. Online lending is a financial service that operates online through a mobile application platform. The existence of these applications creates a lot of conversation in the community, especially in tweets related to online lending on social media Twitter. In a society of opinion, there are both supportive and opposing opinions, and analysis of social media conversations can provide insight into the general sentiment of the public towards online lending. One approach that can be used for this analysis is the Support Vector Machine method.

**Keywords:** Online Loans, Paylater, Support Vector Machine Method.

## ABSTRAK

Analisis sentimen adalah studi komputasional terhadap opini, sentimen, dan ekspresi emosional yang diungkapkan. Tujuan dari analisis sentimen adalah mengklasifikasikan kalimat atau opini ke dalam kategori tertentu, seperti Positif, Negatif, atau Netral. Media sosial, sebagai platform interaktif di internet, memungkinkan pengguna untuk berkomunikasi. Pinjaman online merupakan layanan keuangan yang beroperasi secara daring melalui platform aplikasi mobile. Keberadaan aplikasi tersebut menciptakan banyak perbincangan di masyarakat, terutama dalam cuitan yang berkaitan dengan pinjaman online di media sosial Twitter. Dalam masyarakat berpendapat, terdapat pendapat yang mendukung dan menentang, dan analisis terhadap percakapan di media sosial dapat memberikan wawasan mengenai sentimen umum masyarakat terhadap pinjaman online. Salah satu pendekatan yang dapat digunakan untuk analisis ini adalah metode Support Vector Machine.

**Kata Kunci:** Pinjaman Online, Paylater, Metode Support Vector Machine.

## 1. PENDAHULUAN

Kemajuan teknologi dalam industri finansial telah menciptakan terobosan penting, terutama dalam bentuk kredit digital dan sistem paylater. Konsep peminjaman uang yang dahulu dilakukan secara konvensional kini telah bertransformasi menjadi layanan digital yang memanfaatkan kemajuan teknologi untuk mengefisienkan proses transaksi. Masyarakat kini dapat mengakses layanan peminjaman uang tanpa perlu mengunjungi kantor pemberi pinjaman

secara fisik, cukup melalui perangkat telepon pintar mereka. Kehadiran teknologi finansial (FinTech) bahkan telah mampu mengambil alih fungsi-fungsi yang sebelumnya hanya dapat dilakukan oleh institusi keuangan tradisional seperti bank, termasuk dalam hal menyediakan sistem pembayaran yang inovatif. [1].

Sementara itu, paylater adalah layanan pembiayaan online yang digunakan untuk transaksi di e-commerce, memungkinkan konsumen untuk membeli barang secara online terlebih dahulu, dengan pembayaran dilakukan di

kemudian hari secara angsuran. Layanan ini memudahkan konsumen untuk melakukan pembelian secara instan, dengan pembayaran yang dapat dilakukan di masa mendatang [2].

Meskipun menawarkan kemudahan, layanan ini juga menimbulkan tantangan bagi masyarakat Indonesia. Banyaknya penyedia layanan, termasuk yang ilegal, serta kurangnya regulasi yang memadai menimbulkan keraguan dan risiko bagi pengguna. Situasi ini diperparah oleh kasus penipuan dan praktik tidak etis yang dilaporkan. Beberapa penyedia jasa paylater tidak memiliki izin dari Otoritas Jasa Keuangan, atau dapat dikatakan ilegal, dan sering kali menawarkan suku bunga yang tinggi.

Mengingat beragamnya pandangan masyarakat terhadap layanan ini, penting untuk memahami sentimen publik yang sebenarnya. Media sosial, khususnya Twitter, menjadi wadah utama bagi masyarakat untuk menyuarakan opini mereka[3]. Oleh karena itu, analisis sentimen terhadap percakapan di Twitter tentang pinjaman online dan paylater dapat memberikan wawasan berharga tentang persepsi publik terhadap layanan fintech ini.

Untuk memecahkan masalah ini, penelitian ini akan memanfaatkan teknik text mining dan analisis sentimen terhadap tweet yang berkaitan dengan pinjaman online dan paylater. Penelitian ini menerapkan pendekatan Support Vector Machine (SVM) dalam mengkategorikan sentimen menjadi tiga kelompok yaitu positif, negatif, dan netral, dengan studi kasus pada aplikasi pinjaman online dan paylater yang berada di bawah pengawasan dan terdaftar secara resmi di Otoritas Jasa Keuangan (OJK). Data yang dianalisis akan dibatasi pada tweet dari dua tahun terakhir (2022-2023) untuk mendapatkan gambaran sentimen terkini.

Tujuan penelitian ini adalah:

1. Mengklasifikasikan sentimen tweet pengguna terkait layanan pinjaman online dan paylater ke dalam 3 kelas (positif, negatif, netral) menggunakan metode SVM.
2. Mengevaluasi performa model analisa sentimen terhadap pinjaman online dan paylater dengan metode SVM.

3. Menganalisis perbedaan sentimen pengguna Twitter terhadap layanan pinjaman online dan paylater berdasarkan jenis aplikasinya.

Analisis sentimen merupakan kajian yang menggabungkan bidang pengolahan bahasa alami, linguistik komputasional, dan penambangan teks untuk memahami bagaimana seseorang merasakan dan berpendapat tentang suatu hal. Bidang ini berfokus pada pengkajian berbagai ekspresi seperti opini, perasaan, penilaian, dan respon emosional masyarakat terhadap berbagai aspek, mulai dari produk hingga layanan atau aktivitas tertentu [4]. Dalam konteks fintech, khususnya pinjaman online dan paylater, analisis sentimen dapat memberikan wawasan berharga tentang persepsi publik. Support Vector Machine (SVM) adalah metode yang telah terbukti efektif dalam klasifikasi teks dan analisis sentimen. Metode ini akan digunakan dalam penelitian ini untuk mengklasifikasikan tweet ke dalam kategori sentimen yang berbeda.

Diharapkan hasil penelitian ini dapat memberikan wawasan baru mengenai persepsi, sentimen, dan tanggapan masyarakat terhadap layanan fintech di Indonesia saat ini. Informasi ini dapat bermanfaat bagi regulator untuk mengembangkan kebijakan yang lebih efektif dalam mengatur industri fintech, pelaku industri untuk memahami persepsi publik dan meningkatkan layanan mereka, akademisi sebagai referensi untuk penelitian lebih lanjut di bidang fintech dan analisis sentimen, serta masyarakat untuk mendapatkan pemahaman yang lebih baik tentang layanan pinjaman online dan paylater.

Secara keseluruhan, penelitian ini diharapkan dapat berkontribusi pada pengembangan ekosistem fintech yang sehat dan mendukung kemajuan ekonomi digital di Indonesia.

## 2. TINJAUAN PUSTAKA

Analisis sentimen berbasis algoritma Support Vector Machine (SVM) digunakan dalam penelitian ini untuk mengurai pemahaman tentang sikap masyarakat terhadap layanan pinjaman online dan paylater. Beberapa penelitian terdahulu telah menggunakan metode SVM untuk analisis sentimen di berbagai

konteks. [5] melakukan analisis sentimen terhadap implementasi sistem E-Tilang di Twitter, menghasilkan 2.827 data yang terdiri dari 426 sentimen positif, 491 negatif, dan 1.910 netral. Hasil ini digunakan sebagai umpan balik untuk evaluasi dan pengembangan sistem e-Tilang. [3] mengevaluasi efektivitas algoritma SVM dalam mengklasifikasikan sentimen terkait PPKM di wilayah Jawa dan Bali, dengan membandingkan performa Kernel Linier dan Kernel RBF. Penelitian ini menggunakan dataset seimbang yang terdiri dari 757 tweet positif dan 757 tweet negatif, membuktikan keberhasilan PPKM dalam menekan penyebaran Covid-19. Penelitian [6] menganalisis sentimen masyarakat terhadap program Kartu Prakerja di Twitter menggunakan metode SVM dan evaluasi Confusion Matrix. Hasil menunjukkan dominasi sentimen netral sebesar 98,34%, dengan sentimen negatif 0,99% dan positif 0,66%. [7] menganalisis layanan Shopee Paylater dan Go Paylater, menemukan kecenderungan sentimen netral pada kedua layanan. Model SVM terbaik mencapai akurasi 89,74% untuk Shopee Paylater dan 90,27% untuk Go Paylater dengan rasio data 80:20.[8] menganalisis sentimen pengguna Shopee menggunakan SVM terhadap 3.000 ulasan yang dikumpulkan melalui scraping. Penelitian ini mencapai akurasi 98% dan F1-score 98%, membuktikan efektivitas SVM dalam mengklasifikasikan sentimen pada platform e-commerce.

### **2.1. Analisa Sentimen**

Analisis sentimen merupakan teknik untuk mengidentifikasi dan mengekstraksi opini serta sentimen dari data tekstual secara otomatis. Teknik ini memproses data tekstual untuk menentukan apakah suatu teks mengandung makna positif, negatif ataupun netral. Tujuan dari analisis sentimen adalah menemukan pola sentimen dalam suatu tulisan. Untuk mendapatkan hasil analisis sentimen yang akurat, idealnya dibutuhkan data dan sumber daya yang besar, terutama jika menggunakan pendekatan pembelajaran mesin murni [9].

### **2.2. Natural Language Processing (NLP)**

NLP adalah cabang ilmu yang menggabungkan komputer dan kecerdasan buatan untuk memahami bahasa manusia. Tujuannya adalah membuat mesin bisa

berkomunikasi dengan manusia. Ini tantangan besar karena bahasa manusia rumit dan bisa bermakna ganda.

Menurut [10], ada dua cara utama dalam NLP, Analisis sintaksis adalah mempelajari susunan kata dalam kalimat agar sesuai aturan bahasa. Lalu analisis semantik adalah memahami arti dan maksud dari kalimat tersebut. Kedua cara ini penting untuk memastikan bahasa yang diproses sudah benar dan bisa dipahami oleh komputer. Dengan NLP, komputer bisa "belajar" bahasa manusia untuk berinteraksi lebih baik dengan kita. Meskipun sulit, kemampuan ini membuka banyak peluang untuk teknologi yang lebih canggih dan mudah digunakan.

### **2.3. Text Mining**

Text mining adalah cara untuk mendapatkan informasi penting dari teks yang tidak teratur, seperti dokumen. [11] menjelaskan bahwa tujuan utamanya adalah menemukan kata-kata kunci yang menggambarkan isi dokumen. Ini membantu kita memahami hubungan antara berbagai dokumen.

Text mining bekerja dengan data yang tidak terstruktur, seperti kumpulan teks atau dokumen. Tujuannya adalah menemukan pola-pola menarik dalam data ini. Dengan text mining, kita bisa mengambil informasi berharga dari tumpukan data teks yang tadinya sulit diatur. Ini sangat berguna untuk menganalisis banyak dokumen sekaligus dan menemukan hubungan atau pola yang mungkin tidak terlihat dengan cara biasa.

### **2.4. Support Vector Machine (SVM)**

Support Vector Machine (SVM) adalah teknik supervised learning yang digunakan untuk klasifikasi dan regresi. Menurut Pratama et al. (2018), SVM dianggap sebagai salah satu pendekatan terbaik untuk permasalahan klasifikasi. Tujuan utama SVM adalah menemukan hyperplane optimal sebagai pemisah antara kelas-kelas, dengan maksud memaksimalkan margin di antara mereka.

SVM memiliki keunggulan dalam menerapkan pemisah linear pada data non-linear berdimensi tinggi melalui penggunaan fungsi kernel [12]. Sebuah hyperplane dianggap optimal ketika memiliki margin terbesar, di mana margin adalah dua kali jarak antara hyperplane dan support vector [13].

Banyak penelitian menyatakan bahwa SVM merupakan pendekatan yang paling tepat untuk klasifikasi teks, termasuk dalam analisa sentimen (Santoso et al., 2019). SVM terbukti lebih efektif dalam klasifikasi dibanding Metode K-Nearest Neighbor (KNN) dengan tingkat keakuratan mencapai 89,70% [14]. Menurut [15], sifat umum SVM meliputi, prinsip dasar sebagai klasifikasi linear, kemampuan untuk menangani klasifikasi multi-kelas, keunggulan SVM meliputi, generalisasi yang baik, kemampuan mengatasi curse of dimensionality, implementasi yang relatif mudah

Meskipun demikian, perlu diingat bahwa akurasi model SVM sangat dipengaruhi oleh pilihan fungsi kernel dan parameter yang digunakan.

### 2.5. Crawling data

Crawling data adalah metode pengumpulan informasi dari berbagai website. Awalnya digunakan oleh peramban web untuk mengisi indeks, teknik ini kini diterapkan dalam berbagai konteks, termasuk pengumpulan data untuk analisis dan riset. Melalui crawling, data dapat dikumpulkan secara menyeluruh dari sebuah website atau difokuskan pada informasi tertentu.

Proses crawling menghadapi beberapa tantangan, seperti:

1. Struktur data yang besar dan kompleks
2. Kurangnya kontrol terpusat terhadap konten

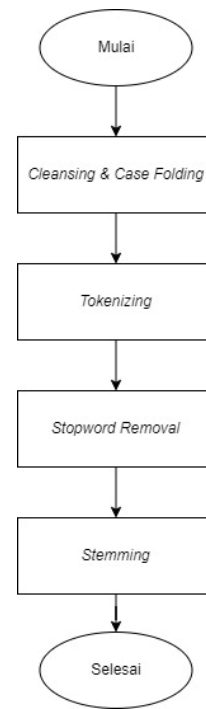
Untuk mengatasi tantangan ini, beberapa strategi dapat diterapkan:

1. Pemilihan algoritma crawling yang efisien
2. Optimalisasi infrastruktur
3. Pengelolaan data yang efektif

Dengan menerapkan strategi-strategi tersebut, efektivitas dan efisiensi proses pengumpulan informasi melalui teknik crawling dapat ditingkatkan.

### 2.6. Preprocessing data

Preprocessing data adalah proses membersihkan dan mempersiapkan data tekstual sebelum diolah lebih lanjut. Berikut tahap dalam preprocessing data, yaitu :



Gambar 1. Tahap Preprocessing

#### 1. Cleansing

Cleansing bertujuan untuk menganalisis dan meningkatkan kualitas data dengan memodifikasi, mengubah, atau menghapus bagian data yang dianggap tidak diperlukan. Tahapan cleansing dilakukan untuk memastikan data yang akan diproses selanjutnya benar-benar berkualitas dan relevan.

Tabel 1. Tahap Cleansing

No	Kalimat	Hasil Cleansing
1.	Kamu tim punya dana dana darurat atau terpaksa pinjol nih hehe?	Kamu tim punya dana darurat atau terpaksa pinjol nih hehe
2.	ada yg open joki lunasin all pinjol ga ya???	ada yg open joki lunasin all pinjol ga ya

#### 2. Case Folding

Merupakan tahap standarisasi teks dengan mengkonversi seluruh karakter dalam kata maupun kalimat ke dalam bentuk huruf kecil, sehingga mempermudah sistem dalam melakukan pencarian dan membandingkan kata atau kalimat.

Tabel 2. Tahap Case Folding

No	Kalimat	Hasil Case Folding
1.	Yang bisa menyelesaikan menyelesaikan	yang bisa menyelesaikan masalah pinjol itu adalah

masalah pinjol itu adalah diri kalian sendiri.. nenek dan teman2 yang lain hanya membantu meluruskan agar kalian tidak panik di serang dc/fc.. yang penting meluruskan agar kalian tidak panik di serang DC/FC.. yang penting 1 jangan panik.. ga usah takut yang kalian hadapi juga manusia juga kok..

2. STOP PAKE JOKI PINJOL KALAU NGGAK MAU TAMBAH BENGKAK!

stop pake joki pinjol kalau nggak mau tambah bengkak!

### 3. Tokenizing

Tokenizing adalah proses pemecahan kalimat menjadi kata-kata. Pada tahap tokenizing, kalimat dalam data tekstual yang sedang diproses akan dipecah menjadi token atau kata-kata penyusun kalimat tersebut. Tujuannya adalah untuk mendapatkan kata per kata dari komentar atau kalimat yang ada.

Tabel 3. Tahap Tokenizing

No	Kalimat	Hasil Tokenizing
1.	disangka pinjol sama anak sendiri	“disangka” “pinjol” “sama” “anak” “sendiri”
2.	stop pake joki pinjol kalau nggak mau tambah bengkak	“stop” “pake” “joki” “pinjol” “kalau” “nggak” “mau” “tambah” “bengkak”

### 4. Stopword Removal

Pada tahap ini dilakukan proses penghapusan stopword, yaitu kata-kata umum yang frekuensi kemunculannya tinggi namun tidak memiliki nilai semantik yang signifikan, seperti kata penghubung "yang", "ini", "dengan" dan kata-kata sejenis lainnya.

Tabel 4. Tahap Stopword Removal

No	Kalimat	Hasil <i>Stopword Removal</i>
1.	Jangan terkecoh sama promo si Pinjol ini resmi berOjk tapi bunga biaya layanan diatas pinjol ilegal Sehari bisa 150ribuan dgn	'Jangan', 'terkecoh', 'promo', 'si', 'Pinjol', 'yes', 'resmi', 'berOjk', 'bunga', 'biaya', 'layanan', 'didas', 'pinjol', 'ilegal', 'Sehari', '150ribuan', 'dgn', 'tenor', 'pinjol', 'ilegal', '14hari', 'Jelas', 'kartel', 'Sehari', 'bisa', 'melenggang', 'bebas', '150ribuan', 'dgn', 'punishment', 'apapun',

tenor cuma 14hari Jelas kartel tapi masih melenggang bebas tanpa ada punishment apapun Eh diperiksa aja gak gimana mau ada punishment dududud

2. SEOJK yang baru mengatur soal kontak darurat pinjol. Dari pasal pertama udah diatur BUKAN untuk melakukan penagihan.

'Eh', 'diperiksa', 'aja', 'gak', 'gimana', 'punishment', 'dududud'

'SEOJK', 'mengatur', 'kontak', 'darurat', 'pinjol', '!', 'Dari', 'pasal', 'udah', 'diatur', 'BUKAN', 'penagihan'

### 5. Stemming

Proses ini bertujuan untuk menemukan bentuk dasar dari sebuah kata, dengan membuang semua imbuhan yang melekat pada kata tersebut.

Tabel 5. Tahap Sremming

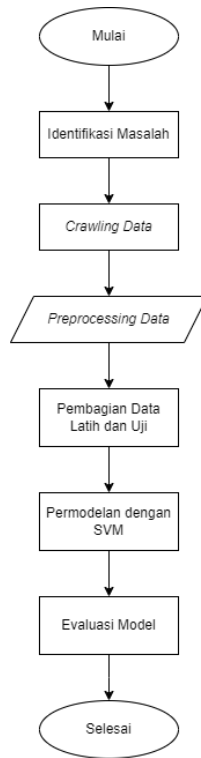
No	Kalimat	Hasil <i>Stemming</i>
1.	Batu sandungan kelen di Indon itu di depan mata. Judi Online Pinjol Aset marketplace milik asing Stunting Hedonisme Kurikulum berubah-ubah tanah adat dirampok monopoli air bersih. Korupsi Kolusi Nepotisme juga jadi akar sandungan Indonesia Emas. Gajah di pelupuk mata.	batu sandungan kelen di indon itu di depan mata judi online pinjol aset marketplace milik asing stunting hedonisme kurikulum ubah tanah adat rampok monopoli air bersih korupsi kolusi nepotisme juga jadi akar sandungan indonesia emas gajah di pelupuk mata
2.	OJK Sebut 3 Kalangan Ini Paling Banyak Terbelit Utang Pinjol Ilegal <a href="https://t.co/qhpJOfqLJF">https://t.co/qhpJOfqLJF</a>	ojk sebut 3 kalang ini paling banyak belit utang pinjol ilegal

### 3. METODELOGI PENELITIAN

Bab ini memaparkan seluruh instrumen dan sumber daya yang diimplementasikan dalam studi analisis sentimen terhadap layanan pinjaman online dan paylater. Data penelitian bersumber dari platform media sosial X (sebelumnya dikenal sebagai Twitter), dengan periode pengambilan data mencakup rentang

waktu dari tahun 2023 hingga saat ini. Seluruh perangkat dan metodologi yang digunakan ditujukan untuk proses pengumpulan, pengolahan, dan analisis data dalam rangka memahami persepsi publik terhadap layanan pinjaman online dan paylater tersebut.

### 3.1. TAHAPAN PENELITIAN



Gambar 2. Alur Penelitian

#### 1. Identifikasi Masalah

Tahap ini merupakan langkah fundamental dalam menganalisis sentimen terhadap aplikasi pinjaman online dan paylater di Twitter. Fokus utama mencakup pengamatan terhadap tren penggunaan layanan finansial digital di Indonesia, termasuk aspek regulasi OJK dan perbedaan persepsi pengguna terhadap aplikasi resmi dan ilegal.

#### 2. Pengumpulan Data

Data tweet dikumpulkan melalui metode crawling menggunakan library Tweet Harvest dengan Python. Pengambilan data dilakukan periode 2023-2024, dengan fokus pada aplikasi yang diawasi OJK seperti Akulaku dan Kredivo. Data disimpan dalam format CSV.

#### 3. Preprocessing Data

Tahap ini meliputi serangkaian proses pengolahan data mentah menjadi data siap

analisis, mencakup cleansing, tokenizing, stopword removal, dan stemming untuk menghasilkan output dengan akurasi tinggi.

#### 4. Pembagian Data

Dataset dibagi menjadi data latih dan data uji dengan rasio 80:20. Data latih digunakan untuk pembelajaran model SVM, sedangkan data uji untuk evaluasi performa prediksi.

#### 5. Pemodelan dengan SVM

Implementasi algoritma SVM menggunakan kernel linear untuk transformasi data, didukung dengan vektorisasi teks menggunakan metode TF-IDF. Pemilihan kernel linear didasarkan pada efisiensinya dalam menangani data teks berdimensi tinggi (Styawati et al., 2021; Wijaya et al., 2021).

#### 6. Evaluasi Model

Evaluasi menggunakan Confusion Matrix dan metrik evaluasi (akurasi, presisi, recall, f1-score).

		Nilai Aktual	
		1 (Positive)	0 (Negative)
Nilai Aktual	1 (Positive)	TP (True Positive)	FP (False Positive)
	0 (Negative)	FN (False Negative)	TN (True Negative)

Gambar 3. Confusion Matrix

Dalam evaluasi model klasifikasi, terdapat empat kategori hasil prediksi: True Positive (TP) yang menunjukkan ketepatan prediksi positif, True Negative (TN) yang mengindikasikan ketepatan prediksi negatif, False Positive (FP) yang menunjukkan kesalahan prediksi positif, dan False Negative (FN) yang mengindikasikan kesalahan prediksi negatif.

Rumus mengenai Confusion Matrix untuk menghitung nilai akurasi, presisi, *recall* :

$$akurasi = \frac{TP + TN}{Total} \tag{1}$$

Akurasi, yang dirumuskan dalam Persamaan (1), mengukur tingkat ketepatan model dengan membandingkan jumlah prediksi yang benar (mencakup kedua kelas positif dan negatif)

dengan total keseluruhan data dalam pengujian. Nilai ini menggambarkan seberapa akurat model dapat melakukan klasifikasi terhadap seluruh dataset yang diuji.

$$presisi = \frac{TP}{TP + TF} \quad (2)$$

Presisi, yang ditunjukkan dalam Persamaan (2), merupakan metrik yang mengevaluasi keakuratan model dalam mengidentifikasi kelas positif, dengan cara membandingkan jumlah prediksi positif yang tepat dengan total seluruh prediksi positif yang dihasilkan model.

$$recall = \frac{TP}{TP + FN} \quad (3)$$

Persamaan (3) adalah rumus recall berguna menunjukkan kemampuan model dalam mengenali data positif yang sebenarnya, dihitung dari perbandingan antara prediksi positif yang tepat dengan total data yang seharusnya positif.

## 7. Penarikan Kesimpulan

Tahap akhir meliputi analisis sentimen masyarakat terhadap aplikasi pinjaman online dan paylater, serta evaluasi performa algoritma SVM dalam kasus ini.

## 4. HASIL DAN PEMBAHASAN

Dalam penelitian ini akan dijabarkan tahapan-tahapan metodologis yang mencakup proses pengumpulan, pengolahan, dan transformasi data, serta evaluasi hasil. Analisis mendalam akan diperkuat dengan berbagai bentuk visualisasi seperti data statistik, tabel, dan grafik untuk meningkatkan pemahaman. Hasil penelitian ini diproyeksikan dapat memberikan fondasi untuk pengembangan studi selanjutnya.

### 4.1. Tahap Pengumpulan Data

Pada tahap ini, data yang berupa tweet dari platform Twitter akan dikumpulkan menggunakan metode crawling. Beberapa kata kunci yang digunakan antara lain, Akulaku dan Kredivo.

```

[1] # Import required python package
pip install pandas

# Install Node.js (because tweet-harvest built using Node.js)
sudo apt-get update
sudo apt-get install -y ca-certificates curl gnupg
sudo mkdir -p /etc/apt/keyrings
curl -fsSL https://deb.nodesource.com/gpgkey/gpgkey.asc | sudo gpg --dearmor -o /etc/apt/keyrings/nodesource.gpg
NODE_MAJOR=20 sudo "sudo" [signed-by=/etc/apt/keyrings/nodesource.gpg] https://deb.nodesource.com/node_20.x_nodistro.tar.xz
sudo apt-get update
sudo apt-get install nodejs -y
make --
  
```

```

[] # Crawl Data
filename = "tweet-kredivo-nov-2023.csv"
search_keywords = "kredivo since:2023-01-01 until:2023-12-31 -is:retweet -is:reply lang:id -'gestun' -'cali' -'joki' -'jasa' -'news' -'berita'"
limit = 500
cmd = "tweet-harvest@2.6.1 -o '{filename}' -s '{search_keywords}' -l {limit} --token {twitter_auth_token}"
  
```

Gambar 4. Proses Crawling data

Gambar 4 diatas menunjukkan tahap implementasi crawling data dari Twitter dengan menggunakan tool bernama tweet-harvest, serta memanfaatkan parameter pencarian spesifik.

```

[8] num_tweet = len(tweet_kredivo)
print(f"Jumlah dataset adalah {num_tweet} data")

Jumlah dataset adalah 1515 data
  
```

Gambar 5. Hasil Crawling data Kredivo

Hasil akhir crawling dari bulan November 2023 – November 2024 menunjukkan bahwa terdapat 1515 tweet yang sesuai dengan filter pencarian.

### 4.2. Tahap Pengumpulan Data

```

[9] tweet_kredivo.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1515 entries, 0 to 1514
Data columns (total 16 columns):
 #   Column                Non-Null Count  Dtype
---  ---
 0   conversation_id_str    1515 non-null  object
 1   created_at            1514 non-null  object
 2   favorite_count        1514 non-null  object
 3   full_text             1514 non-null  object
 4   id_str               1514 non-null  object
 5   image_url            366 non-null   object
 6   in_reply_to_screen_name 232 non-null   object
 7   lang                 1514 non-null  object
 8   location              1058 non-null  object
 9   quote_count          1514 non-null  object
10  reply_count           1514 non-null  object
11  retweet_count         1514 non-null  object
12  tweet_url             1514 non-null  object
13  user_id_str           1514 non-null  object
14  username              1514 non-null  object
15  language              1364 non-null  object
dtypes: object(16)
memory usage: 189.5+ KB
  
```

Gambar 6. Data Kredivo sebelum melalui preprocessing

Data mentah ini terdiri dari tweet yang dikumpulkan melalui proses crawling, dan diketahui dengan jumlah 16 kolom, dan 1515 baris data mentah yang siap untuk diolah lebih lanjut.

```
def clean_text(text):
    # Menghapus Username twitter
    text = re.sub('@[\w]{1,20}', '', text)

    # Menghapus https dan http
    text = re.sub('(?:https?://|http?://|www)', '', text)

    # Menghilangkan Tanda Baca, kecuali spasi
    text = ''.join(char for char in text if char not in string.punctuation or char == ' ')

    # Mengganti karakter html dengan tanda petik
    text = re.sub('<.*?>', '', text)

    # Mengganti Line baru dengan spasi
    text = re.sub("\n", " ", text)

    # Menghapus angka
    text = re.sub("[0-9]+", '', text)]

    # Mengubah ke huruf kecil
    text = text.lower()

    # Normalisasi kata dengan huruf berulangi
    text = re.sub(r'(\w)\1{2,}', r'\1\1', text)

    # Menghapus single char, tapi mempertahankan spasi
    text = ' '.join(word for word in text.split() if len(word) > 1])

    # Memisahkan dan menggabungkan kata (menghapus spasi berlebih)
    text = ' '.join(text.split())

    return text
```

Gambar 7. Proses Case Folding dan Cleansing

Proses dimulai dengan penghapusan elemen-elemen yang spesifik untuk Twitter, seperti nama pengguna dan URL, menggunakan substitusi regex. Selanjutnya, fungsi ini menghilangkan tanda baca (kecuali spasi) dan karakter HTML, serta menormalisasi pemformatan teks dengan mengganti baris baru menjadi spasi tunggal.

```
[ ] tweet_kredivo['clean'] = [clean_text(i) for i in tweet]
tweet = tweet_kredivo['clean']

[ ] tweet_kredivo.iloc[:, [3, 16]].head(20)
```

	full_text	clean
0	HELP RT(tlh sewa buzzer jki luga) - a thr...	help rtfhah sewa buzzer jki luga thread ta...
1	Terimakasih om agus om koko kredivo alukaku S...	terimakasih om agus om koko kredivo alukaku se...
2	- Spjnjam : 9 juta - Alukaku : 3 juta - Kredit...	spjnjam juta alukaku juta kredit pintar juta a...
3	-star koreo YEAH YEAH yang bikin canduuuu http...	star koreo yeah yeah yang bikin canduu...
4	Pamer aman bayar tagihan kredivo 100k dapat vo...	pamer aman bayar tagihan kredivo dapat voucher...
5	Tadi iseng beli paket telkomsel disambungi ke...	tadi iseng beli paket telkomsel disambungi ke...
6	Buat kamu pecinta Brand H&amp;M ada kabar gemb...	buat kamu pecinta brand hamp ada kabar gemb...
7	PENCARAAN HARI III @jasa_pencarianpaylaterby...	pencarian hari iii open selapp hari pencarian ...
8	Belanja pake Kredivo di SELURUH Brand MAP disk...	belanja pake kredivo di seluruh brand map disk...
9	Belanja kebutuhan bulanan ngga perlu masuk daf...	belanja kebutuhan bulanan ngga perlu masuk daf...
10	Wikenan belanja hemat? Gaskuen scroll kebawah ...	wikenan belanja hemat gaskuen scroll kebawah d...
11	Yuk baca S&mp;Krya disini!!!- https://t.co/M...	yuk baca sampkrya disini
12	Dirgahayu Indonesiakul !!! stay tune info d...	dirgahayu indonesiaul !!! stay tune info d...
13	Ini dia yang ditunggu-tunggu!!! Untuk mempergi...	ini dia yang ditunggu-tunggu untuk mempergi...
14	Yang belum daftar kredivo mending incer promo...	yang belum daftar kredivo mending incer promo...
15	Udah tau belum kalo top up listrik pake Krediv...	udah tau belum kalo top up listrik pake krediv...
16	kredivo riveuh ih ya aloh sabar atuh nanti ju...	kredivo riveuh ih ya aloh sabar atuh nanti ju...
17	Beli produk belajar Ruanggru LEBIH BANYAK DIS...	beli produk belajar ruanggru lebih banyak dis...
18	Alhamdulillah tercahkan in case aku mengajuk...	ahamduillah tercahkan in case aku mengajuk...
19	Makasi loh udah kangen @kredivo https://t.co/...	makasi loh udah kangen

Gambar 8. Hasil dari proses Case Folding dan Cleansing

Hasil diatas menunjukkan 20 data teratas untuk di tampilkan.

```
[ ] import pandas as pd

# Membaca file CSV ke dalam DataFrame
df = pd.read_csv("emoticon_arti.csv")

# Mengonversi data emoticon dan artinya ke dalam dictionary
emoticon_dict = dict(zip(df['emoticon'], df['arti']))

# Fungsi untuk mengganti emoticon dengan teks berdasarkan dictionary, dengan spasi
def replace_emoticons(text, emoticon_dict):
    for emoticon, arti in emoticon_dict.items():
        text = text.replace(emoticon, f" {arti} ")
    return text

# Contoh teks dengan emoticon
tweet = "Hari ini sangat menyenangkan! 🥰❤️ Semangat untuk semuanya 🙌🏻"

# Mengganti emoticon dengan arti dari teks
cleaned_tweet = replace_emoticons(tweet, emoticon_dict)

print("Sebelum:", tweet)
print("Setelah:", cleaned_tweet)
```

Gambar 9. Mengubah emoticon menjadi teks

Fungsi ini bekerja dengan melakukan iterasi untuk setiap emoticon dan arti dalam dictionary. Pada setiap iterasi, fungsi text.replace() digunakan untuk mengganti emoticon dengan artinya, ditambah spasi di akhir. Hasil akhirnya adalah teks yang sudah diproses dikembalikan. Contoh penerapannya dapat dilihat dari output kode diatas.

```
[ ] pip install Sastrawi

Collecting Sastrawi
  Downloading Sastrawi-1.0.1-py2.py3-none-any.whl (289 bytes)
  Downloading Sastrawi-1.0.1-py2.py3-none-any.whl (289 kB)
Installing collected packages: Sastrawi
Successfully installed Sastrawi-1.0.1

[ ] from Sastrawi.SastrawiRemover.SastrawiRemoverFactory import SastrawiRemoverFactory

# Inisialisasi Sastrawi SastrawiRemover
sastrawi_remover = SastrawiRemoverFactory().create_sastrawi_remover()

# Tokenisasi
def tokenisasi(text):
    # Hapus stopwords
    text = sastrawi_remover.remove(text)
    # Tokenisasi (split berdasarkan spasi)
    tokens = text.split()
    return tokens

tweet_kredivo['token'] = tweet_kredivo['clean'].apply(tokenisasi)
```

```
[ ] tweet_kredivo.iloc[:, [3, 16, 17]].head(10)
```

	full_text	clean	token
0	HELP RT(tlh sewa buzzer jki luga) - a thr...	help rtfhah sewa buzzer jki luga thread ta...	help rtfhah sewa buzzer jki luga th...
1	Terimakasih om agus om koko kredivo alukaku S...	terimakasih om agus om koko kredivo alukaku se...	terimakasih om agus om koko kredivo alu...
2	- Spjnjam : 9 juta - Alukaku : 3 juta - Kredit...	spjnjam juta alukaku juta kredit pintar juta a...	spjnjam juta alukaku juta kredit pintar...
3	-star koreo YEAH YEAH yang bikin canduuuu http...	star koreo yeah yeah yang bikin canduu...	star koreo yeah yeah yang bikin canduu...
4	Pamer aman bayar tagihan kredivo 100k dapat vo...	pamer aman bayar tagihan kredivo dapat voucher...	pamer aman bayar tagihan kredivo vouch...
5	Tadi iseng beli paket telkomsel disambungi ke...	tadi iseng beli paket telkomsel disambungi ke...	tadi iseng beli paket telkomsel disambun...
6	Buat kamu pecinta Brand H&amp;M ada kabar gemb...	buat kamu pecinta brand hamp ada kabar gemb...	buat kamu pecinta brand hamp ada kabar gem...
7	PENCARAAN HARI III @jasa_pencarianpaylaterby...	pencarian hari iii open selapp hari pencarian ...	pencarian hari open hari pencarian payl...
8	Belanja pake Kredivo di SELURUH Brand MAP disk...	belanja pake kredivo di seluruh brand map disk...	belanja pake kredivo di seluruh brand map...
9	Belanja kebutuhan bulanan ngga perlu masuk daf...	belanja kebutuhan bulanan ngga perlu masuk daf...	belanja kebutuhan bulanan ngga perlu mas...

Gambar 10. Hasil proses Tokenizing Memisahkan teks menjadi token-token individual berdasarkan spasi.

```
[ ] import nltk
nltk.download('stopwords')
from nltk.corpus import stopwords

stopword = stopwords.words('indonesian')
txt_stopword = pd.read_csv('stopwords-id.txt', names=['stopwords'], header=None)

stopword.extend(["kak","ya","si","mahaka","nyak","deak","ukak","gag","gag","sar","lar","melih","ga","nappi","susuak","lila","tuh","rak","kisar","lidi","gagak","gus","guk","ct","lon","kok","guk","hest","sabr","sor","huh","linglung","pau","pau","saur","nggak","susu","kaku","om"])

stopword.extend(txt_stopword["stopwords"].split(" "))
stopword = set(stopword)

def stopword(text):
    text = [word for word in text if word not in stopword]
    return text

tweet_kredivo['stopword'] = tweet_kredivo['token'].apply(stopword)
tweet_kredivo.iloc[:, [16, 17, 18]].head(10)
```

```
[ ] [nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] package stopwords is already up-to-date!
```

	full_text	clean	token	stopwords
0	HELP RT(tlh sewa buzzer jki luga) - a thr...	help rtfhah sewa buzzer jki luga thread ta...	help rtfhah sewa buzzer jki luga th...	
1	Terimakasih om agus om koko kredivo alukaku S...	terimakasih om agus om koko kredivo alu...	terimakasih om agus om koko kredivo alu...	
2	spjnjam juta alukaku juta kredit pintar juta a...	spjnjam juta alukaku juta kredit pintar...	spjnjam juta alukaku juta kredit pintar...	
3	star koreo yeah yeah yang bikin canduu...	star koreo yeah yeah yang bikin canduu...	star koreo yeah yeah yang bikin canduu...	
4	pamer aman bayar tagihan kredivo dapat voucher...	pamer aman bayar tagihan kredivo dapat voucher...	pamer aman bayar tagihan kredivo vouch...	
5	tadi iseng beli paket telkomsel disambungi ke...	tadi iseng beli paket telkomsel disambungi ke...	tadi iseng beli paket telkomsel disambun...	
6	buat kamu pecinta brand hamp ada kabar gemb...	buat kamu pecinta brand hamp ada kabar gemb...	peceinta brand hamp kabar gembira katar gem...	
7	pencarian hari iii open selapp hari pencarian ...	pencarian hari iii open selapp hari pen...	pencarian open pencarian paylater i lin...	
8	belanja pake kredivo di seluruh brand map disk...	belanja pake kredivo di seluruh brand map...	belanja pake kredivo brand map diskom...	
9	belanja kebutuhan bulanan ngga perlu masuk daf...	belanja kebutuhan bulanan ngga perlu mas...	belanja kebutuhan bulanan ngga masuk daf...	

Gambar 11. Hasil proses Stopword Removal

Proses penghapusan stopwords diawali dengan perluasan daftar stopwords baku untuk



Namun hasil dari proses diatas masih dianggap terlalu banyak tweet yang tidak relevan sehingga penulis menyeleksi data tweet dengan cara manual. Berikut adalah proses menyeleksi data secara manual.

```

hum_tweet = len(tweet_kredivo)
print(f"Jumlah dataset adalah {num_tweet} data")

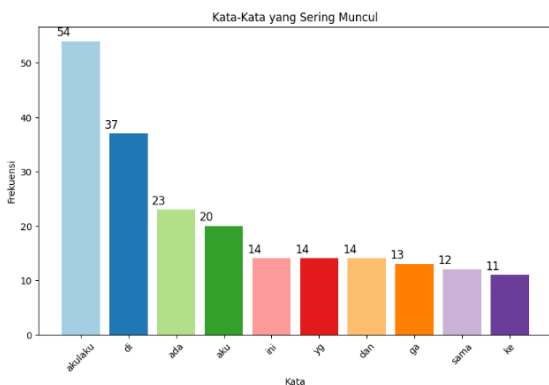
[133] tweet_kredivo.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 68 entries, 0 to 67
Data columns (total 17 columns):
 #   Column              Non-Null Count  Dtype
---  -
 0   conversation_id_str  67 non-null     object
 1   created_at          67 non-null     object
 2   favorite_count      67 non-null     float64
 3   full_text           67 non-null     object
 4   sentiment           67 non-null     float64
 5   id_str              67 non-null     object
 6   image_url           11 non-null     object
 7   in_reply_to_screen_name 4 non-null     object
 8   lang                67 non-null     object
 9   location            24 non-null     object
10  quote_count         67 non-null     float64
11  reply_count         67 non-null     float64
12  retweet_count       67 non-null     float64
13  tweet_url           67 non-null     object
14  user_id_str         67 non-null     object
15  username            67 non-null     object
16  language            63 non-null     object
dtypes: float64(5), object(12)
memory usage: 9.2+ KB

```

Gambar 15. Data seleksi manual kredivo

Setelah data diseleksi secara manual data yang didapat menjadi sebanyak 68 data yang benar mengandung sentimen.



Gambar 16. Visualisasi data seleksi manual

Visualisasi data setelah melewati preprocessing data, sehingga data yang ditampilkan diharapkan akan berisi akar kata, dan kata paling banyak muncul tetap aplikasi “Akulaku”.

Untuk data kredivo setelah melewati seleksi manual dan rangkaian preprocessing data yang diperoleh adalah sebanyak 95 data, yang dinilai penulis sebagai data yang mengandung sentimen dari masyarakat di twitter.

Tabel 6. Hasil perbandingan data preprocessing

Data	Hasil dari preprocessing
Data kredivo Nov 23 – Nov 24	Data sebelum di preprocessing sebanyak 1515 data namun setelah melewati preprocessing menjadi 1424 data
Data akulaku Nov 23 – Nov 24	Data sebelum di preprocessing sebanyak 1759 data namun setelah melewati preprocessing menjadi 1390 data
Data kredivo Nov 23 – Nov 24	Data setelah melewati seleksi manual sebelum dan setelah di preprocessing berjumlah sebanyak 95 data
Data akulaku Nov 23 – Nov 24	Data setelah melewati seleksi manual sebelum di preprocessing sebanyak 68 data namun setelah melewati preprocessing menjadi 67 data

Data yang digunakan dalam penelitian ini dikumpulkan pada periode November 23 - November 24 untuk kedua aplikasi, Kredivo dan Akulaku. Untuk data Kredivo, jumlah awal sebelum preprocessing adalah 1515 data yang kemudian berkurang menjadi 1424 data setelah melalui tahap preprocessing. Sedangkan untuk data Akulaku, dari 1759 data awal, tersisa 1390 data setelah preprocessing. Setelah dilakukan seleksi manual untuk memastikan relevansi data, jumlah data Kredivo yang tersisa adalah 95 data, sementara data Akulaku mengalami penyusutan dari 68 data menjadi 67 data setelah melalui tahap preprocessing. Penurunan jumlah data yang signifikan ini menunjukkan ketatnya proses seleksi manual yang dilakukan untuk memastikan bahwa data yang digunakan benar-benar relevan dengan tujuan penelitian.

Tabel 7. Hasil perbandingan pelabelan

Data	Hasil dari pelabelan
Data kredivo Nov 23 – Nov 24	Data dari setelah melewati preprocessing adalah 1424 data, dengan hasil pelabelan antara lain, 609 sentimen

negatif, 581 sentimen netral, dan 234 sentimen positif.

Data akulaku Nov 23 – Nov 24 Data dari setelah melewati preprocessing adalah 1390 data, dengan hasil pelabelan antara lain, 569 sentimen negatif, 613 sentimen netral, dan 208 sentimen positif.

Data kredivo Nov 23 – Nov 24 Data dari setelah melewati pelabelan manual data jumlah data sebanyak 95 tweet, dengan 47 tweet bersifat sentimen negatif, 27 tweet sentimen netral dan 21 tweet sentimen positif.

Data akulaku Nov 23 – Nov 24 Data dari setelah melewati pelabelan manual data jumlah data sebanyak 67 tweet, dengan 33 tweet bersifat sentimen negatif, 26 tweet sentimen netral dan 8 tweet sentimen positif

Data hasil pelabelan sentimen untuk kedua aplikasi menunjukkan pola yang menarik, baik sebelum maupun setelah seleksi manual. Untuk Kredivo, dari 1424 data setelah preprocessing, ditemukan dominasi sentimen negatif dengan 609 tweet, diikuti sentimen netral 581 tweet, dan positif 234 tweet. Sementara Akulaku dari 1390 data menunjukkan distribusi yang sedikit berbeda dengan dominasi sentimen netral sebanyak 613 tweet, diikuti negatif 569 tweet, dan positif 208 tweet. Setelah dilakukan seleksi manual, pola sentimen tetap konsisten di mana sentimen negatif mendominasi. Pada data Kredivo yang tersisa 95 tweet, terdapat 47 sentimen negatif, 27 netral, dan 21 positif. Sedangkan Akulaku dengan 67 tweet yang tersisa menunjukkan 33 sentimen negatif, 26 netral, dan 8 positif. Hal ini mengindikasikan bahwa baik sebelum maupun sesudah seleksi manual, kedua aplikasi cenderung mendapatkan sentimen negatif dari pengguna Twitter, dengan Kredivo memiliki proporsi sentimen positif yang relatif lebih baik dibandingkan Akulaku.

```
[3] X_train, X_test, y_train, y_test = train_test_split(df['Normalisasi'], df['Sentimen'], test_size=0.2, random_state=42)
```

Gambar 17. Proses pembagian data latihan dan uji

Dalam proses pembagian dataset, digunakan fungsi `train_test_split` dari `scikit-learn` yang membagi data menjadi dua bagian: data training

dan testing. Dengan parameter `test_size=0.2`, (80%) untuk training (20%) untuk testing. Parameter `random_state=42` digunakan untuk memastikan konsistensi hasil pembagian data setiap kali kode dijalankan.

```
#Cetak nama fitur dari TF-IDF Vectorizer
feature_names = tfidf_vectorizer.get_feature_names_out()

# Cetak beberapa contoh hasil transformasi dari data pelatihan print("hasil transformasi dari data pelatihan:")
print(X_train_tfidf)

# Cetak beberapa contoh fitur dari vektor TF-IDF
print("fitur dari vektor TF-IDF:")
print(feature_names.head())

(0, 941) 0.35378492946547907
(0, 1070) 0.3445840893958168
(0, 1220) 0.1955589425593149
(0, 52) 0.32320485482504844
(0, 1318) 0.4193587165348815
(0, 3039) 0.4193587165348815
(0, 236) 0.3779862632184866
(0, 2116) 0.3445840893958168
(1, 1220) 0.2515108898616352
(1, 129) 0.4615388383715856
(1, 2761) 0.4615388383715856
(1, 1491) 0.515883462796963
(1, 1444) 0.49318273284748626
(2, 2782) 0.5288415973868985
(2, 3077) 0.5288415973868985
```

Gambar. 18. Tahap pembobotan setiap kata dengan TF-IDF

Langkah selanjutnya adalah melakukan pembobotan pada setiap kosakata di dalam dokumen dengan menggunakan metode TF-IDF (Term Frequency - Inverse Document Frequency), yang menghitung frekuensi token pada setiap dokumen.

```
# Inisialisasi model SVM
svm_model = SVC(kernel='linear', random_state=42)

#Melatih model dengan data pelatihan yang sudah di-vektorisasi
svm_model.fit(X_train_tfidf, y_train)

#Cetak parameter model
print("Parameter model SVM:")
print(f"Kernel: {svm_model.kernel}")
print(f"C: {svm_model.C}")
print(f"Intercept: {svm_model.intercept_}")
print(f"Support Vectors: {svm_model.support_vectors_}")

Parameter model SVM:
Kernel: linear
C: 1.0
Intercept: [-0.93588693 -0.32312629 0.88997316]
Support Vectors: (0, 8) 0.1537337559079804
(0, 87) 0.14762441308128219
(0, 112) 0.1184741962582615
(0, 226) 0.1727111080671815
(0, 238) 0.11255432499647464
(0, 487) 0.12253771809546286
(0, 533) 0.19168846022630262
(0, 534) 0.19168846022630262
(0, 670) 0.19168846022630262
(0, 708) 0.19168846022630262
```

Gambar 19. Inisialisasi model SVM

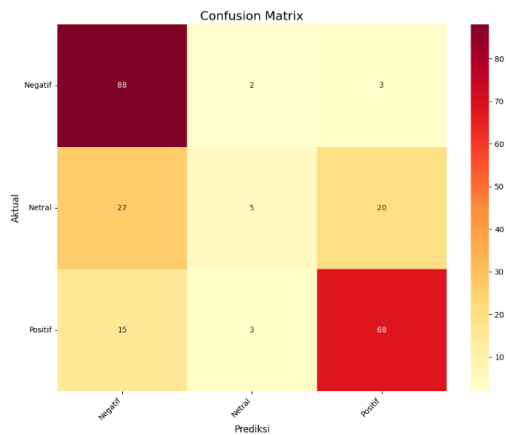
Gambar diatas menunjukkan metode yang digunakan untuk membangun model SVM (Support Vector Machine) dengan kernel linear, serta output yang menampilkan parameter dan hasil dari model SVM yang telah dilatih.

```
# Output classification report
print(classification_report(y_test_encoded, y_pred_encoded))
```

	precision	recall	f1-score	support
0	0.61	0.77	0.68	115
1	0.75	0.68	0.71	125
2	0.70	0.42	0.53	45
accuracy			0.68	285
macro avg	0.69	0.63	0.64	285
weighted avg	0.69	0.68	0.67	285

Gambar. 20. Laporan dari model SVM

Gambar tersebut adalah laporan model klasifikasi. Fungsi ini membandingkan nilai sebenarnya ( $y_{test}$ ) dengan prediksi model ( $y_{pred}$ ) dan menghasilkan berbagai metrik evaluasi.



Gambar 21. Hasil Confusion Matrix

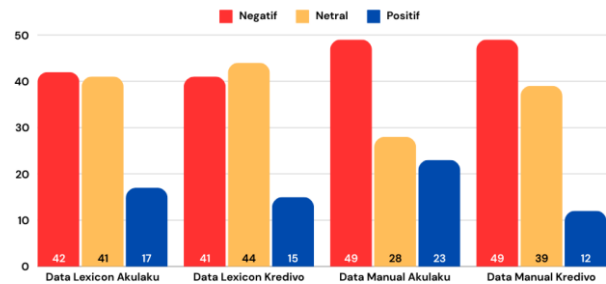
Secara keseluruhan, model klasifikasi ini memiliki akurasi yang cukup baik dalam memprediksi kelas "Negatif" dan "Netral", tetapi memiliki kesulitan dalam mengklasifikasikan kelas "Positif" dengan tepat, model menunjukkan performa yang paling rendah. Hanya 15 sampel yang diklasifikasikan dengan benar sebagai "Positif". Model juga salah mengklasifikasikan 3 sampel "Prediksi" sebagai "Positif" dan 68 sampel "Positif" sebagai "Prediksi".

Tabel 8. Hasil perbandingan hasil akurasi seleksi manual

Data	Hasil dari pelabelan
Data kredivo Nov 23 – Nov 24 Dengan library	Hasil dari pembangunan data uji dan data latih dengan SVM berdasarkan data kredivo memiliki hasil akurasi 68%, precision 69%, recall 68% dan f1score 67%.
Data akulaku Nov 23 – Nov 24 Dengan library	Hasil dari pembangunan data uji dan data latih dengan SVM berdasarkan data akulaku memiliki hasil akurasi 63%, precision 63%, recall 63% dan f1score 61%.
Data kredivo Nov 23 – Nov 24 Seleksi manual	Hasil dari pembangunan data uji dan data latih dengan SVM berdasarkan data kredivo yang berjumlah 95 data memiliki hasil akurasi 68%, precision 58%, recall 68% dan

Data akulaku Nov 23 – Nov 24  
Seleksi manual

f1score 62%.  
Hasil dari pembangunan data uji dan data latih dengan SVM berdasarkan data akulaku yang berjumlah 67 data memiliki hasil akurasi 36%, precision 41%, recall 36% dan f1score 37%.



Gambar 22. Persentase data tweet

Berdasarkan hasil perbandingan distribusi sentimen antara data Lexicon dan Manual untuk Akulaku dan Kredivo, terlihat beberapa pola penting. Pada analisis Lexicon, Akulaku mendapatkan 42% sentimen negatif, 41% netral, dan 17% positif, sementara Kredivo memperoleh 44% netral, 41% negatif, dan 15% positif. Pada analisis Manual, kedua aplikasi menunjukkan dominasi sentimen negatif yang sama sebesar 49%, namun dengan distribusi berbeda untuk sentimen lainnya - Akulaku memiliki 28% netral dan 23% positif, sedangkan Kredivo memiliki 39% netral dan 12% positif. Secara keseluruhan, sentimen negatif mendominasi di semua kategori, sementara sentimen positif konsisten berada pada proporsi terendah.

Tabel 9. Hasil perbandingan metode SVM lain

Metode	Data	Hasil akurasi
SVM linear	Kredivo	68%
	Akulaku	37%
RBF (Radial Basis Function)	Kredivo	58%
	Akulaku	57%
Polynomial	Kredivo	58%
	Akulaku	57%
Sigmoid	Kredivo	68%
	Akulaku	43%
Multi-class SVM dengan One-vs-Rest	Kredivo	74%
	Akulaku	36%

Hasil perbandingan akurasi berbagai metode SVM pada data Kredivo dan Akulaku menunjukkan variasi yang signifikan. Untuk

Kredivo, Multi-class SVM One-vs-Rest mencapai akurasi tertinggi sebesar 74%, diikuti oleh SVM Linear dan Sigmoid (masing-masing 68%), serta RBF dan Polynomial (masing-masing 58%). Sementara untuk Akulaku, metode RBF dan Polynomial memberikan akurasi terbaik sebesar 57%, diikuti Sigmoid (43%), SVM Linear (37%), dan Multi-class SVM One-vs-Rest dengan akurasi terendah 36%.

## 5. PENUTUP

### 5.1. Kesimpulan

Berdasarkan penelitian analisis sentimen terhadap aplikasi pinjaman online Kredivo dan Akulaku menggunakan metode Support Vector Machine (SVM), dapat ditarik kesimpulan, penelitian analisis sentimen terhadap aplikasi pinjaman online Kredivo dan Akulaku menggunakan metode Support Vector Machine (SVM) menghasilkan beberapa temuan penting. Dari proses crawling data, terkumpul 1515 tweet untuk Kredivo dan 1759 tweet untuk Akulaku, yang kemudian berkurang menjadi 1424 dan 1390 tweet setelah preprocessing. Hasil analisis menggunakan metode Lexicon menunjukkan dominasi sentimen negatif pada kedua aplikasi, dengan Kredivo memiliki 473 tweet negatif dan Akulaku 569 tweet negatif. Hal ini juga tercermin dalam hasil pelabelan manual, di mana dari 95 tweet Kredivo, 47 di antaranya bernada negatif, sementara dari 67 tweet Akulaku, 33 bersifat negatif. Dalam hal performa model, SVM mencapai akurasi 68% untuk Kredivo dan 63% untuk Akulaku pada dataset umum, namun pada dataset pelabelan manual, akurasi Kredivo tetap di 68% sedangkan Akulaku turun menjadi 36%. Meskipun proses seleksi dan pelabelan manual berhasil meningkatkan kualitas dataset, pengurangan jumlah data yang signifikan mempengaruhi performa model, terutama dalam mengklasifikasikan sentimen positif.

### 5.2. Saran

Untuk pengembangan penelitian selanjutnya, terdapat beberapa saran penting yang perlu dipertimbangkan. Pertama, perlu adanya peningkatan akurasi dalam fitur filter pada library Tweet Harvest Python, karena saat ini masih menghasilkan banyak data yang kurang relevan dengan topik penelitian. Kedua,

disarankan untuk memperluas sumber pengambilan data ke platform media sosial lain seperti Instagram atau Tiktok untuk mendapatkan perspektif yang lebih beragam. Terakhir, kamus yang digunakan dalam pelabelan lexicon sebaiknya diperluas dan disesuaikan dengan bahasa gaul atau informal yang umum digunakan di media sosial, sehingga dapat meningkatkan akurasi dalam pelabelan sentimen.

## 6. DAFTAR PUSTAKA

- [1] B. Rahardjo, K. Ikhwan, dan A. K. Siharis, "Pengaruh Financial Technology (Fintech) Terhadap Perkembangan Umkm Di Kota Magelang," *Prosiding Seminar Nasional Fakultas Ekonomi Untidar 2019*, no. 0, Art. no. 0, 2019, Diakses: 20 November 2023. [Daring]. Tersedia pada: <https://jurnal.untidar.ac.id/index.php/semnasfe/article/view/2093>
- [2] D. M. Maulida, "Pandangan Ekonomi Islam terhadap Sikap Konsumerisme Akibat Metode Pembayaran Tunda Bayar (PayLater)," *Jurnal Transformatif (Islamic Studies)*, vol. 5, no. 2, Art. no. 2, Okt 2021, doi: 10.23971/tf.v5i2.2980.
- [3] R. Wati dan S. Ernawati, "Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python," *Jurnal Teknik Informatika UNIKA Santo Thomas*, hlm. 240–247, Nov 2021, doi: 10.54367/jtiust.v6i2.1465.
- [4] L. Ardiani, H. Sujaini, dan T. Tursina, "Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak," *justin*, vol. 8, no. 2, hlm. 183, Apr 2020, doi: 10.26418/justin.v8i2.36776.
- [5] D. Oktavia, Y. R. Ramadhan, dan M. Minarto, "Analisis Sentimen Terhadap Penerapan Sistem E-Tilang Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 1, Art. no. 1, Agu 2023, doi: 10.30865/klik.v4i1.1040.

- [6] S. Styawati, N. Hendrastuty, dan A. R. Isnain, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 6, no. 3, Art. no. 3, Okt 2021, doi: 10.30591/jpit.v6i3.2870.
- [7] F. J. Wahidna dan P. Nerisafitra, "Analisis Sentimen Pengguna Sistem Pay Later Menggunakan Support Vector Machine Metode Pembobotan Lexicon," *JINACS*, hlm. 334–343, Feb 2023, doi: 10.26740/jinacs.v4n03.p334-343.
- [8] I. S. K. Idris, Y. A. Mustofa, dan I. A. Salihi, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM)," *Jambura Journal of Electrical and Electronics Engineering*, vol. 5, no. 1, Art. no. 1, Jan 2023, doi: 10.37905/jjee.v5i1.16830.
- [9] I. Pakpahan dan J. Pardede, "Analisis Sentimen Penanganan Covid-19 Menggunakan Metode Long Short-Term Memory Pada Media Sosial Twitter," *Jurnal Publikasi Teknik Informatika*, vol. 2, no. 1, Art. no. 1, Jan 2023, doi: 10.55606/jupti.v1i1.767.
- [10] V. R. Prasetyo, N. Benarkah, dan V. J. Chrisintha, "Implementasi Natural Language Processing Dalam Pembuatan Chatbot Pada Program Information Technology Universitas Surabaya," *Teknika*, vol. 10, no. 2, Art. no. 2, Jul 2021, doi: 10.34148/teknika.v10i2.370.
- [11] H. A. R. Harpizon, R. Kurniawan, I. Iskandar, R. Salambue, E. Budianita, dan F. Syafria, "Analisis Sentimen Komentar Di YouTube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naïve Bayes," *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, vol. 5, no. 1, Art. no. 1, Feb 2022, doi: 10.32672/jnkti.v5i1.4008.
- [12] B. P. Zen, D. Wicaksana, dan H. Alfidzar, "ANALISIS SENTIMEN TWEET VAKSIN COVID 19 SINO VAC MENGGUNAKAN METODE SUPPORT VECTOR MACHINE," *JDMSI*, vol. 3, no. 2, hlm. 21, Agu 2022, doi: 10.33365/jdmsi.v3i2.1926.
- [13] E. D. N. Sari dan I. Irhamah, "Analisis Sentimen Nasabah pada Layanan Perbankan Menggunakan Metode Regresi Logistik Biner, Naïve Bayes Classifier (NBC), dan Support Vector Machine (SVM)," *JSSITS*, vol. 8, no. 2, hlm. D177–D184, Feb 2020, doi: 10.12962/j23373520.v8i2.44565.
- [14] M. R. A. Nasution dan M. Hayaty, "Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter," *Jurnal Informatika*, vol. 6, no. 2, Art. no. 2, Sep 2019, doi: 10.31294/ji.v6i2.5129.
- [15] S. Wahyuni dan N. Wisnuri, "Penerapan Metode Klasifikasi Support Vector Machine (SVM) Untuk Menentukan Karyawan Putus Kontrak Pada PT. Tae Hang Indonesia," *Jurnal Informatika SIMANTIK*, vol. 4, no. 2, Art. no. 2, Sep 2019.