



# Analisis Sentimen Bull dan Bear Market Bitcoin Pada Komentar YouTube Menggunakan Algoritma Support Vector Machine

Hardi Wirkan\*, Andri Firmansyah

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Pelita Bangsa, Cikarang, Indonesia

Email: <sup>1,\*</sup>wirkanhardi@gmail.com, <sup>2</sup>andrifirmansyah@pelitabangsa.ac.id

Email Penulis Korespondensi: wirkanhardi@gmail.com

**Abstrak**—Perkembangan pasar cryptocurrency, khususnya Bitcoin, memengaruhi opini publik yang banyak disampaikan melalui media sosial seperti YouTube. Penelitian ini bertujuan untuk menganalisis sentimen bull dan bear market Bitcoin pada komentar YouTube menggunakan algoritma Support Vector Machine (SVM). Data dikumpulkan melalui YouTube Data API sebanyak 25.000 komentar, kemudian dilakukan tahap preprocessing meliputi case folding, cleaning, tokenizing, normalisasi kata, dan stopword removal sehingga diperoleh 8.991 data yang layak digunakan. Dataset dibagi menjadi data training sebanyak 7.192 (80%) dan data testing sebanyak 1.799 (20%). Proses pembobotan menggunakan TF-IDF dan klasifikasi dilakukan dengan algoritma SVM. Hasil pengujian menunjukkan model memperoleh akurasi sebesar 98%, dengan macro F1-score sebesar 0,91 dan weighted F1-score sebesar 0,98. Distribusi sentimen pada data testing menunjukkan sentimen netral sebanyak 50,42% (907 komentar), sentimen positif sebesar 47,08% (847 komentar), dan sentimen negatif sebesar 2,50% (45 komentar). Dominasi sentimen netral dan positif menunjukkan kecenderungan opini publik yang relatif stabil dan optimis, sejalan dengan kondisi pasar crypto pada tahun 2024 yang mengalami tren kenaikan. Penelitian ini berkontribusi dalam menyediakan pendekatan analisis sentimen berbasis komentar YouTube untuk menggambarkan kecenderungan opini publik terhadap kondisi bull dan bear market Bitcoin, sekaligus memberikan bukti empiris mengenai efektivitas algoritma Support Vector Machine dalam mengklasifikasikan sentimen pada data media sosial berbasis teks berdimensi tinggi. Selain itu, hasil penelitian memberikan gambaran mengenai keterkaitan kecenderungan opini publik dengan kondisi pasar cryptocurrency sehingga dapat digunakan sebagai referensi dalam memahami psikologi pasar (market sentiment) berbasis media sosial. Penelitian ini membuktikan bahwa algoritma SVM efektif dalam mengklasifikasikan sentimen komentar YouTube terkait kondisi pasar Bitcoin.

**Kata Kunci:** Analisis Sentimen; Bitcoin; Support Vector Machine; YouTube; Bull Market

**Abstract**—The development of the cryptocurrency market, particularly Bitcoin, significantly influences public opinion expressed through social media platforms such as YouTube. This study aims to analyze bull and bear market sentiments in YouTube comments about Bitcoin using the Support Vector Machine (SVM) algorithm. A total of 25,000 comments were collected using the YouTube Data API. After preprocessing stages including case folding, cleaning, tokenizing, word normalization, and stopword removal, 8,991 valid data were obtained. The dataset was divided into 7,192 training data (80%) and 1,799 testing data (20%). TF-IDF weighting was applied before classification using the SVM algorithm. The evaluation results show that the model achieved an accuracy of 98%, with a macro F1-score of 0.91 and a weighted F1-score of 0.98. Sentiment distribution in the testing data indicates 50.42% neutral (907 comments), 47.08% positive (847 comments), and 2.50% negative (45 comments). The dominance of neutral and positive sentiments reflects relatively stable and optimistic public opinion, consistent with the upward trend of the cryptocurrency market in 2024. This study contributes by providing a YouTube comment-based sentiment analysis approach to describe public opinion tendencies toward Bitcoin bull and bear market conditions, while also offering empirical evidence regarding the effectiveness of the Support Vector Machine algorithm in classifying sentiment within high-dimensional text-based social media data. Furthermore, the findings provide insights into the relationship between public opinion tendencies and cryptocurrency market conditions, which may serve as a reference for understanding market psychology through social media-based sentiment analysis. This study demonstrates that the SVM algorithm is effective in classifying YouTube comment sentiments related to Bitcoin market conditions.

**Keywords:** Sentiment Analysis; Bitcoin; Support Vector Machine; YouTube; Bull Market

## 1. PENDAHULUAN

Perkembangan teknologi finansial telah membawa berbagai inovasi yang berdampak besar terhadap sistem ekonomi global, salah satunya adalah munculnya mata uang kripto. Bitcoin merupakan salah satu jenis mata uang kripto yang saat ini memiliki tingkat popularitas paling tinggi. Mata uang ini bersifat digital dan tidak dikendalikan oleh otoritas pusat (terdesentralisasi). Bitcoin pertama kali diperkenalkan pada tahun 2009 oleh individu atau kelompok yang menggunakan nama samaran Satoshi Nakamoto. Dalam operasionalnya, Bitcoin memanfaatkan teknologi *blockchain* sebagai sistem pencatatan transaksi yang memungkinkan proses transaksi dilakukan dengan biaya relatif rendah (Bitcoin, 2024). Bitcoin dikenal sebagai mata uang kripto pertama sekaligus yang paling banyak digunakan secara global, dengan karakteristik pergerakan harga yang sangat fluktuatif. Perubahan harga tersebut umumnya membentuk dua kondisi pasar utama, yaitu *bull market* dan *bear market*. *Bull market* menggambarkan fase ketika harga Bitcoin mengalami peningkatan yang signifikan dan berlangsung dalam periode tertentu, bahkan dapat mencapai ratusan hingga ribuan persen. Kondisi ini biasanya dipengaruhi oleh meningkatnya optimisme investor, pertumbuhan permintaan, serta faktor tertentu seperti peristiwa *halving*. Sebaliknya, *bear market* merupakan kondisi ketika harga mengalami penurunan yang cukup tajam, umumnya berkisar antara 70% hingga 80% dari harga tertinggi, yang disebabkan oleh sentimen pasar yang negatif, meningkatnya tekanan jual, serta adanya koreksi setelah periode kenaikan yang tinggi (Nur, 2026). Kedua kondisi tersebut tidak hanya memengaruhi perilaku investor, tetapi juga bisa menimbulkan kerugian yang cukup besar serta membentuk persepsi dan opini publik yang banyak disampaikan melalui media sosial.



Sebagai salah satu media sosial yang memiliki jumlah pengguna besar, YouTube berperan sebagai platform berbasis video yang banyak dimanfaatkan dalam penyebaran berbagai informasi, termasuk yang berkaitan dengan ekonomi dan keuangan digital seperti *cryptocurrency*. Melalui berbagai kanal yang tersedia, para kreator konten menyajikan pembahasan mengenai pergerakan harga Bitcoin, analisis kondisi pasar, serta potensi keuntungan dan risiko investasi dalam bentuk video yang mudah diakses oleh masyarakat luas. Hal ini menjadikan YouTube sebagai salah satu sumber informasi penting yang berkontribusi dalam membentuk pemahaman dan persepsi publik terhadap dinamika *bull market* dan *bear market* pada Bitcoin (Suharsono & Nurahman, 2024).

Ketika pasar kripto berada dalam kondisi *bull market* maupun *bear market*, perhatian publik terhadap pergerakan Bitcoin cenderung meningkat, termasuk dari kalangan influencer yang aktif membahas topik tersebut. Dalam menyampaikan pandangan terkait fluktuasi harga Bitcoin, YouTube menyediakan fitur komentar yang memungkinkan audiens untuk menyampaikan berbagai tanggapan, seperti opini, kekhawatiran, maupun dukungan terhadap konten yang disajikan. Isi komentar tersebut umumnya mengandung beragam sentimen, baik positif, negatif, maupun netral, yang mencerminkan respons terhadap kondisi pasar, misalnya saat terjadi kenaikan harga yang signifikan (*bull market*) atau penurunan yang tajam (*bear market*). Maka dari itu, kolom komentar pada video YouTube dapat dimanfaatkan sebagai representasi dari kondisi psikologis pasar serta persepsi masyarakat terhadap Bitcoin (Tri Ayu Mareta et al., 2025).

Oleh karena itu, analisis sentimen terhadap komentar pengguna di YouTube dapat memberikan gambaran yang lebih jelas mengenai bagaimana respon publik terhadap kondisi *bull market* maupun *bear market*. Analisis sentimen sendiri bertujuan untuk menemukan pola teks dan mengelompokkan sentimen menjadi positif, negatif, atau netral melalui tahapan preprocessing, pembobotan, dan klasifikasi. Metode *Support Vector Machine* (SVM) sangat cocok untuk data teks dan analisis sentimen karena mampu memisahkan kelas sentimen (positif, negatif, netral), metode ini bekerja berdasarkan prinsip mencari hyperplane optimal yang memisahkan kelas data dengan margin terbesar, sehingga efektif untuk data *non-linear* melalui *kernel trick*. SVM dikenal memiliki keunggulan dalam hal akurasi tinggi pada dataset kecil hingga menengah, kemampuan menangani data berdimensi tinggi seperti fitur TF-IDF dari teks komentar, serta *robust* terhadap *noise* dan *outlier*. Berdasarkan hal tersebut, penelitian ini bertujuan untuk menganalisis perbandingan kecenderungan sentimen publik (positif, negative, netral) pada komentar YouTube selama kondisi *bull* dan *bear market* Bitcoin dan korelasi antara hasil analisis sentimen (positif, negatif, netral) komentar YouTube terkait *bull market* dan *bear market* Bitcoin dengan kondisi pasar pada periode yang sama menggunakan metode SVM.

Kontribusi penelitian ini terletak pada pemanfaatan komentar YouTube sebagai sumber data opini publik untuk menganalisis kecenderungan sentimen terhadap kondisi *bull* dan *bear market* Bitcoin menggunakan algoritma Support Vector Machine. Penelitian ini tidak hanya berfokus pada klasifikasi sentimen positif, negatif, dan netral, tetapi juga memberikan gambaran empiris mengenai kecenderungan persepsi masyarakat terhadap dinamika pasar *cryptocurrency* berdasarkan opini yang berkembang di media sosial. Dengan demikian, penelitian ini diharapkan dapat memperkaya kajian analisis sentimen pada domain *cryptocurrency* serta menjadi referensi dalam memahami psikologi pasar berbasis data media sosial.

## 2. METODOLOGI PENELITIAN

### 2.1 Kerangka Dasar Penelitian

Penelitian ini menggunakan pendekatan *Knowledge Discovery in Database* (KDD) yang terdiri dari beberapa tahapan utama. *Knowledge Discovery in Database* (KDD) merupakan proses sistematis untuk mengekstraksi pola dan pengetahuan dari data berukuran besar. Tahapannya meliputi pemilihan data, preprocessing, transformasi data (seperti TF-IDF), penerapan teknik data mining, evaluasi, serta penyajian hasil (Putra & Putra, 2025). KDD berfungsi sebagai kerangka kerja yang komprehensif untuk mengolah data teks mentah menjadi pola sentimen yang dapat ditindaklanjuti (*actionable*). Proses ini melibatkan penerapan metode klasifikasi seperti *Support Vector Machine* (SVM) sehingga mampu menghasilkan insight mengenai hubungan antara opini publik dan pergerakan tren pasar *cryptocurrency* (Sari Siregar et al., 2024).

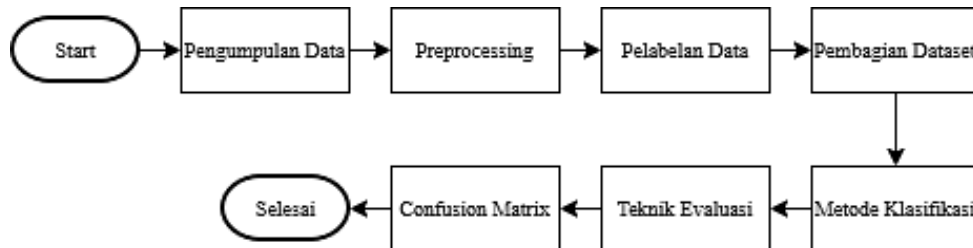
Data penelitian diperoleh melalui YouTube Data API v3 dengan teknik web *crawling* pada video yang memiliki keterkaitan dengan pembahasan Bitcoin. Jumlah data yang dikumpulkan sebanyak 25.000 komentar yang kemudian diproses untuk menghasilkan dataset yang siap digunakan dalam proses klasifikasi. Penelitian dilakukan menggunakan lingkungan pemrograman Python pada Google Colaboratory dengan bantuan beberapa library seperti *pandas*, *scikit-learn*, *NLTK*, dan *googleapiclient*.

Metode klasifikasi yang digunakan dalam penelitian ini adalah *Support Vector Machine* (SVM) karena memiliki kemampuan yang baik dalam menangani data teks berdimensi tinggi dan mampu menghasilkan tingkat akurasi yang tinggi pada analisis sentimen. Dataset dibagi menjadi data *training* sebesar 80% dan data *testing* sebesar 20%. Model SVM kemudian digunakan untuk mengklasifikasikan komentar ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral.

Evaluasi model dilakukan menggunakan *confusion matrix* dengan parameter *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengetahui performa klasifikasi yang dihasilkan. Hasil evaluasi kemudian dianalisis untuk mengetahui kecenderungan sentimen publik terhadap kondisi *bull* dan *bear market* Bitcoin berdasarkan komentar pengguna YouTube.

## 2.2 Tahapan Penelitian

Tahapan penelitian dilakukan secara sistematis untuk menganalisis sentimen publik terhadap kondisi *bull* dan *bear market* Bitcoin berdasarkan komentar pengguna YouTube menggunakan algoritma *Support Vector Machine* (SVM). Untuk memberikan gambaran sistematis mengenai tahapan penelitian yang dilakukan, alur penelitian ditunjukkan pada Gambar 1. Tahapan tersebut dimulai dari proses pengumpulan data komentar YouTube, *preprocessing* data, pelabelan sentimen, pembagian dataset, proses klasifikasi menggunakan *Support Vector Machine* (SVM), evaluasi model, hingga analisis menggunakan *confusion matrix*.



Gambar 1. Tahapan Penelitian

### 2.2.1 Pengumpulan Data

Pengumpulan data merupakan tahap awal dalam penelitian yang dilakukan untuk memperoleh informasi yang sesuai dengan kebutuhan analisis. Data yang diperoleh selanjutnya digunakan sebagai dasar dalam proses pengolahan serta dalam penarikan kesimpulan penelitian (William et al., 2022). Sumber data primer dalam penelitian ini berupa komentar pengguna pada video YouTube yang membahas Bitcoin, khususnya yang mencerminkan kondisi *bull market* (kenaikan harga signifikan) dan *bear market* (penurunan harga tajam). Data dipilih dari platform YouTube karena karakteristiknya yang kaya akan opini publik, emosi investor, serta diskusi yang berlangsung secara *real-time* terkait fluktuasi pasar kripto, sehingga sangat relevan untuk digunakan dalam analisis sentimen.

### 2.2.2 Preprocessing Data

*Text preprocessing* merupakan tahap awal dalam pengolahan data teks yang terdiri dari serangkaian proses untuk mentransformasikan teks mentah menjadi bentuk yang lebih terstruktur, sehingga dapat digunakan secara optimal dalam penerapan algoritma *text mining* maupun *machine learning* (Albab et al., 2023). Tahapan *preprocessing* bertujuan untuk membersihkan dan menyiapkan data agar dapat diproses lebih lanjut. Proses ini meliputi : *Case folding*, *Cleaning*, *Tokenizing*, Normalisasi kata, dan *Stopword removal* (Yanto et al., 2025).

### 2.2.3 Pelabelan Data

Pelabelan data merupakan tahapan pengklasifikasian data teks ke dalam kategori tertentu, seperti positif, negatif, dan netral, yang bertujuan untuk memfasilitasi algoritma dalam mengenali serta mengekstraksi pola opini dari data yang dianalisis (Rifka Alkhilyatul Ma'rifat, I Made Suraharta, 2024). Proses ini menghasilkan file Hasil\_Labelling\_Data\_3class.csv yang berisi kolom *stopword removal* (fitur teks) dan Sentiment (target label) untuk pelatihan SVM. Oleh karena itu, pelabelan data menjadi tahapan penting dalam *pipeline machine learning* karena memengaruhi kinerja model dalam melakukan klasifikasi dan prediksi pada data baru (Syahrohim et al., 2024).

### 2.2.4 Pembagian Dataset

Pembagian dataset dilakukan menggunakan metode *train-test split* dengan rasio 80:20 yang merupakan standar emas dalam *machine learning* untuk memastikan keseimbangan antara ukuran data latih yang memadai dan validasi independen. Pendekatan ini memungkinkan model SVM untuk belajar pola sentimen dari 80% data sambil menguji generalisasi pada 20% data yang belum pernah dilihat (*hold-out set*).

### 2.2.5 Metode *Support Vector Machine* (SVM)

*Support Vector Machine* (SVM) merupakan algoritma pembelajaran terawasi yang digunakan dalam proses klasifikasi dengan cara menentukan *hyperplane* optimal sebagai batas pemisah antar kelas dengan margin maksimum (Hidayat et al., 2024). Dalam penerapan pada klasifikasi teks, SVM merepresentasikan setiap dokumen, seperti komentar, ke dalam ruang fitur berdimensi tinggi menggunakan metode seperti TF-IDF, kemudian membentuk batas keputusan untuk membedakan kelas sentimen (I Putu Gede Hendra Suputra et al., 2025). Metode SVM dipilih sebagai algoritma klasifikasi multi-kelas untuk mengklasifikasikan sentimen komentar Bitcoin menjadi tiga kategori: Positif, Negatif, dan Netral. SVM optimal untuk dataset teks *high-dimensional* (*sparse matrix* dari *CountVectorizer*) karena kemampuan *maximum margin hyperplane* dan *robustness* terhadap *noise*.

### 2.2.6 Teknik Evaluasi

Evaluasi performa model *Support Vector Machine* (SVM) dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur kemampuan model dalam mengklasifikasikan sentimen ke dalam kategori positif,



negatif, dan netral. Nilai evaluasi diperoleh berdasarkan *confusion matrix* yang terdiri dari True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (Zahra et al., 2025). *Accuracy* digunakan untuk mengukur ketepatan prediksi model secara keseluruhan, *precision* untuk mengetahui ketepatan prediksi setiap kelas, *recall* untuk mengukur kemampuan model dalam menemukan data yang relevan, serta *F1-score* untuk menunjukkan keseimbangan antara *precision* dan *recall* dalam proses klasifikasi sentimen (Arimbi Puspitasari et al., 2025).

### 2.2.7 Confusion Matrix

*Confusion matrix* merupakan teknik evaluasi pada model klasifikasi yang digunakan untuk mengukur kinerja model dengan membandingkan hasil prediksi terhadap nilai aktual. Melalui metode ini, dapat diketahui tingkat akurasi serta jenis kesalahan yang terjadi selama proses klasifikasi (Suryadewiansyah & Tju, 2022). *Confusion matrix* adalah sebuah tabel yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan cara membandingkan label prediksi model dengan label aktual pada data uji. Penggunaan *confusion matrix* dalam penelitian *machine learning* memiliki peran penting karena mampu menyajikan informasi yang lebih detail mengenai kinerja model serta jenis kesalahan yang dihasilkan. Informasi ini dapat dimanfaatkan sebagai dasar dalam proses evaluasi dan perbaikan model agar menghasilkan performa yang lebih optimal (Pinaria et al., 2024).

## 3. HASIL DAN PEMBAHASAN

Pada bagian ini dijelaskan hasil penelitian yang telah dilakukan dalam menganalisis sentimen komentar pengguna YouTube terkait kondisi *bull* dan *bear market* Bitcoin menggunakan algoritma *Support Vector Machine* (SVM). Penelitian dilakukan untuk mengetahui kecenderungan opini publik terhadap pergerakan pasar *cryptocurrency* berdasarkan komentar yang diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Tahapan penelitian dimulai dari proses pengumpulan data komentar, *preprocessing*, pelabelan data sentimen, pembagian dataset menjadi data training dan data testing, proses klasifikasi menggunakan algoritma *Support Vector Machine* (SVM), teknik evaluasi model menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*, serta analisis hasil klasifikasi menggunakan *confusion matrix*.

Hasil penelitian menunjukkan bahwa algoritma *Support Vector Machine* (SVM) memiliki performa yang baik dalam proses klasifikasi sentimen komentar YouTube terkait Bitcoin. Model yang dibangun mampu mengidentifikasi berbagai bentuk opini publik terhadap kondisi pasar *cryptocurrency*, seperti sentimen optimis terhadap kenaikan harga, sentimen negatif akibat penurunan pasar, serta komentar yang bersifat netral. Hasil klasifikasi tersebut dapat memberikan gambaran mengenai persepsi masyarakat terhadap kondisi pasar Bitcoin, sehingga diharapkan dapat menjadi referensi tambahan bagi peneliti maupun pelaku investasi *cryptocurrency* dalam memahami pola sentimen publik pada media sosial.

### 3.1 Hasil Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan menggunakan YouTube Data API melalui proses web *crawling* pada platform Google Colaboratory. Data yang dikumpulkan berupa komentar pengguna pada video YouTube yang membahas kondisi *bull market* dan *bear market* Bitcoin. Pemilihan video dilakukan berdasarkan kata kunci yang berkaitan dengan Bitcoin, *cryptocurrency*, kenaikan harga pasar, serta penurunan harga pasar agar data yang diperoleh sesuai dengan topik penelitian. Berdasarkan proses *crawling* yang telah dilakukan, diperoleh sebanyak 25.000 komentar sebagai dataset awal penelitian. Data komentar yang diperoleh berasal dari berbagai pengguna YouTube dengan beragam bentuk opini dan tanggapan terhadap kondisi pasar Bitcoin. Komentar tersebut terdiri dari sentimen positif, negatif, maupun netral yang mencerminkan persepsi publik terhadap pergerakan harga *cryptocurrency*. Namun, data hasil *crawling* masih berupa teks mentah (*raw data*) yang mengandung berbagai elemen tidak relevan seperti simbol, angka, URL, emoji, karakter khusus, serta kemungkinan adanya data duplikat. Kondisi tersebut dapat mempengaruhi kualitas proses klasifikasi apabila data langsung digunakan tanpa pengolahan lebih lanjut. Oleh karena itu, diperlukan tahap *preprocessing* untuk membersihkan dan menyiapkan data sebelum digunakan dalam proses klasifikasi sentimen menggunakan algoritma *Support Vector Machine* (SVM).

Proses pengambilan data komentar YouTube menggunakan teknik web *crawling* melalui YouTube Data API ditunjukkan pada Gambar 2. Gambar tersebut memperlihatkan tahapan pengambilan data mentah (*raw data*) yang selanjutnya digunakan sebagai dataset awal penelitian. Pada tahap ini, proses *crawling* dilakukan untuk memperoleh komentar yang relevan dengan topik penelitian, khususnya yang berkaitan dengan kondisi *bull market* dan *bear market* Bitcoin. Pemanfaatan YouTube sebagai sumber data memungkinkan penelitian memperoleh opini publik yang bersifat aktual dan dinamis karena komentar yang dituliskan pengguna umumnya merepresentasikan respons langsung terhadap kondisi pasar *cryptocurrency* yang sedang berlangsung. Selain itu, pengumpulan data dalam jumlah besar juga memberikan peluang bagi model klasifikasi untuk mengenali pola sentimen secara lebih optimal sehingga hasil analisis yang diperoleh menjadi lebih representatif dalam menggambarkan kecenderungan persepsi masyarakat terhadap Bitcoin. Dengan demikian, tahapan pengumpulan data menjadi bagian penting dalam penelitian karena kualitas dan relevansi data yang diperoleh akan sangat memengaruhi proses *preprocessing*, klasifikasi sentimen, hingga hasil evaluasi model yang dihasilkan.

```
# ID video YouTube yang ingin diambil
video_ids = ['MPK1DI0eddw']
total_comments = 25000
output_filename = 'BulldanBearmarketBitcoin.csv'

# Ambil komentar dan simpan ke file CSV
comments = scrape_comments_from_videos(video_ids, total_comments, output_filename)

df = pd.DataFrame(comments)
df.info()
df.head()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 25000 entries, 0 to 24999
Data columns (total 4 columns):
#   Column          Non-Null Count  Dtype
---  ---            -
0   author           25000 non-null  object
1   comment          25000 non-null  object
2   likes            25000 non-null  int64
3   published_at     25000 non-null  object
dtypes: int64(1), object(3)
memory usage: 781.4+ KB
```

**Gambar 2.** Proses Pengambilan Data

### 3.2 Hasil Preprocessing Data

Berdasarkan proses pengambilan data menggunakan kata kunci yang berkaitan dengan topik penelitian, diperoleh sebanyak 25.000 komentar YouTube sebagai dataset awal penelitian. Data yang diperoleh masih berupa teks mentah yang mengandung berbagai karakter tidak relevan, seperti simbol, angka, URL, emoji, serta kemungkinan adanya data duplikat sehingga perlu dilakukan tahap preprocessing sebelum digunakan pada proses klasifikasi sentimen.

Tahap *preprocessing* dilakukan untuk membersihkan dan menyiapkan data agar dapat diproses oleh algoritma klasifikasi dengan lebih optimal. Proses *preprocessing* yang dilakukan meliputi *case folding* untuk mengubah seluruh huruf menjadi *lowercase*, *cleaning* untuk menghapus karakter yang tidak diperlukan, *tokenizing* untuk memisahkan kalimat menjadi kata-kata, normalisasi kata untuk mengubah kata tidak baku menjadi kata baku, serta *stopword removal* untuk menghapus kata-kata umum yang tidak memiliki pengaruh besar terhadap analisis sentimen.

Setelah seluruh proses *preprocessing* selesai dilakukan, jumlah data mengalami pengurangan yang cukup signifikan dibandingkan dataset awal. Dari total 25.000 komentar, diperoleh sebanyak 8.991 komentar yang layak digunakan dalam proses klasifikasi sentimen menggunakan algoritma *Support Vector Machine* (SVM).

**Tabel 1.** Jumlah Data Setelah *Preprocessing*

Keterangan	Jumlah Data
Data awal hasil <i>crawling</i>	25.000
Data akhir siap klasifikasi	8.991

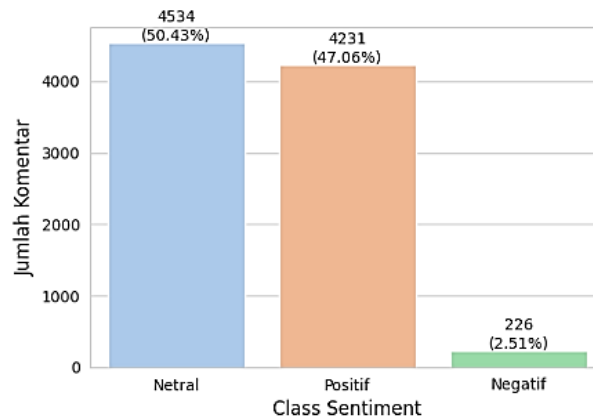
Berdasarkan Tabel 1 tersebut, terjadi penurunan jumlah data yang cukup signifikan dibandingkan dengan dataset awal. Hal ini menunjukkan bahwa terdapat sejumlah komentar yang tidak relevan, mengandung duplikasi, atau tidak memenuhi kriteria untuk digunakan dalam proses analisis sentimen.

### 3.3 Hasil Pelabelan Sentimen

Setelah tahap *preprocessing* selesai dilakukan, data kemudian melalui proses pelabelan sentimen dengan mengelompokkan komentar ke dalam tiga kategori, yaitu positif, netral, dan negatif. Proses pelabelan dilakukan untuk memberikan identitas kelas pada setiap komentar sehingga data dapat digunakan dalam proses pelatihan model klasifikasi menggunakan algoritma *Support Vector Machine* (SVM).

Distribusi sentimen pada dataset dapat dilihat pada gambar 3 di bawah ini. Berdasarkan hasil distribusi tersebut, terlihat bahwa dataset didominasi oleh sentimen netral dan positif, sedangkan jumlah sentimen negatif relatif lebih sedikit dibandingkan kedua kelas lainnya. Dominasi sentimen netral menunjukkan bahwa sebagian besar komentar pengguna YouTube cenderung bersifat informatif atau diskusi umum terkait kondisi pasar Bitcoin. Sementara itu, tingginya sentimen positif menunjukkan adanya kecenderungan optimisme masyarakat terhadap perkembangan pasar *cryptocurrency* pada periode pengambilan data penelitian.

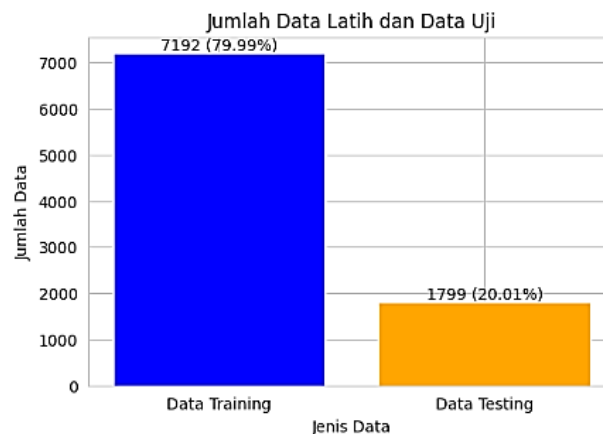
Jumlah sentimen negatif yang lebih sedikit menunjukkan bahwa komentar bernada pesimis atau kekhawatiran terhadap kondisi pasar tidak terlalu mendominasi dataset. Perbedaan jumlah data pada masing-masing kelas sentimen tersebut menunjukkan adanya kondisi ketidakseimbangan kelas (*class imbalance*) yang dapat mempengaruhi performa model klasifikasi, terutama dalam mengenali pola pada kelas minoritas seperti sentimen negatif.

**Gambar 3.** Distribusi Sentimen

### 3.4 Pembagian Data *Training* dan *Testing*

Dataset yang telah melalui tahap *preprocessing* dan pelabelan sentimen kemudian dibagi menjadi dua bagian menggunakan metode *train-test split* dengan rasio 80:20. Pembagian dataset ini dilakukan untuk memisahkan data yang digunakan pada tahap pelatihan model dan data yang digunakan pada tahap pengujian model. Proses pembagian data ditunjukkan pada Gambar 4. Sebanyak 80% data digunakan sebagai data training untuk melatih model *Support Vector Machine* (SVM) dalam mengenali pola dari masing-masing kelas sentimen, yaitu positif, netral, dan negatif. Data training digunakan agar model dapat mempelajari hubungan antar fitur teks yang telah direpresentasikan menggunakan metode TF-IDF. Sementara itu, 20% data sisanya digunakan sebagai data testing untuk menguji performa model terhadap data yang belum pernah dipelajari sebelumnya.

Pembagian dataset ini bertujuan untuk mengetahui kemampuan model dalam melakukan klasifikasi sentimen secara akurat pada data baru. Selain itu, penggunaan rasio 80:20 dipilih karena mampu memberikan keseimbangan antara jumlah data pelatihan dan data pengujian sehingga proses evaluasi model dapat dilakukan dengan lebih optimal. Dengan jumlah data training yang lebih besar, model diharapkan mampu mempelajari pola sentimen dengan lebih baik tanpa mengurangi kualitas pengujian pada data *testing*.

**Gambar 4.** Pembagian Dataset

### 3.5 Distribusi Sentimen Pada Data *Testing*

Distribusi kelas sentimen pada data *testing* ditunjukkan pada Tabel 2. Distribusi ini menggambarkan jumlah data pada masing-masing kategori sentimen yang digunakan dalam proses pengujian model klasifikasi. Berdasarkan hasil distribusi tersebut, terlihat bahwa data testing didominasi oleh sentimen netral dan positif, sedangkan jumlah sentimen negatif memiliki jumlah yang jauh lebih sedikit dibandingkan dengan kedua kelas lainnya. Dominasi sentimen netral menunjukkan bahwa sebagian besar komentar pengguna YouTube cenderung bersifat informatif, diskusi umum, atau tidak menunjukkan kecenderungan opini yang terlalu kuat terhadap kondisi pasar Bitcoin. Komentar netral umumnya berisi tanggapan mengenai informasi pasar, prediksi harga, maupun diskusi antar pengguna terkait perkembangan *cryptocurrency*. Sementara itu, tingginya jumlah sentimen positif menunjukkan adanya pandangan optimis masyarakat terhadap kondisi pasar *cryptocurrency* pada periode pengambilan data penelitian. Hal ini sejalan dengan kondisi pasar Bitcoin pada tahun 2024 yang cenderung mengalami tren kenaikan sehingga memunculkan lebih banyak komentar positif dari pengguna media sosial.

Sebaliknya, sentimen negatif memiliki jumlah yang relatif kecil sehingga menunjukkan bahwa komentar bernada pesimis atau kekhawatiran terhadap kondisi pasar tidak terlalu mendominasi dataset. Perbedaan jumlah data pada setiap



kelas sentimen tersebut menunjukkan adanya kondisi *class imbalance* pada data testing. Kondisi ini dapat mempengaruhi performa model klasifikasi, terutama dalam mengenali pola pada kelas minoritas seperti sentimen negatif, karena jumlah data yang digunakan untuk proses pembelajaran dan pengujian lebih sedikit dibandingkan kelas lainnya.

**Tabel 2.** Distribusi Sentimen Data *Testing*

Kelas Sentimen	Jumlah Data
Netral	907
Positif	847
Negatif	45
Total	1.799

### 3.6 Klasifikasi Menggunakan *Support Vector Machine* (SVM)

Pada tahap ini dilakukan proses klasifikasi sentimen menggunakan algoritma *Support Vector Machine* (SVM). Model dibangun berdasarkan fitur teks yang telah diekstraksi menggunakan metode TF-IDF sehingga setiap komentar dapat direpresentasikan dalam bentuk vektor numerik. Representasi tersebut memungkinkan algoritma SVM untuk mengenali pola kata yang berkaitan dengan masing-masing kategori sentimen, yaitu positif, netral, dan negatif. Proses ekstraksi fitur menggunakan TF-IDF juga membantu model dalam memberikan bobot terhadap kata-kata yang dianggap penting pada setiap komentar sehingga proses klasifikasi dapat dilakukan dengan lebih efektif.

Data yang telah melalui tahap *preprocessing* kemudian dibagi menjadi data *training* dan data *testing* dengan rasio 80:20. Sebanyak 7.192 data digunakan sebagai data *training* untuk melatih model dalam mengenali pola sentimen pada komentar YouTube, sedangkan 1.799 data digunakan sebagai data *testing* untuk melihat kemampuan model dalam mengklasifikasikan data baru yang belum pernah dipelajari sebelumnya. Pembagian data ini dilakukan agar model tidak hanya mampu mengenali pola pada data pelatihan, tetapi juga dapat melakukan prediksi dengan baik terhadap data baru.

Penggunaan algoritma *Support Vector Machine* (SVM) pada penelitian ini dipilih karena memiliki kemampuan yang baik dalam menangani data teks berdimensi tinggi hasil ekstraksi TF-IDF. Selain itu, SVM juga dikenal mampu menghasilkan performa klasifikasi yang tinggi pada analisis sentimen, terutama pada data teks media sosial yang memiliki karakteristik sparse dan tidak terstruktur. Dengan kemampuan membentuk *hyperplane* optimal sebagai batas pemisah antar kelas, SVM mampu mengklasifikasikan sentimen komentar YouTube secara lebih akurat dan stabil.

### 3.7 Evaluasi Model

Evaluasi model dilakukan menggunakan beberapa metrik klasifikasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengetahui performa algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen komentar YouTube. Penggunaan beberapa metrik evaluasi bertujuan agar performa model dapat dianalisis secara lebih menyeluruh, tidak hanya berdasarkan tingkat ketepatan prediksi, tetapi juga berdasarkan kemampuan model dalam mengenali masing-masing kelas sentimen.

Berdasarkan hasil pengujian dapat dilihat pada tabel 3 di bawah ini, model memperoleh tingkat akurasi sebesar 98%, yang menunjukkan bahwa sebagian besar data testing berhasil diklasifikasikan dengan benar sesuai kategori sentimennya. Tingginya nilai *accuracy* menunjukkan bahwa kombinasi tahapan *preprocessing*, pembobotan TF-IDF, dan algoritma SVM mampu menghasilkan model klasifikasi yang efektif pada data komentar YouTube terkait Bitcoin.

Nilai *weighted F1-score* yang tinggi menunjukkan bahwa model memiliki performa yang sangat baik pada kelas sentimen yang dominan, khususnya sentimen netral dan positif. Sementara itu, nilai *macro average* yang sedikit lebih rendah menunjukkan adanya pengaruh ketidakseimbangan jumlah data antar kelas (*class imbalance*), terutama karena jumlah data sentimen negatif relatif lebih sedikit dibandingkan kelas lainnya. Kondisi tersebut menyebabkan model memiliki keterbatasan dalam mengenali pola pada kelas minoritas secara optimal.

Meskipun demikian, secara keseluruhan model *Support Vector Machine* (SVM) mampu menunjukkan performa klasifikasi yang baik pada dataset komentar YouTube terkait kondisi pasar Bitcoin. Hasil evaluasi ini menunjukkan bahwa algoritma SVM efektif digunakan dalam proses analisis sentimen berbasis teks, khususnya pada data media sosial yang memiliki karakteristik tidak terstruktur dan berdimensi tinggi. Selain ditinjau dari nilai akurasi dan F1-score, hasil evaluasi model juga menunjukkan bahwa kombinasi tahapan *preprocessing*, pembobotan TF-IDF, dan algoritma *Support Vector Machine* (SVM) memberikan kontribusi yang baik terhadap kemampuan model dalam mengenali pola sentimen pada komentar pengguna YouTube. Tahapan *preprocessing* membantu mengurangi noise pada data teks sehingga komentar menjadi lebih terstruktur untuk dianalisis, sedangkan pembobotan TF-IDF memungkinkan model memberikan tingkat kepentingan yang berbeda pada setiap kata yang muncul dalam komentar. Dengan demikian, proses klasifikasi tidak hanya dipengaruhi oleh algoritma yang digunakan, tetapi juga oleh kualitas representasi fitur teks yang dihasilkan selama tahapan pengolahan data. Hasil evaluasi ini memperlihatkan bahwa model yang dibangun mampu mempertahankan performa klasifikasi yang tinggi meskipun dihadapkan pada karakteristik data media sosial yang cenderung tidak terstruktur, beragam, serta mengandung variasi bahasa yang cukup tinggi. Temuan ini sekaligus memperkuat bahwa pendekatan analisis sentimen berbasis komentar YouTube dapat digunakan sebagai salah satu

pendekatan untuk memahami kecenderungan opini publik terhadap kondisi pasar cryptocurrency, khususnya pada dinamika bull market dan bear market Bitcoin.

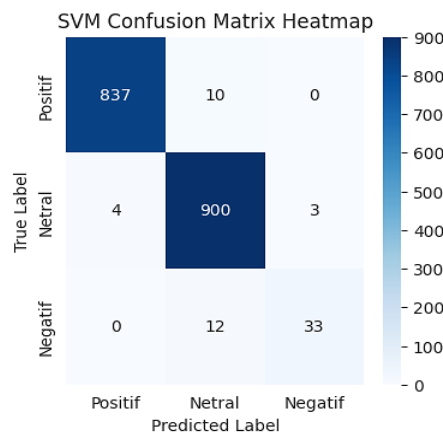
**Tabel 3.** Hasil Evaluasi Model SVM

Metrik	Nilai
<i>Accuracy</i>	98%
<i>Macro Precision</i>	0.91
<i>Macro Recall</i>	0.90
<i>Macro F1-Score</i>	0.91
<i>Weighted F1-Score</i>	0.98

### 3.8 Analisis Confusion Matrix

Berdasarkan confusion matrix pada Gambar 5, sebagian besar hasil prediksi berada pada bagian diagonal matriks yang menunjukkan bahwa model mampu melakukan klasifikasi dengan tingkat ketepatan yang tinggi. Hal tersebut menandakan bahwa sebagian besar data testing berhasil diprediksi sesuai dengan label sentimen aslinya oleh algoritma Support Vector Machine (SVM). Pada kelas sentimen positif, sebanyak 834 data berhasil diklasifikasikan dengan benar sesuai label aslinya. Selanjutnya, pada kelas sentimen netral terdapat 892 data yang berhasil dikenali dengan benar oleh model. Tingginya jumlah prediksi benar pada kedua kelas tersebut menunjukkan bahwa model mampu mempelajari pola kata dan karakteristik komentar positif maupun netral dengan baik. Dominasi jumlah data pada kedua kelas tersebut juga membantu model dalam mengenali pola sentimen secara lebih optimal selama proses pelatihan.

Sementara itu, pada kelas sentimen negatif, sebanyak 33 dari 45 data berhasil diklasifikasikan dengan benar, sedangkan 12 data lainnya salah diprediksi sebagai sentimen netral. Kesalahan klasifikasi tersebut menunjukkan bahwa beberapa komentar negatif memiliki karakteristik kata yang mirip dengan komentar netral sehingga model mengalami kesulitan dalam membedakan kedua kelas tersebut secara konsisten. Hasil tersebut menunjukkan bahwa model Support Vector Machine (SVM) memiliki performa yang sangat baik dalam mengenali kelas sentimen positif dan netral. Namun, model masih mengalami keterbatasan dalam mengidentifikasi sentimen negatif karena jumlah data pada kelas tersebut relatif lebih sedikit dibandingkan kelas lainnya. Kondisi class imbalance ini menyebabkan pola pada sentimen negatif tidak dipelajari secara optimal sehingga beberapa data negatif masih mengalami kesalahan klasifikasi.



**Gambar 5.** Confusion Matrix

### 3.9 Analisis Dan Pembahasan

Hasil penelitian menunjukkan bahwa algoritma *Support Vector Machine* (SVM) mampu mengklasifikasikan sentimen komentar YouTube terkait *bull* dan *bear market* Bitcoin dengan performa yang sangat baik, ditunjukkan oleh tingkat akurasi sebesar 98%. Berdasarkan hasil evaluasi, model memperoleh nilai *macro F1-score* sebesar 0,91 dan *weighted F1-score* sebesar 0,98, yang menunjukkan kemampuan klasifikasi yang sangat baik secara keseluruhan pada ketiga kelas sentimen.

Berdasarkan distribusi data *testing*, sentimen netral mendominasi dengan jumlah 907 komentar (50,42%), diikuti oleh sentimen positif sebanyak 847 komentar (47,08%), sedangkan sentimen negatif hanya sebesar 45 komentar (2,50%). Dominasi sentimen netral dan positif dalam dataset berkontribusi terhadap tingginya nilai *accuracy* dan *weighted F1-score* yang diperoleh model. Hal ini menunjukkan bahwa model SVM sangat baik dalam mengenali pola pada kelas mayoritas.

Distribusi tersebut menunjukkan bahwa mayoritas opini publik terhadap kondisi *bull* dan *bear market* Bitcoin pada komentar YouTube cenderung bersifat netral dan positif, sedangkan komentar bernada negatif relatif sangat rendah. Temuan ini sejalan dengan kondisi pasar crypto sepanjang tahun 2024 yang secara umum mengalami fase tren naik (*bullish*), meningkatnya minat investor, serta berkembangnya sentimen optimis dari komunitas kripto. Karena



suasana pasar yang cenderung positif tersebut, wajar apabila komentar masyarakat lebih banyak mengandung unsur netral dan optimisme, sementara komentar negatif hanya muncul dalam proporsi yang kecil.

Hasil penelitian ini juga sejalan dengan beberapa penelitian terbaru yang menunjukkan bahwa algoritma Support Vector Machine (SVM) memiliki performa yang baik dalam analisis sentimen berbasis teks pada media sosial. Penelitian oleh Saputra dkk. (2025) (Saputra et al., 2025) menunjukkan bahwa penerapan SVM pada analisis sentimen komentar YouTube dengan pendekatan normalisasi bahasa tidak baku mampu meningkatkan akurasi klasifikasi sentimen, sehingga memperlihatkan efektivitas SVM dalam menangani karakteristik data komentar pengguna pada platform YouTube. Selain itu, penelitian Siregar dkk. (2024) (Tjut Adek et al., 2025) menunjukkan bahwa kombinasi tahapan preprocessing, pembobotan TF-IDF, dan metode SVM menghasilkan performa klasifikasi yang tinggi pada analisis sentimen berbasis komentar YouTube. Temuan tersebut mendukung hasil penelitian ini, di mana model SVM memperoleh tingkat akurasi sebesar 98% dalam mengklasifikasikan komentar terkait kondisi bull market dan bear market Bitcoin. Penelitian Putri dkk. (2025) (Putri et al., 2025) juga menunjukkan bahwa metode SVM efektif digunakan untuk memahami persepsi pengguna melalui data teks media sosial, khususnya pada platform YouTube. Dengan demikian, hasil penelitian ini memperkuat temuan penelitian sebelumnya bahwa algoritma Support Vector Machine merupakan metode yang efektif untuk diterapkan pada analisis sentimen berbasis komentar media sosial yang bersifat tidak terstruktur dan berdimensi tinggi.

Namun demikian, performa pada kelas negatif masih dapat ditingkatkan mengingat jumlah datanya yang relatif sedikit dibandingkan kelas lainnya. Ketidakseimbangan distribusi kelas (*class imbalance*) ini berpotensi memengaruhi sensitivitas model terhadap kelas minoritas. Secara keseluruhan, penelitian ini membuktikan bahwa metode SVM efektif digunakan untuk analisis sentimen berbasis teks pada media sosial dalam mengidentifikasi kecenderungan opini publik terhadap kondisi pasar Bitcoin. Selain menunjukkan performa klasifikasi yang tinggi, hasil penelitian ini juga memberikan gambaran deskriptif bahwa dinamika sentimen di media sosial memiliki kecenderungan sejalan dengan kondisi pasar, sehingga dapat merefleksikan psikologi pasar (*market sentiment*) dalam ekosistem *cryptocurrency*.

## 4. KESIMPULAN

Berdasarkan hasil penelitian, algoritma Support Vector Machine (SVM) mampu mengklasifikasikan sentimen komentar YouTube terkait kondisi bull dan bear market Bitcoin dengan performa yang sangat baik, ditunjukkan oleh nilai akurasi sebesar 98%, macro F1-score sebesar 0,91, dan weighted F1-score sebesar 0,98. Hasil analisis menunjukkan bahwa kecenderungan opini publik didominasi oleh sentimen netral (50,42%) dan positif (47,08%), sedangkan sentimen negatif relatif rendah (2,50%), yang mengindikasikan bahwa persepsi masyarakat terhadap kondisi pasar Bitcoin selama periode penelitian cenderung stabil dan optimis. Temuan ini menjawab tujuan penelitian bahwa analisis sentimen berbasis komentar YouTube mampu menggambarkan kecenderungan opini publik terhadap bull dan bear market Bitcoin serta menunjukkan bahwa metode Support Vector Machine efektif digunakan untuk klasifikasi sentimen pada data teks media sosial. Penelitian ini juga memberikan kontribusi berupa bukti empiris bahwa komentar YouTube dapat dimanfaatkan sebagai sumber data untuk memahami kecenderungan opini publik terhadap kondisi pasar Bitcoin, sekaligus menunjukkan keterkaitan antara dinamika sentimen masyarakat dan kondisi pasar *cryptocurrency*. Dengan demikian, hasil penelitian diharapkan dapat mendukung pengembangan studi analisis sentimen pada domain *cryptocurrency* serta menjadi referensi dalam memahami psikologi pasar (*market sentiment*) berbasis opini publik digital.

## REFERENCES

- Albab, M. U., P., Y. K., & Fawaiq, M. N. (2023). Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic. *Jurnal Transformatika*, 20(2), 1–12. <https://doi.org/10.26623/transformatika.v20i2.5374>
- Arimbi Puspitasari, Diana Sava Salsabila, & Dwi Roliawati. (2025). Penerapan ResNet50-CNN untuk Optimalisasi Klasifikasi pada Data Fashion. *Indonesian Journal on Data Science*, 3(1), 1–12. <https://doi.org/10.30989/ijds.v3i1.1533>
- Bitcoin, C. (2024). Return Dan Risiko Investasi Terhadap Volume Perdagangan Cryptocurrency Bitcoin. *Jurnal Bisnis & Akuntansi Unsuraya*, 9(1), 1–11. <https://doi.org/10.35968/jbau.v9i1.1168>
- Hidayat, S., Herlina, N., & Nurul, G. (2024). Penerapan Model Support Vector Machine Pada Kasus Klasifikasi Teks Berdasarkan Tujuan SDGS Ke Tiga, Empat, Dan Enam. *SisInfo*, 6(2), 28–37. <https://doi.org/10.37278/sisinfo.v6i2.893>
- I Putu Gede Hendra Suputra, Linawati, Sukadarmika, I. G., & Sastra, N. P. (2025). Klasifikasi Judul Berita Bahasa Indonesia Menggunakan Support Vector Machine Dan Seleksi Fitur Mutual Information. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 22(1), 69–79. <https://doi.org/10.23887/jptkuniksha.v22i1.89158>
- Nur, M. (2026). Apakah Bitcoin Masih Layak sebagai Instrumen Investasi di Tengah Volatilitas Ekstrem? Studi terhadap Perilaku Investasi Generasi Z. *Salewangang: Jurnal Ekonomi*, 20(1), 43–53.
- Pinaria, A. P. P., Widodo, & Nugraheni, M. (2024). Analisis Sentimen Rancangan Undang-Undang Tindak Pidana Kekerasan Seksual/Undang-Undang Tindak Pidana Kekerasan Seksual Pengguna Twitter Menggunakan Algoritma Naïve Bayes Classifier. *PINTER: Jurnal Pendidikan Teknik Informatika Dan Komputer*, 8(2), 1–10.



- <https://doi.org/10.21009/pinter.8.2.1>
- Putra, I. G. S. D., & Putra, I. N. T. A. (2025). Implementasi Metode Naïve Bayes Pada Analisis Sentimen Pengguna Aplikasi Mobile Kita Bisa. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(2). <https://doi.org/10.23960/jitet.v13i2.6423>
- Putri, S. R., Asrianda, A., & Rosnita, L. (2025). Sentiment Analysis of Youtube and Gotube Reviews on Google Play Using the Support Vector Machine (SVM) Method in Indonesia. *Journal of Applied Informatics and Computing*, 9(3), 1025–1033. <https://doi.org/10.30871/jaic.v9i3.9461>
- Rifka Alkhilyatul Ma'rifat, I Made Suraharta, I. I. J. (2024). Analisis Sentimen Untuk Mengukur Ulasan Pengguna Aplikasi Mobile Legend Menggunakan Algoritma Naive Bayes, SVM, Random Fores, Decision Tree, dan Logistic Regression. *Jurnal Sistem Informasi*, 2(1), 306–312. <https://doi.org/10.18495/jsi.v16i1.152>
- Saputra, A. N. A., Saputro, R. E., & Saputra, D. I. S. (2025). Enhancing Sentiment Analysis Accuracy Using SVM and Slang Word Normalization on YouTube Comments. *Sinkron*, 9(2), 687–699. <https://doi.org/10.33395/sinkron.v9i2.14613>
- Sari Siregar, Y., Handoko, D., Khairani, M., Syahputri, N. I., & Harahap, H. (2024). Implementasi Data Mining Klasifikasi Algoritma Chaid Dalam Menentukan Pola Penerima Mahasiswa Baru. *Digital Transformation Technology*, 3(2), 978–989. <https://doi.org/10.47709/digitech.v3i2.3612>
- Suharsono, J. P., & Nurahman, D. (2024). Pemanfaatan Youtube Sebagai Media Peningkatan Pelayanan Dan Informasi. *Ganaya : Jurnal Ilmu Sosial Dan Humaniora*, 7(1), 298–304. <https://doi.org/10.37329/ganaya.v7i1.3157>
- Suryadewiansyah, M. K., & Tju, T. E. E. (2022). Naïve Bayes dan Confusion Matrix untuk Efisiensi Analisa Intrusion Detection System Alert. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 8(2), 81–88. <https://doi.org/10.25077/teknosi.v8i2.2022.81-88>
- Syahrohimi, I., Saputra, S. D., Saputra, R. W., Pranatawijaya, V. H., & Priskila, R. (2024). Perbandingan Analisis Sentimen Setelah Pilpres 2024 Di Twitter Menggunakan Algoritma Machine Learning. *JITET (Jurnal Informatika Dan Teknik Elektro Terapan)*, 12(2). <https://doi.org/10.23960/jitet.v12i2.4249>
- Tjut Adek, R., Fitri, Z., & Siregar, S. C. (2025). Analisis Sentimen Komentar Pada Saluran Youtube Beauty Vlogger Berbahasa Indonesia Menggunakan Metode Support Vector Machine. *Jurnal Algoritme*, 5(2), 164–175. <https://doi.org/10.35957/algoritme.v5i2.9692>
- Tri Ayu Mareta, Desty Endrawati Subroto, Lailaturrohmah Aulia, Siti Nuryanah, & Ratu Najwa Fadilah. (2025). Peran Media Sosial Youtube sebagai Media Edukasi dalam Pendidikan Generasi Z. *Guruku: Jurnal Pendidikan Dan Sosial Humaniora*, 3(1), 98–106. <https://doi.org/10.59061/guruku.v3i1.894>
- William, W., Maryati, M., Khesi, K., Alvina, J., & Selina. Ng, S. N. (2022). Analisis Kebijakan Pemerintah Terkait Ancaman Pengangguran Pasca Inflasi di Negara Indonesia. *Jurnal Multidisiplin Indonesia*, 1(4), 1066–1073. <https://doi.org/10.58344/jmi.v1i4.101>
- Yanto, R., Kesuma, H. Di, Afrudi, S., Komputer, I., Indo, U., Mandiri, G., Rawas, U. M., Rawas, M., & Info, A. (2025). Implementasi Text Mining Dalam Mengidentifikasi Similarity Judul Penelitian. *Djtechno : Jurnal Teknologi Informasi*, 6(1), 237–251. <https://doi.org/10.46576/djtechno>
- Zahra, F. A., Setiaji, P., & Triyanto, W. A. (2025). Klasifikasi Ekspresi Emosi Wajah Bahagia dan Tidak Bahagia Menggunakan Arsitektur Mobilenetv2 Berbasis Deep Learning. *Jurnal Sistem Informasi Dan Teknologi (SINTEK)*, 1(2), 1–6. <https://doi.org/10.24176/sitech.v8i1.15546>