

## PERBANDINGAN METODE SVM DAN NAÏVE BAYES DALAM ANALISIS SENTIMEN KONFLIK PERANG DAGANG CINA-AMERIKA PADA APLIKASI X

Muhammad Ridha

Manajemen Informatika, Politeknik Astra, Bekasi, 17530, Indonesia

*E-mail* : m.ridha@polytechnic.astra.ac.id

*Abstract-- The trade war issue between China and the United States has attracted significant public attention across various social media platforms, including Application X. Although many previous studies have compared the performance of Support Vector Machine (SVM) and Naïve Bayes algorithms in sentiment analysis, research specifically examining the performance of these two methods in analyzing public sentiment regarding the China-U.S. trade war on Application X is still limited. This study aims to fill this gap by comparing the performance of SVM and Multinomial Naïve Bayes in classifying sentiments in tweets related to this topic. The evaluation results show that Multinomial Naïve Bayes achieved an accuracy of 82%, with precision of 0.80 for negative sentiment and 0.82 for positive sentiment. However, this model was less optimal in recognizing negative sentiment, with a recall of only 0.64 and an F1-score of 0.71. In contrast, the SVM model showed superior performance, with an accuracy of 95% and an F1-score of 0.93 for negative sentiment and 0.96 for positive sentiment. SVM also exhibited more stable and balanced predictions, with low prediction errors for both sentiment classes. The advantage of SVM lies in its ability to handle class imbalance and recognize more complex patterns in the data. This study is expected to provide deeper insights into sentiment analysis for global strategic issues like the trade war and offer recommendations for the development of more effective classification models on social media.*

*Keywords* : Trade War, China-America, SVM, Naïve Bayes

*Abstrak-- Isu perang dagang antara Cina dan Amerika Serikat telah menjadi topik yang menarik perhatian publik di berbagai platform media sosial, termasuk Aplikasi X. Meskipun banyak penelitian sebelumnya telah membandingkan algoritma Support Vector Machine (SVM) dan Naïve Bayes dalam konteks analisis sentimen, studi yang secara khusus mengkaji performa kedua metode ini dalam menganalisis sentimen publik terhadap perang dagang Cina-Amerika di Aplikasi X masih sangat terbatas. Penelitian ini bertujuan untuk mengisi kekosongan tersebut dengan membandingkan kinerja SVM dan Multinomial Naïve Bayes dalam mengklasifikasikan sentimen cuitan terkait topik tersebut. Hasil evaluasi menunjukkan bahwa Multinomial Naïve Bayes menghasilkan akurasi sebesar 82%, dengan precision 0.80 untuk sentimen negatif dan 0.82 untuk sentimen positif. Namun, model ini kurang optimal dalam mengenali sentimen negatif, dengan recall negatif hanya mencapai 0.64 dan F1-score 0.71. Sebaliknya, model SVM menunjukkan kinerja yang lebih unggul, dengan akurasi 95% serta F1-score 0.93 untuk sentimen negatif dan 0.96 untuk sentimen positif. SVM juga menunjukkan prediksi yang lebih stabil dan seimbang, dengan kesalahan prediksi yang rendah pada kedua kelas sentimen. Keunggulan SVM terletak pada kemampuannya untuk menangani ketidakseimbangan kelas dan mengenali pola yang lebih kompleks dalam data. Hasil penelitian diharapkan dapat memberikan wawasan yang lebih dalam mengenai analisis sentimen isu-isu strategis global seperti perang dagang, serta memberikan rekomendasi untuk pengembangan model klasifikasi yang lebih efektif di media sosial.*

*Kata Kunci* : Perang Dagang, Cina-Amerika, SVM, Naïve Bayes

### I. PENDAHULUAN

Perang dagang antara Cina dan Amerika Serikat merupakan konflik ekonomi global yang dimulai sejak tahun 2018 dan terus berlanjut hingga kini. Perseteruan ini ditandai dengan pemberlakuan tarif tinggi terhadap produk impor, pembatasan teknologi, serta berbagai kebijakan proteksionis dari kedua negara [1]. Dampaknya tidak hanya dirasakan oleh sektor perdagangan dan keuangan, tetapi juga memengaruhi opini publik secara luas di berbagai negara, termasuk di dunia maya.

Media sosial, khususnya Aplikasi X (sebelumnya dikenal sebagai Twitter), menjadi salah satu ruang

digital utama bagi masyarakat untuk mengekspresikan opini mereka terhadap isu ini. Melalui komentar, unggahan, dan diskusi daring, publik menyampaikan pandangan mereka yang beragam baik dukungan, kritik, maupun kecemasan terhadap dampak kebijakan perang dagang tersebut. Hal ini menjadikan Aplikasi X sebagai sumber data yang potensial dalam menggali dan menganalisis sentimen publik secara cepat dan real-time [2].

Di tengah melimpahnya data tidak terstruktur dari media sosial, teknologi analisis sentimen menjadi solusi penting untuk mengekstrak informasi berharga. Salah satu pendekatan yang digunakan secara luas

adalah *machine learning*, yang memungkinkan pemodelan otomatis terhadap opini publik berdasarkan data teks. Dalam ranah ini, metode seperti Support Vector Machine (SVM) dan Naïve Bayes dikenal sebagai algoritma andalan untuk klasifikasi sentimen, masing-masing dengan kelebihan dan kekurangannya. SVM terkenal karena akurasinya yang tinggi [3], sementara Naïve Bayes unggul dalam kesederhanaan dan kecepatan proses [4].

Sejumlah penelitian sebelumnya telah melakukan perbandingan kinerja antara algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam berbagai domain aplikasi. Sebagai contoh, penelitian yang dikemukakan oleh Aini dkk. [5] menunjukkan bahwa algoritma SVM memiliki keunggulan signifikan dalam menganalisis dampak ekonomi selama masa pandemi Covid-19, dengan capaian akurasi yang sangat tinggi, yaitu mencapai 100%. Di sisi lain, penelitian lain mengungkapkan bahwa algoritma Naïve Bayes tetap mampu bersaing secara kompetitif, khususnya ketika digunakan dalam kondisi dataset yang berukuran kecil, seperti yang terlihat pada studi klasifikasi pesan singkat (SMS) spam [6]. Selain itu, algoritma Naïve Bayes juga menunjukkan performa yang baik ketika fitur-fitur dalam data bersifat independen. Hal ini dibuktikan dalam penelitian yang dilaporkan oleh Setiawati dkk. [7], yaitu algoritma Naïve Bayes berhasil mencapai akurasi sebesar 81,50%, sementara algoritma SVM pada studi yang sama memperoleh akurasi sebesar 80,17%.

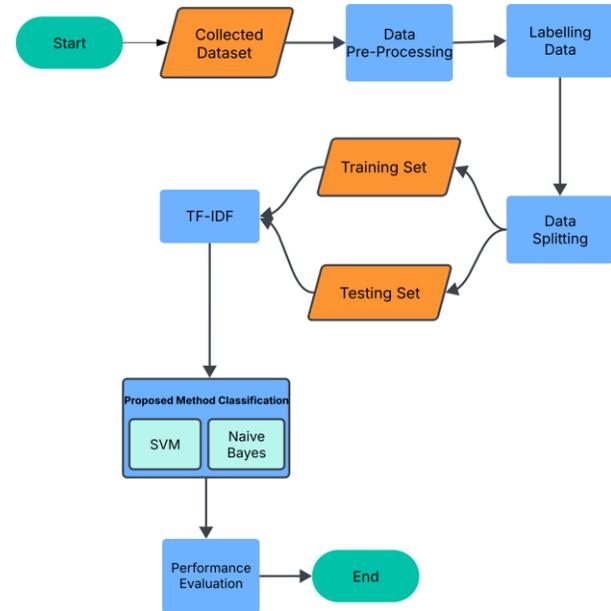
Penelitian Alyafeai dkk. tahun 2021 [8] menyatakan bahwa efektivitas metode sangat bergantung pada karakteristik data dan tahap prapemrosesan. Hal ini terbukti dari studi [9] yang mengkaji lagi sentimen publik terhadap kebijakan pemerintah di Indonesia, hasilnya algoritma SVM mampu mencapai tingkat akurasi hingga 88%, sementara algoritma Naïve Bayes memperoleh akurasi sebesar 82%. Temuan ini mengindikasikan bahwa SVM memiliki kapabilitas yang lebih unggul dalam menangani data teks, terutama dalam mengenali pola-pola yang bersifat kompleks dibandingkan dengan Naïve Bayes.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membandingkan performa metode Support Vector Machine (SVM) dan Naïve Bayes dalam mengklasifikasikan sentimen publik terhadap konflik perang dagang antara Cina dan Amerika Serikat dengan menggunakan data dari Aplikasi X. Studi yang secara spesifik membandingkan kedua algoritma dalam konteks isu ini pada platform media sosial tersebut masih sangat terbatas. Melalui pendekatan ini, diharapkan dapat diperoleh pemahaman yang lebih komprehensif mengenai

persepsi publik terhadap isu perdagangan internasional, serta rekomendasi metode klasifikasi yang paling efektif untuk analisis sentimen pada topik-topik strategis berskala global.

**II. METODOLOGI**

Penelitian ini menganalisis persepsi publik terhadap konflik perang dagang Cina-AS melalui data dari media sosial Aplikasi X. Metodologi penelitian dapat di lihat pada gambar 1 di bawah ini.



Gambar 1. Metodologi Penelitian

**1. Collected Dataset**

Pada tahap pengumpulan data, penelitian ini menggunakan metode *tweet harvesting* melalui pemanfaatan API dari Aplikasi X. Data yang dikumpulkan merupakan unggahan pengguna yang mengandung kata kunci relevan dengan topik perang dagang antara Cina dan Amerika Serikat. Proses pengumpulan dilakukan dalam rentang waktu mulai dari tanggal 1 Januari 2025 hingga 16 April 2025. Data ini nantinya akan digunakan sebagai dasar dalam proses prapemrosesan dan pelabelan untuk keperluan klasifikasi sentimen menggunakan algoritma machine learning.

**2. Data Preprocessing**

Pada tahap prapemrosesan data, dilakukan serangkaian tahapan untuk membersihkan dan menyiapkan data teks sebelum dianalisis. Langkah-langkah *preprocessing* ini mencakup penghapusan mention pengguna (misalnya @username), tautan URL, dan hashtag yang tidak memberikan kontribusi

langsung terhadap konteks sentimen. Selanjutnya, dilakukan penghapusan angka dan tanda baca untuk meningkatkan konsistensi format teks. Teks juga dikonversi menjadi huruf kecil (*lowercasing*) agar model tidak membedakan kata berdasarkan kapitalisasi. Selain itu, *stopword* atau kata-kata umum yang tidak bermakna signifikan dalam analisis sentimen seperti “yang”, “oleh”, dan lain lain akan dihapus dari teks. Untuk menghindari duplikasi informasi, juga diterapkan penghapusan data ganda. Proses ini diimplementasikan menggunakan bahasa pemrograman Python, dan dijalankan pada platform Google Colaboratory dengan memanfaatkan pustaka pemrosesan teks seperti *re* dan daftar *stopword* manual.

### 3. Labeling data

Pada tahap *labeling* atau pelabelan data, penelitian ini menggunakan pendekatan lexicon-based [10]. Metode ini merupakan teknik pelabelan otomatis yang memanfaatkan kamus sentimen berbahasa Indonesia yang telah terdefinisi, di mana setiap kata memiliki nilai sentimen positif atau negatif. Proses dilakukan dengan mendeteksi kemunculan kata-kata dalam setiap cuitan yang sesuai dengan daftar pada lexicon. Jika cuitan mengandung lebih banyak kata bernilai positif, maka akan dilabeli sebagai *positif*, dan sebaliknya. Metode ini dipilih karena bersifat *unsupervised* dan tidak memerlukan data latih, serta cukup efisien digunakan untuk pelabelan awal sebelum proses klasifikasi dengan algoritma pembelajaran mesin.

### 4. Data Splitting

Setelah data selesai melalui proses pelabelan dan prapemrosesan, tahap berikutnya adalah pembagian data (*data splitting*). Proses ini dilakukan untuk memisahkan data ke dalam dua bagian, yaitu data latih (*training data*) dan data uji (*testing data*).

### 5. Proposed Method Classification

Pada tahap klasifikasi, penelitian ini menggunakan dua algoritma pembelajaran mesin, yaitu Support Vector Machine (SVM) dan Multinomial Naïve Bayes, untuk mengklasifikasikan sentimen dari cuitan yang berkaitan dengan isu perang dagang antara Cina dan Amerika Serikat. Sebelum data dimasukkan ke dalam model, dilakukan konversi teks ke dalam bentuk numerik menggunakan teknik TF-IDF (*Term Frequency – Inverse Document Frequency*)[11]. Pendekatan ini digunakan untuk memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam dokumen dan keseluruhan korpus, sehingga kata-kata yang lebih bermakna memiliki pengaruh lebih besar pada model.

Setelah data direpresentasikan dalam bentuk vektor TF-IDF, tahap pelatihan model dilakukan. SVM [12] dipilih karena kemampuannya dalam menangani data berdimensi tinggi dan keunggulannya dalam menghasilkan hyperplane optimal yang mampu memisahkan kelas dengan margin maksimum. Sementara itu, digunakan Multinomial Naïve Bayes [13], yang merupakan varian dari Naïve Bayes yang paling sesuai untuk data teks, karena mengasumsikan bahwa fitur (kata) mengikuti distribusi multinomial dan bekerja efektif pada frekuensi kata.

### 6. Performance Evaluation

Evaluasi kinerja model dilakukan untuk menilai sejauh mana efektivitas algoritma Naïve Bayes dan SVM dalam melakukan klasifikasi sentimen pada data yang telah disiapkan. Penilaian dilakukan dengan menerapkan sejumlah metrik evaluasi yaitu confusion matrix [14], diantaranya akurasi, presisi, recall, dan F1-score, yang masing-masing memberikan gambaran dari berbagai sudut pandang terhadap performa model. Dengan penerapan metrik-metrik tersebut, penelitian ini dapat mengevaluasi dan membandingkan performa masing-masing model secara menyeluruh, serta menentukan metode klasifikasi yang paling optimal untuk analisis sentimen pada isu perang dagang antara Cina dan Amerika Serikat.

## III. HASIL PENELITIAN

### 1. Collected Dataset

Hasil dari proses pengumpulan data menunjukkan bahwa sebanyak 2.812 cuitan berhasil diperoleh dari Aplikasi X dalam periode 1 Januari hingga 16 April 2025. Seluruh data yang terkumpul berkaitan dengan topik perang dagang antara Cina dan Amerika Serikat, yang diambil berdasarkan kata kunci yang relevan. Data ini masih dalam bentuk mentah saat diperoleh, sehingga selanjutnya dilakukan proses prapemrosesan dan pelabelan untuk keperluan analisis sentimen. Jumlah data yang diperoleh menunjukkan bahwa topik ini cukup ramai diperbincangkan selama periode waktu tersebut.

### 2. Data Preprocessing

Proses prapemrosesan data dilakukan menggunakan Python di lingkungan Google Colaboratory, dan terdiri dari beberapa tahap pembersihan teks. Beberapa proses yang dilakukan *cleaning*, *case folding*, *stopword removal*, dan *duplicate remove*. Setelah proses ini dijalankan, diketahui bahwa tidak terdapat data yang terduplikasi dari total 2.812 cuitan, sehingga tidak ada pengurangan jumlah data. Table 1 di bawah ini adalah

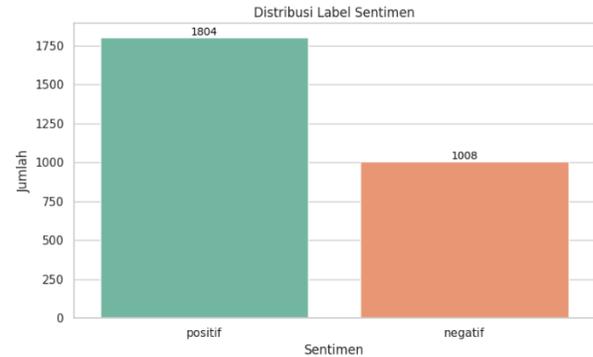
contoh kalimat sebelum dan setelah data *preprocessing*.

Tabel 1. Contoh Kalimat Sebelum dan Setelah Data *Preprocessing*

No	Sebelum	Setelah
1	@hegar wibawa @WidasSatyo Kebijakan tarif Trump bikin pasar gonjang-ganjing Dow Jones anjlok 2.200 poin Nasdaq masuk bear market. Ini bisa picu inflasi dan resesi global termasuk di Indonesia karena harga impor naik dan ekspor turun. Tarif mungkin lindungi industri lokal tapi risiko perang dagang <a href="https://t.co/UnsVBF1qo3">https://t.co/UnsVBF1qo3</a>	kebijakan tarif trump bikin pasar gonjang-ganjing dow jones anjlok poin nasdaq masuk bear market ini bisa picu inflasi dan resesi global termasuk di indonesia karena harga impor naik dan ekspor turun tarif mungkin lindungi industri lokal tapi risiko perang dagang
2	@garysenal @mohmahfudmd Dampak terburuk dari perang dagang AS-China untuk Indonesia bisa meliputi penurunan ekspor terutama di sektor elektronik dan tekstil serta kenaikan biaya impor yang memicu inflasi. Studi memperkirakan kerugian ekspor hingga \$336 juta. Tekanan pada rupiah juga mungkin terjadi	dampak terburuk dari perang dagang aschina untuk indonesia bisa meliputi penurunan ekspor terutama di sektor elektronik dan tekstil serta kenaikan biaya impor yang memicu inflasi studi memperkirakan kerugian ekspor hingga juta tekanan pada rupiah juga mungkin terjadi
3	Kerugian Defisit Perdagangan: Pada tahun 2024 AS mengimpor \$462.5 miliar dari China tetapi hanya mengekspor \$199.2 miliar. Negara-negara defisit seperti AS lebih menderita dalam perang dagang menghadapi kekurangan barang yang dibutuhkan	kerugian defisit perdagangan pada tahun as mengimpor miliar dari china tetapi hanya mengekspor miliar negaranegara defisit seperti as lebih menderita dalam perang dagang menghadapi kekurangan barang yang dibutuhkan

### 3. Labeling data

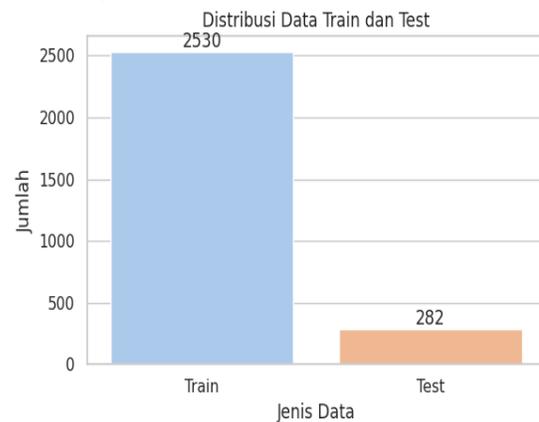
Setelah dilakukan proses *preprocessing*, tahap selanjutnya adalah pelabelan data menggunakan metode *lexicon-based*. Berdasarkan proses pelabelan yang dilakukan terhadap 2.812 cuitan, diperoleh hasil sebanyak 1.804 cuitan dengan sentimen positif dan 1.008 cuitan dengan sentimen negatif. Pelabelan dilakukan secara otomatis menggunakan *Python* dan memastikan bahwa setiap data telah diberi kategori yang sesuai untuk keperluan tahap klasifikasi selanjutnya. Gambar 2 di bawah ini distribusi *labeling data*.



Gambar 2. Distribusi *Labeling Data*

### 4. Data Splitting

Pada penelitian ini, digunakan proporsi pembagian sebesar 90:10, di mana 90% data digunakan untuk melatih model klasifikasi, dan 10% sisanya digunakan untuk menguji performa model. Berdasarkan jumlah total data sebanyak 2.812 cuitan, Dengan rasio tersebut, diperoleh 2.530 data yang digunakan sebagai data latih (*training set*) dan 282 data sebagai data uji (*testing set*). Gambar 3 merupakan distribusi data latih dan data uji.



Gambar 3. Distribusi Data Latih dan Data Uji

### 5. Klasifikasi dengan SVM dan Naïve Bayes

Pada tahap klasifikasi, dua model pembelajaran mesin yaitu Multinomial Naïve Bayes dan Support Vector Machine (SVM) diuji untuk mengklasifikasikan sentimen cuitan terkait perang dagang Cina–Amerika. Berdasarkan hasil evaluasi seperti gambar 4, model ini memperoleh akurasi sebesar 82%, dengan nilai precision sebesar 0.80 untuk sentimen negatif dan 0.82 untuk sentimen positif. Meskipun cukup akurat dalam mengenali cuitan dengan sentimen positif (*recall* = 0.91), model ini masih kurang mampu menangkap sentimen negatif secara optimal (*recall* = 0.64). Hal ini berdampak pada nilai F1-score untuk kelas negatif yang hanya mencapai 0.71, lebih rendah dibandingkan kelas

positif (0.86). Perbedaan ini menunjukkan adanya kemungkinan bias dari model terhadap kelas yang jumlahnya lebih dominan dalam data, yaitu sentimen positif.

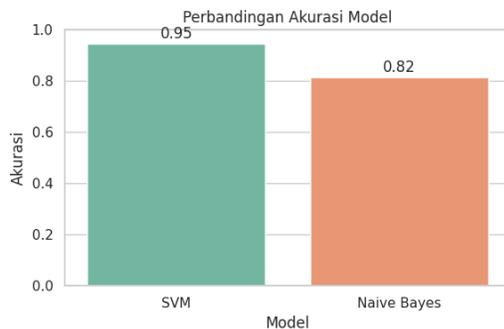
Sebaliknya, model Support Vector Machine (SVM) menunjukkan performa klasifikasi yang jauh lebih unggul dan stabil. Dengan akurasi mencapai 95%, SVM berhasil mempertahankan nilai precision dan recall yang tinggi di kedua kelas. Untuk sentimen negatif, precision dan recall masing-masing berada di angka 0.92 dan 0.93, sedangkan untuk sentimen positif mencapai 0.96 dan 0.95. F1-score yang diperoleh pun sangat memuaskan, yakni 0.93 untuk negatif dan 0.96 untuk positif. Hasil ini menunjukkan bahwa SVM mampu mengatasi ketidakseimbangan kelas serta menghasilkan prediksi yang lebih konsisten. Oleh karena itu, dalam konteks penelitian ini, SVM dapat dianggap sebagai metode klasifikasi yang lebih efektif dibandingkan Naïve Bayes dalam mengolah data sentimen terkait perang dagang Cina–Amerika. Perbandingan akurasi kedua model dapat dilihat pada gambar 5.

Classification Report - SVM:				
	precision	recall	f1-score	support
negatif	0.92	0.93	0.93	101
positif	0.96	0.96	0.96	181
accuracy			0.95	282
macro avg	0.94	0.94	0.94	282
weighted avg	0.95	0.95	0.95	282

Classification Report - Naive Bayes:				
	precision	recall	f1-score	support
negatif	0.80	0.64	0.71	101
positif	0.82	0.91	0.86	181
accuracy			0.82	282
macro avg	0.81	0.78	0.79	282
weighted avg	0.81	0.82	0.81	282

Gambar 4. Hasil klasifikasi SVM dan Naïve Bayes



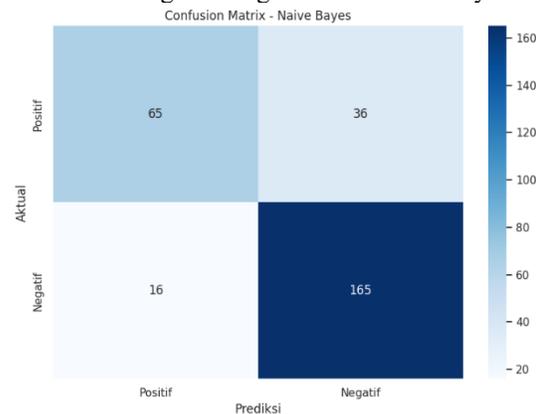
Gambar 5. Perbandingan Akurasi Model SVM dan Naïve Bayes

### 6. Hasil Confusion Matrix

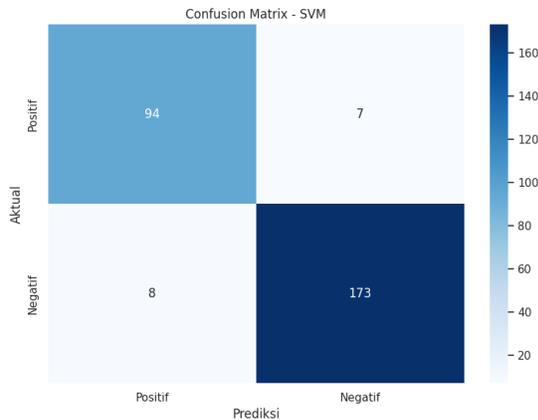
Hasil visualisasi pada gambar 6 menunjukkan confusion matrix dari model Naïve Bayes yang

digunakan untuk klasifikasi sentimen. Evaluasi terhadap data uji menghasilkan jumlah prediksi yang benar pada kelas negatif (*True Negative/TN*) sebanyak 165, dan jumlah kesalahan dalam memprediksi data negatif sebagai positif (*False Positive/FP*) sebanyak 16. Pada kelas positif, model berhasil memprediksi 65 data positif dengan benar (*True Positive/TP*), sementara sebanyak 36 data positif diprediksi salah sebagai negatif (*False Negative/FN*). Dari hasil ini, terlihat bahwa model cukup baik dalam mengenali data negatif, namun cenderung kurang akurat dalam mengenali sebagian data positif, yang bisa mengindikasikan bahwa model belum sepenuhnya seimbang dalam menangani kedua kelas sentimen.

Sementara itu, gambar 7 menampilkan confusion matrix dari model Support Vector Machine (SVM) yang juga digunakan untuk tugas klasifikasi sentimen. Model ini menunjukkan kinerja yang lebih stabil dan seimbang. Berdasarkan confusion matrix, model berhasil mengklasifikasikan 173 data negatif dengan benar (*True Negative/TN*) dan hanya melakukan kesalahan pada 8 data negatif yang diprediksi sebagai positif (*False Positive/FP*). Untuk data positif, sebanyak 94 berhasil diklasifikasikan dengan benar (*True Positive/TP*), dan hanya 7 data positif yang diklasifikasikan salah sebagai negatif (*False Negative/FN*). Hasil ini menunjukkan bahwa model SVM mampu memisahkan kedua kelas dengan baik dan menghasilkan jumlah kesalahan yang relatif rendah dibandingkan dengan model Naïve Bayes.



Gambar 6. Confusion Matrix Naïve Bayes



Gambar 7. Confusion Matrix SVM

#### IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa model Support Vector Machine (SVM) menunjukkan performa klasifikasi sentimen yang lebih unggul dibandingkan dengan model Multinomial Naïve Bayes. Model SVM berhasil mencapai akurasi sebesar 95%, sementara Multinomial Naïve Bayes menghasilkan akurasi sebesar 82%. SVM mampu mengklasifikasikan sentimen positif dan negatif dengan lebih seimbang dan akurat, sementara Multinomial Naïve Bayes cenderung bias terhadap kelas positif dan gagal memprediksi data negatif dengan tepat. Hal ini menunjukkan bahwa SVM lebih cocok digunakan dalam tugas klasifikasi sentimen pada data cuitan terkait perang dagang antara Cina dan Amerika Serikat.

Sebagai saran, penelitian selanjutnya dapat mempertimbangkan penerapan teknik penyeimbangan data seperti oversampling atau SMOTE [15] untuk mengurangi ketimpangan distribusi kelas, terutama agar model seperti Naïve Bayes dapat mengenali sentimen negatif dengan lebih baik. Selain itu, penggunaan fitur lanjutan seperti n-gram, stemming, atau word embedding dapat meningkatkan representasi data dan kinerja model secara keseluruhan. Eksplorasi model lain seperti Random Forest, XGBoost, atau pendekatan berbasis deep learning juga bisa menjadi alternatif yang menjanjikan dalam klasifikasi sentimen pada data media sosial.

#### V. DAFTAR PUSTAKA

[1] M. Kholid, "The US-China Trade War Phenomenon (2018-2020) In Economic Diplomacy Perspective," *Politik Indonesia: Indonesian Political Science Review*, vol. 7, no. 1, pp. 83–97, Apr. 2022, doi: 10.15294/ipsr.v7i1.38488.

[2] N. Arifah Wahyuni, D. Rizky Amelia, and tishom Al Khoiry, "Exploring Public Sentiment on ChatGPT-An Analysis of Twitter Data Article Info," *JEE*, vol. 13, no. 1, pp. 51–60, 2024, [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/jee>

[3] M. F. Alfauzan, Y. Sibaroni, and F. Fitriyani, "Sentiment Classification of Fuel Price Rise in Economic Aspects Using Lexicon and SVM Method," *sinkron*, vol. 8, no. 4, pp. 2526–2536, Oct. 2023, doi: 10.33395/sinkron.v8i4.12851.

[4] F. Yudistira and A. Rahman Isnain, "Analisis Sentimen Terhadap Seleksi CPNS Tahun 2024 Berbasis Media Sosial X Menggunakan Algoritma Naïve Bayes," *Jurnal Pendidikan dan Teknologi Indonesia (JPTI)*, vol. 5, no. 3, pp. 887–897, 2025, doi: 10.52436/1.jpti.731.

[5] Q. Aini, R. R. Fauzi, and E. Khudzaeva, "Economic Impact due Covid-19 Pandemic: Sentiment Analysis on Twitter Using Naïve Bayes Classifier and Support Vector Machine," *JOIV: International Journal on Informatics Visualization*, vol. 7, no. 3, pp. 733–741, Sep. 2023, doi: 10.30630/joiv.7.3.1474.

[6] Widyawati and Sutanto, "PERBANDINGAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE (SVM) DALAM KLASIFIKASI SMS SPAM BERBAHASA INDONESIA," *Jurnal Sains & Teknologi*, vol. 3, no. 2, pp. 178–194, 2019.

[7] P. A. Setiawati, I. M. A. D. Suarjaya, and I. N. P. Trisna, "Sentiment Analysis of Unemployment in Indonesia During and Post COVID-19 on X (Twitter) Using Naïve Bayes and Support Vector Machine," *Journal of Information Systems and Informatics*, vol. 6, no. 2, pp. 662–675, Jun. 2024, doi: 10.51519/journalisi.v6i2.713.

[8] Z. Alyafeai, M. S. Al-shaibani, M. Ghaleb, and I. Ahmad, "Evaluating Various Tokenizers for Arabic Text Classification," Jun. 2021, [Online]. Available: <http://arxiv.org/abs/2106.07540>

[9] Ardiyansah and Parjito, "Perbandingan Metode Naïve Bayes dan Support Vector Machine Dalam Analisis Sentimen Terhadap Tokoh Publik," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 6, pp. 2813–2821, 2024, doi: 10.30865/klik.v4i6.1928.

[10] H. Mardiana, M. Walid, and A. Kisnu Darmawan, "Lexicon-Based and Naïve Bayes Sentiment ... LEXICON-BASED AND NAÏVE BAYES SENTIMENT ANALYSIS FOR RECOMMENDING THE BEST MARKETPLACE SELECTION AS A MARKETING STRATEGY FOR MSMEs," vol. 19, no. 2, 2023, doi: 10.33480/pilar.v19i2.4176.

- [11] S. Singh, K. Kumar, and B. Kumar, "Sentiment Analysis of Twitter Data Using TF-IDF and Machine Learning Techniques," in *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, IEEE, May 2022, pp. 252–255. doi: 10.1109/COM-IT-CON54601.2022.9850477.
- [12] R. Malha and P. Amaral, "A Maximal Margin Hypersphere SVM," 2021, pp. 304–319. doi: 10.1007/978-3-030-86976-2\_21.
- [13] Yuyun, Nurul Hidayah, and Supriadi Sahibu, "Algoritma Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Pemerintah Terhadap Penanganan Covid-19 Menggunakan Data Twitter," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 4, pp. 820–826, Aug. 2021, doi: 10.29207/resti.v5i4.3146.
- [14] M. Rakha, M. Dwi Sulistiyo, D. Nasien, and M. Ridha, "A Combined MobileNetV2 and CBAM Model to Improve Classifying the Breast Cancer Ultrasound Images," *Journal of Applied Engineering and Technological Science*, vol. 6, no. 1, pp. 561–578, 2024.
- [15] Muhammad Ridha, Dade Nurjanah, and Muhammad Rakha, "Multilabel Classification Abusive Language and Hate Speech on Indonesian Twitter using Transformer Model: IndoBERTweet & IndoRoBERTa," in *The 4th International Conference on Intelligent Cybernetics Technology & Applications 2024 (ICICyTA 2024)*, 2024.