

# IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN ISU RESESI EKONOMI 2023 DI INDONESIA PADA PLATFORM TWITTER

Ribka Tesalonika<sup>1)</sup>, Evangs Mailoa<sup>2)</sup>

1. Universitas Kristen Satya Wacana, Salatiga
2. Universitas Kristen Satya Wacana, Salatiga

## Article Info

**Kata Kunci:** Analisis Sentimen; Naive Bayes; Resesi Ekonomi

**Keywords:** *Economic Recession*; Naive Bayes; *Sentiment Analysis*

## Article history:

Received 4 November 2023

Revised 18 November 2023

Accepted 2 December 2023

Available online 1 March 2024

## DOI :

<https://doi.org/10.29100/jipi.v9i1.4288>

\* Corresponding author.

Ribka Tesalonika

E-mail address:

[672019327@student.uksw.edu](mailto:672019327@student.uksw.edu)

## ABSTRAK

Resesi ekonomi merupakan kondisi dimana suatu negara mengalami kemerosotan aktivitas ekonomi. Isu mengenai resesi 2023 di Indonesia mulai hangat diperbincangkan semenjak terjadinya perang antara Rusia dan Ukraina, pandemi Covid-19 dan inflasi yang tak terkendali. Munculnya isu tersebut menimbulkan berbagai tanggapan dari masyarakat dan beberapa diantaranya menyampaikan pendapatnya melalui media sosial salah satunya Twitter. Dari permasalahan tersebut, diperlukan analisa mengenai sentimen masyarakat terhadap tweet untuk melihat respon mengenai Isu Resesi Ekonomi 2023 di Indonesia dari masyarakat yang aktif. Penelitian ini menerapkan algoritma Naive Bayes dalam melakukan pengklasifikasian data dan tahapannya dimulai dengan pengumpulan data tweet dengan teknik scraping, lalu pembersihan data, pembagian data latih dan data uji, pelabelan data, pra-pemrosesan data, klasifikasi kemudian diakhiri dengan evaluasi model menggunakan confusion matrix. Jumlah data yang dikumpulkan setelah melewati proses pembersihan data sebanyak 3301 tweet. Akurasi model yang diperoleh mencapai 93,46% dengan prediksi label sentimen positif 1302 tweet, sentimen negatif 1236 tweet dan sentimen netral sejumlah 764 tweet.

## ABSTRACT

Economic recession is a condition where a country experiences a decline in economic activity. The issue of the 2023 recession in Indonesia has become a hot topic of discussion since the outbreak of the war between Russia and Ukraine, the Covid-19 pandemic, and uncontrollable inflation. The emergence of this issue has generated various responses from the public, and some of them have expressed their opinions through social media, particularly Twitter. From these issues, an analysis of public sentiment towards tweets is needed to understand the response of the active Indonesian citizens regarding the 2023 Economic Recession issue. This research applies the Naive Bayes algorithm for data classification, and the process starts with data collection through scraping techniques, followed by data cleaning, splitting the data into training and testing sets, data labeling, data pre-processing, classification process, and concludes with model evaluation using a confusion matrix. The number of data collected after the data cleaning process is 3301 tweets. The obtained model accuracy is 93.46%, with a prediction of 1301 tweets having a positive sentiment, 1237 tweets having a negative sentiment, and 763 tweets being labeled as neutral sentiment.

## I. PENDAHULUAN

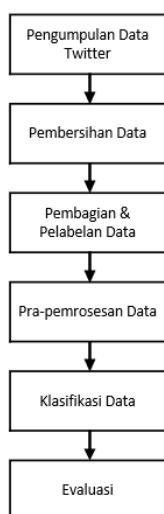
ISU mengenai resesi perekonomian Indonesia di tahun 2023 mendatang kini sedang menjadi perbincangan hangat di kalangan masyarakat, khususnya sejak terjadinya perang antara Rusia dan Ukraina, kebijakan pandemi di China, serta inflasi yang tak terkendali [1]. Faktor-faktor tersebut membuat kondisi perekonomian global di tahun yang akan datang menjadi masalah kepentingan umum bahkan juga menyebabkan kecemasan. Pembahasan tentang kondisi perekonomian global memunculkan kata resesi. Resesi ekonomi itu sendiri merupakan suatu kondisi ekonomi negara yang sedang melemah atau mengalami kemerosotan [2]. Banyak masyarakat yang kini menyampaikan keluh kesahnya mengenai isu resesi 2023 melalui media sosial. Salah satu media sosial yang sering digunakan untuk menyampaikan opini yaitu Twitter [3]. Opini masyarakat mengenai resesi ekonomi tentu

beragam, terdapat opini masyarakat yang siap dan optimis menghadapi isu resesi mendatang, namun ada sentimen masyarakat yang justru khawatir akan kondisi ekonomi kedepannya. Dengan memanfaatkan data tweet dari Twitter, penelitian ini melakukan analisis sentimen terhadap isu resesi ekonomi 2023 di Indonesia. Rumusan dari penelitian ini yaitu untuk mengklasifikasikan sentimen tersebut ke dalam 3 kelas yaitu positif, netral dan negatif dan hanya untuk tweet berbahasa Indonesia. Dalam menganalisis sentimen, penelitian ini menggunakan algoritma Naïve Bayes Classifier dikarenakan Algoritma tersebut merupakan salah satu algoritma pengklasifikasian yang memiliki kinerja sangat baik terutama untuk klasifikasi teks [4]. Diharapkan hasil dari analisa penelitian ini dapat membantu pemerintah dan peneliti terkait dalam melihat respon mengenai Isu Resesi Ekonomi 2023 di Indonesia dari masyarakat yang aktif.

Beberapa penelitian mengenai analisis sentimen sudah dilakukan sebelumnya, di antaranya adalah penelitian yang menerapkan naive bayes dalam menganalisa sentimen dampak corona, penelitian ini melibatkan sebanyak 796 dataset dengan perbandingan 75 : 25 untuk data latih dan uji, dan menghasilkan tingkat akurasi sebesar 67% [5]. Penelitian lainnya yang dilakukan oleh Romaito, dkk dengan membandingkan Algoritma Naive Bayes Classifier dan Support Vector Machine, Hasil dari perbandingannya yaitu algoritma Naive Bayes memperoleh hasil terbaik dengan tingkat akurasi 81,7% [6]. Telah banyak penelitian yang menganalisis sentimen berbagai topik namun sejauh ini belum ada penelitian yang membahas terkait sentimen masyarakat terhadap Isu Resesi Ekonomi 2023 di Indonesia. Berdasarkan penelitian yang telah dilakukan sebelumnya, dapat disimpulkan bahwa algoritma Naive Bayes dapat digunakan untuk analisis sentimen yang memberikan tingkat akurasi yang tinggi. Seluruh studi terdahulu tersebut menjadi referensi utama untuk dilakukannya penelitian mengenai Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Isu Resesi Ekonomi 2023 di Indonesia pada Platform Twitter.

## II. METODE PENELITIAN

Metodologi penelitian ini terdiri dari beberapa tahap antara lain Pengumpulan data, Pembersihan data, Pelabelan dan Pembagian Data, Pra-pemrosesan Data, Klasifikasi dan Hasil Evaluasi. Tahap penelitian ini dapat dilihat pada Gambar 1.



Gambar. 1. Tahapan Penelitian

### A. Pengumpulan Data

Penelitian ini mengumpulkan data tweet dengan teknik scraping menggunakan *library sncscrape* pada bahasa pemrograman Python. Saat melakukan scraping data, *library sncscrape* ini tidak memerlukan API dari Twitter [7], namun teknik ini memerlukan konfigurasi tambahan seperti *search*, *since*, dan *until*. Konfigurasi *search* digunakan untuk mengisi kata kunci dari topik tweet yang ingin dicari, *since* dan *until* digunakan untuk mengisi tanggal awal dan tanggal akhir periode tweet yang ingin dicari. Data yang diambil pada penelitian ini adalah data tweet berbahasa Indonesia. Penelitian ini memperoleh sejumlah 20.000 data tweet, dimana 10.000 tweet didapatkan dari kata kunci #Resesi2023 dan 10.000 tweet dari kata kunci Resesi Ekonomi 2023 Indonesia dalam periode waktu tweet sejak tanggal 1 Juli 2022 – 21 April 2023. Tweet diambil secara acak baik dari pengguna maupun berita online di Twitter.

### B. Pembersihan Data

Data tweet yang terkumpul masih berupa data mentah seperti tweet yang masih mengandung data duplikat,

karakter spesial, emoji, tagar, angka dan *tag*, Oleh karena itu, pembersihan data perlu dilakukan untuk menghilangkan beberapa masalah tersebut yang dapat menghambat proses klasifikasi teks [8]. Penelitian ini menggunakan *tools* RapidMiner dalam melakukan proses analisis sentimen. Proses pembersihan data dilakukan menggunakan fitur *replace* pada Rapidminer.

### C. Pembagian dan Pelabelan Data

Data yang sudah bersih kemudian dibagi menjadi data latih dan uji dengan perbandingan 70:30. Data latih digunakan untuk melatih model klasifikasi berdasarkan pola-pola yang ada, sedangkan data uji untuk menguji model klasifikasi yang telah dibuat. Proses berikutnya yaitu pemberian label untuk data latih. Contoh pemberian label dengan kategori positif, netral dan negatif pada penelitian ini ditunjukkan pada Tabel I.

TABEL I  
 CONTOH PELABELAN DATA TWEET

Tweet	Label
Yuk kita lawan resesi dengan investasi Biar optimis menuju 2023 yang terang! (jangan lupa riset-riset lagi ya bestie untuk semua investasi yang kalian pilih). Peace, Love and Gawl~	Positive
Semoga dengan adanya G20, issue resesi 2023 ini gak terjadi yaa... Dengan adanya pandemi aja, aku trauma banget gaji dipotong dan sampe sekarang belum balik normal. Sedih..	Negative
Iya dong. Kan sdh saya buka jalannya. Silahkan dilanjutkan. Saya butuh uangnya. 2023 resesi, saya tidak peduli. I want my money!! Hahaha etic field strength	Netral

### D. Pra-pemrosesan Data

Terdapat beberapa tahapan dalam pra-pemrosesan data yang dilakukan, antara lain:

- Tokenisasi : Proses memecah kalimat menjadi atribut kata (token) dan dimasukkan ke dalam *array*.
- Transform Cases* : Mengubah isi teks menjadi *lowercase* (huruf kecil).
- Stopwords Removal* : Proses penghilangan beberapa kata yang tidak diperlukan atau kurang penting.
- Filter Tokens by length* : Menyaring kata berdasarkan panjang karakternya

Setiap atribut kata yang dihasilkan dari pra-pemrosesan data, berikutnya diberikan pembobotan dengan teknik TF-IDF. Teknik tersebut mengintegrasikan dua konsep perhitungan antara *Term Frequency* dengan *Inverse Document Frequency* guna membantu dalam penentuan apakah kata tersebut relevan dengan dokumen atau tidak, Kata yang paling penting adalah kata yang relevansinya paling tinggi atau skor TF-IDF nya tinggi [9].

### E. Klasifikasi dengan Naive Bayes

Pada tahap ini, dilakukan pemodelan data dengan menerapkan algoritma Naive Bayes Classifier. Metode yang digunakan pada Algoritma ini yaitu peluang dan statistik, untuk kelebihan yang dimiliki algoritma ini diantaranya yaitu bisa diterapkan pada jumlah data yang sedikit, hanya memerlukan waktu perhitungan yang singkat dan dapat digunakan untuk klasifikasi *multiclass data* [10]. Bentuk umum teorema Bayes ditunjukkan pada persamaan (1):

$$P(Y|X) = \frac{P(X|Y).P(Y)}{P(X)} \quad (1)$$

Nilai X pada persamaan (1) merupakan data yang kelas-nya belum ditentukan, sedangkan Y adalah kelas tertentu, dan untuk P (Y|X) merupakan kemungkinan hipotesis yang bergantung pada kondisi X, lalu untuk P(X|Y) adalah probabilitas sebelumnya dari suatu kelas berdasarkan kondisi hipotesis, P(Y) yaitu probabilitas dari hipotesis Y dan untuk P(X) adalah probabilitas dari X [11].

### F. Evaluasi

Implementasi model Naive Bayes menghasilkan sebuah *rules* yang akan digunakan sebagai dasar untuk memprediksi kategori sentimen pada data uji. *Rules* atau model tersebut perlu dilakukan evaluasi sehingga dapat diketahui nilai akurasi dari hasil prediksinya. Evaluasi pada tahap ini dilakukan menggunakan *confusion matrix*. *Confusion Matrix* merupakan alat bantu dalam mengevaluasi performa dari berbagai algoritma dan matrix ini

menghasilkan beberapa istilah dari kelas aktual dan kelas prediksi[12]. Terdapat sembilan istilah hasil klasifikasi yang terdapat pada *confusion matrix* dengan tiga kelas, istilah tersebut ditunjukkan pada Tabel II.

TABEL II  
 CONFUSION MATRIX UNTUK TIGA KELAS

	True Positive	True Negative	True Netral
<b>Pred. Positive</b>	TP	FP-Neg	FP-Net
<b>Pred. Negative</b>	FN-Pos	TN	FN-Net
<b>Pred. Netral</b>	FNet-Pos	FNet-Neg	TNet

Berdasarkan data *confusion matrix* pada Tabel II, nilai akurasi, *recall* dan *precision* diperoleh dengan persamaan (1) sampai (3). Akurasi dihitung untuk mengetahui nilai tingkat keberhasilan algoritma dalam memprediksi secara benar[13], dan *recall* dihitung untuk mengetahui tingkat ketepatan model dalam mengukur pola positif yang diprediksi benar dari total data aktual yang positif sedangkan *precision* untuk mengetahui tingkat ketepatan model dalam mengukur pola positif yang diprediksi benar dari keseluruhan pola yang diprediksi positif [14].

Nilai akurasi didapat dari jumlah tweet yang berhasil diprediksi sesuai dengan kelasnya dibagi dengan total keseluruhan tweet yang diprediksi, dan untuk nilai *recall* didapat dari jumlah sentimen positif yang diprediksi dengan benar dibagi dengan jumlah keseluruhan sentimen yang secara aktual memang positif, sedangkan untuk nilai *precision* diperoleh dari hasil bagi antara jumlah sentimen positif yang diprediksi benar dengan jumlah keseluruhan sentimen yang diprediksi positif.

$$\text{Akurasi} = \frac{TP+TN+TNet}{Total} \times 100\% \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+FNPos+FNetPos} \times 100\% \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FPNeg+FPNet} \times 100\% \quad (3)$$

Pada persamaan (1) terdapat nilai TP (True Positive) yang merupakan tweet berlabel positif dan diprediksi dengan benar, selanjutnya terdapat TN (True Negative) yaitu tweet berlabel negatif yang diprediksi dengan benar, dan TNet (True Netral) merupakan tweet berlabel netral yang diprediksi secara benar, Berikutnya pada persamaan (2) terdapat nilai FN-Pos (False Negative – Positive) yaitu tweet yang berlabel positif namun diprediksi berlabel negatif, sedangkan nilai FNet-Pos (False Netral – Positive) yaitu tweet yang berlabel positif namun diprediksi berlabel netral, Selanjutnya pada persamaan (3) terdapat nilai FP-Neg (False Positive- Negative) yaitu tweet yang secara aktual berlabel negatif namun diprediksi sebagai positif, sedangkan FP-Net (False Positive- Netral) yaitu tweet yang berlabel Netral diprediksi sebagai positif [15].

### III. HASIL DAN PEMBAHASAN

Data tweet yang semula terkumpul sebanyak 20.000, sesudah melalui proses pembersihan data, tersisa sejumlah 3.301 data. Berkurangnya jumlah data disini dikarenakan terdapat proses penghapusan data duplikat, karakter spesial dan tweet yang tidak relevan. Hasil tweet yang sudah bersih dibagi menjadi 2.387 untuk data latih dan 914 untuk data uji. Pada penelitian ini, data latih diberi label secara manual dengan mengkategorikannya ke tiga kelas antara lain positif, netral dan negatif. Pada tahap pra-pemrosesan data dilakukan beberapa sub-proses yang terdiri dari tokenisasi untuk memecah kalimat tweet menjadi kata (token), *transform cases* untuk mengubah seluruh huruf pada tweet menjadi huruf kecil, *filter stopword* untuk membuang sejumlah kata yang tidak digunakan atau kurang penting, dan *filter token by length* untuk membuang kata yang jumlah karakternya  $\leq 3$  dan  $\geq 25$ . Hasil contoh perubahan data tweet sesaat dan sebelum pra-pemrosesan ditunjukkan pada Tabel III.

TABEL III  
 CONTOH HASIL PRA-PEMROSESAN DATA

Tahap Pra-Pemrosesan	Sebelum pra-pemrosesan	Setelah pra-pemrosesan
Data Cleaning	@ABemates @luvv_188 @CelotehAB Iya dong. Kan sdh saya buka jalannya. Silahkan dilanjutkan. Saya butuh uangnya. 2023 resesi, saya tidak peduli. I want my money!! Hahaha	Iya dong Kan sdh saya buka jalannya Silahkan dilanjutkan Saya butuh uangnya resesi saya tidak peduli I want my money Hahaha

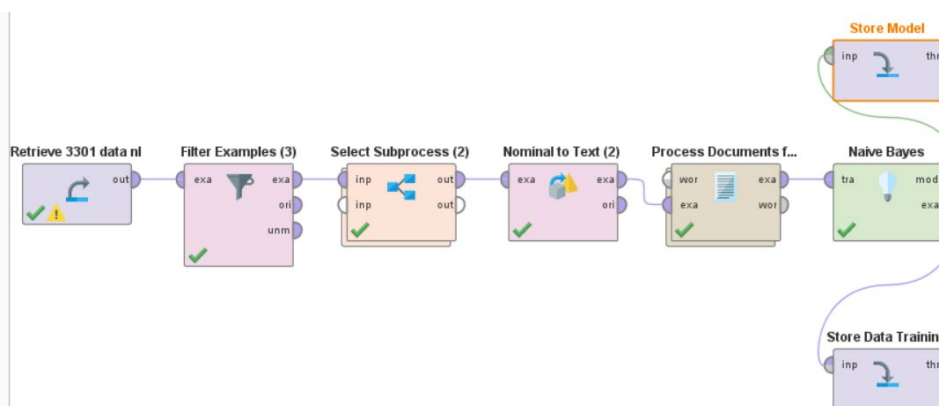
Tokenizing	Iya dong Kan sdh saya buka jalannya Silahkan dilanjutkan Saya butuh uangnya resesi saya tidak peduli I want my money Hahaha	“Iya” “dong” “Kan” “sdh” “saya” “buka” “jalannya” “Silahkan” “dilanjutkan” “Saya” “butuh” “uangnya” “resesi” “saya” “tidak” “peduli I” “want” “my” “money” “Hahaha”
Transform Case	Iya dong Kan sdh saya buka jalannya Silahkan dilanjutkan Saya butuh uangnya resesi saya tidak peduli I want my money Hahaha	iya dong kan sdh saya buka jalannya silahkan dilanjutkan saya butuh uangnya resesi saya tidak peduli i want my money hahaha
Stopword Removal	iya dong kan sdh saya buka jalannya silahkan dilanjutkan saya butuh uangnya resesi saya tidak peduli i want my money hahaha	iya kan sdh buka jalannya silahkan dilanjutkan butuh uangnya resesi tidak peduli i want my money
Filter Token (by length)	iya kan sdh buka jalannya silahkan dilanjutkan butuh uangnya resesi tidak peduli want money	buka jalannya silahkan dilanjutkan butuh uangnya resesi tidak peduli want money

Pada penelitian ini, sejumlah 6.297 atribut kata dihasilkan dari tahap pra-pemrosesan data latih, dan setiap kata diberikan pembobotan menggunakan TF-IDF, tujuannya untuk membantu menentukan relevansi kata dengan dokumen. Nilai pembobotan untuk setiap kata tergantung akan seberapa sering kata tersebut muncul dalam suatu dokumen. Jika frekuensi kemunculan kata-nya tinggi maka nilai bobotnya akan semakin tinggi dan diasumsikan bahwa kata tersebut merupakan kata yang penting, begitupun sebaliknya. Contoh hasil pemberian bobot untuk atribut kata yang ada ditunjukkan pada Tabel IV.

TABEL IV  
 CONTOH PEMBOBOTAN KATA DENGAN TF-IDF

Teks	antisipasi	bikin	bukti	hadapin	kerja	nyata	strategi	...
bukti nyata kerja ganjar antisipasi resesi	0,419	0	0,578	0	0,291	0,390	0	
bikin strategi hadapin resesi	0	0,392	0	0,745	0	0	0,540	

Algoritma Naive bayes selanjutnya diterapkan ke data latih yang sudah melalui tahap pra-pemrosesan data dan pembobotan. Pada proses ini, pola-pola data yang ada akan dipelajari dan nantinya akan dihasilkan sebuah model klasifikasi atau *rules* dan juga data hasil pelatihan. Model dan data pelatihan ini digunakan untuk mengklasifikasikan data uji secara otomatis. Keseluruhan proses pelatihan model pada penelitian ini dapat dilihat pada Gambar 2.



Gambar. 2. Proses Pembuatan Model

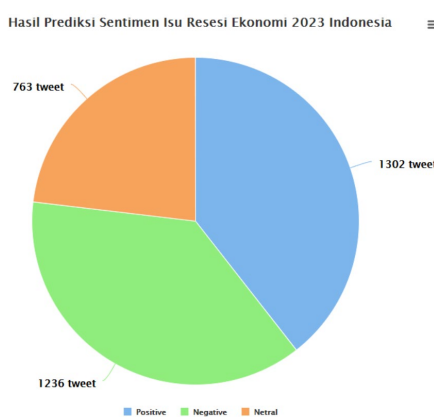
Performa dari model klasifikasi yang sudah dibuat, perlu dievaluasi terlebih dahulu untuk mengetahui seberapa baik model dalam memprediksi kelas sentimen. Performa yang dihitung antara lain akurasi, *precision* dan *recall* dan nilai tersebut didapatkan dengan metode *confusion matrix*. Hasil data *confusion matrix* pada penelitian ini ditunjukkan pada Tabel V.



TABEL V  
 HASIL PERFORMANCE MODEL

	True Positive	True Negative	True Netral	Precision
Prediction Positive	958	15	0	98.46%
Prediction Negative	8	885	0	
Prediction Netral	67	66	388	
Recall	92.74%			
Accuracy	93.46%			

Berdasarkan data performa model pada Tabel V, dilakukan perhitungan dan didapatkan nilai akurasi model klasifikasi sebesar 93,46%, *precision* 98.46% dan *recall* 92.74%. Tahap berikutnya yaitu proses penerapan model klasifikasi pada data uji, hal ini berguna untuk membuktikan apakah model memang akurat di kenyataannya atau tidak. Proses pengujian akan memperoleh hasil yaitu label sentimen yang diklasifikasikan secara otomatis.



Gambar. 3. Diagram Hasil Prediksi Label Sentimen Isu Resesi 2023 Indonesia

Gambar 3 menunjukkan distribusi label untuk data tweet isu resesi ekonomi 2023 Indonesia yang telah diklasifikasikan secara otomatis. Terlihat bahwa sentimen isu resesi ekonomi di Indonesia tahun 2023 didominasi oleh sentimen positif sebanyak 1302 tweet, sedangkan untuk sentimen negatif sejumlah 1236 dan sentimen netral 763 tweet.

Penelitian ini menemukan beberapa kata yang sering muncul dalam data tweet isu resesi ekonomi di Indonesia seperti yang ditunjukkan pada Gambar 4, diantaranya yaitu "ancaman", "negara", "semoga", "inflasi", "bahas", "takut", "uang", "harga", "kerja", "investasi" dan lainnya. Berdasarkan kata-kata tersebut dapat diasumsikan bahwa terdapat respon masyarakat Indonesia yang menganggap resesi ini adalah suatu ancaman bagi negara dan menjadikan ini sebagai peristiwa yang menakutkan baik dari segi kenaikan harga, kondisi keuangan, dan lainnya. Namun beberapa respon masyarakat lainnya saat diperhadapkan dengan isu resesi ekonomi di tahun 2023 masih terdapat yang optimis, memiliki harapan dan melakukan berbagai persiapan untuk bisa bertahan seperti dengan cara bekerja, berinvestasi, menabung dan lainnya. Berdasarkan hasil penelitian ini terhadap penelitian terdahulu yang dilakukan oleh Romaito, dkk[6], ditemukan bahwa Algoritma Naive Bayes pada kedua penelitian tersebut menghasilkan akurasi yang tinggi dan disimpulkan bahwa Algoritma ini cukup baik dalam menganalisis sentimen, baik dengan menggunakan jumlah data latihan yang kecil maupun besar.



Gambar. 4. Wordcloud Sentimen Isu Resesi 2023 Indonesia

#### IV. KESIMPULAN

Pada penelitian ini, algoritma Naive Bayes berhasil mengklasifikasikan sentimen isu resesi ekonomi 2023 di Indonesia ke kelas positif, netral dan negatif. Data diperoleh dengan teknik scraping menggunakan kata kunci #Resesi 2023 dan Resesi Ekonomi 2023 Indonesia dalam periode waktu tweet sejak tanggal 1 Juli 2022 – 21 April 2023. Akurasi model klasifikasi yang dihasilkan pada penelitian ini mencapai 93,46%. Hasil analisa menunjukkan bahwa sentimen isu resesi ekonomi 2023 di Indonesia pada periode tersebut didominasi dengan tweet positif sebanyak 1.301 tweet, sedangkan sentimen negatif sebanyak 1.237 tweet dan sentimen netral sejumlah 763 tweet. Berdasarkan jumlah prediksi dan melihat *wordcloud* yang ada, dapat diasumsikan bahwa saat diperhadapkan dengan isu resesi 2023, lebih banyak masyarakat yang optimis, memiliki harapan dan melakukan berbagai persiapan untuk bertahan seperti dengan cara bekerja, berinvestasi, dan menabung. Diharapkan hasil dari analisa penelitian ini dapat membantu pemerintah dan peneliti terkait dalam melihat respon mengenai Isu Resesi Ekonomi 2023 di Indonesia dari masyarakat yang aktif. Saran untuk penelitian serupa kedepannya perlu dilakukan peningkatan dalam tahap pra-pemrosesan teks, dikarenakan masih didapatkan kata yang tidak baku dan beberapa singkatan kata pada data tweet sehingga hal ini bisa menghambat jalannya proses klasifikasi sentimen.

#### DAFTAR PUSTAKA

- [1] N. Laoli, "Ekonom Sebut Perang Rusia dan Ukraina Penyebab Pelemahan Ekonomi Global," Oct. 04, 2022. <https://nasional.kontan.co.id/news/ekonom-sebut-perang-rusia-dan-ukraina-penyebab-pelemahan-ekonomi-global> (accessed Dec. 15, 2022).
- [2] H. Alexander, "Penjelasan Sederhana soal Penyebab Resesi Ekonomi Global 2023." <https://tirtoid.penjelasan-sederhana-soal-penyebab-resesi-ekonomi-global-2023-gwZN> (accessed Dec. 14, 2022).
- [3] M. Zaskya, A. Boham, L. Jackelin, and H. Lotulung, "Twitter Sebagai Media Mengungkapkan Diri Pada Kalangan Milenial," 2021. Accessed: Jun. 01, 2023. [Online]. Available: <https://ejournal.unsrat.ac.id/v3/index.php/actadiumakomunikasi/article/download/31979/30381/66820>
- [4] Kim, Han, Rim, and Myaeng, "Some Effective Techniques for Naive Bayes Text Classification.," *IEEE Trans Knowl Data Eng*, vol. 18, no. 11, pp. 1457–1466, 2006, Accessed: Feb. 03, 2022. [Online]. Available: [https://www.researchgate.net/publication/3297622\\_Some\\_Effective\\_Techniques\\_for\\_Naive\\_Bayes\\_Text\\_Classification](https://www.researchgate.net/publication/3297622_Some_Effective_Techniques_for_Naive_Bayes_Text_Classification)
- [5] N. M. A. J. Astari, D. G. H. Divayana, and G. Indrawan, "Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier," *Jurnal Sistem dan Informatika (JSI)*, vol. 15, no. 1, pp. 27–29, Nov. 2020, doi: 10.30864/jsi.v15i1.332.
- [6] E. S. Romaito and dkk, "Perbandingan Algoritma SVM Dan NBC Dalam Analisa Sentimen Pilkada Pada Twitter," *CSRID Journal*, vol. 13, no. 3, pp. 169–179, 2021, Accessed: Feb. 15, 2022. [Online]. Available: <http://repository.sar.ac.id/id/eprint/60/>.
- [7] F. S. Darusman, A. A. Arifiyanti, and S. F. A. Wati, "Sentiment Analysis Pedulilindungi Tweet Using Support Vector Machine Method," *Applied Technology and Computing Science Journal*, vol. 4, no. 2, pp. 113–118, May 2022, doi: 10.33086/atcsj.v4i2.2836.
- [8] Suripto, Rr Nurul Rahmanita, and Ajeng Sekar Kirana, "Teknik pre-processing dan classification dalam data science." <https://mie.binus.ac.id/2022/08/26/teknik-pre-processing-dan-classification-dalam-data-science/> (accessed Apr. 29, 2023).
- [9] W. Scott, "TF-IDF from scratch in python on a real-world dataset," Feb. 15, 2019. <https://towardsdatascience.com/tf-idf-for-document-ranking-from-scratch-in-python-on-real-world-dataset-796d339a4089> (accessed May 30, 2023).
- [10] Mochammad Haldi Widianto, "Algoritma Naive Bayes." <https://binus.ac.id/bandung/2019/12/algoritma-naive-bayes/> (accessed May 01, 2023).
- [11] Y. Pratama, A. Roberto Tampubolon, L. Diantri Sianturi, R. Diana Manalu, and D. Friez Pangaribuan, "Implementation of Sentiment Analysis on Twitter Using Naive Bayes Algorithm to Know the People Responses to Debate of DKI Jakarta Governor Election," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Jun. 2019. doi: 10.1088/1742-6596/1175/1/012102.
- [12] I. Markoulidakis, I. Rallis, I. Georgoulas, G. Kopsiaftis, A. Doulamis, and N. Doulamis, "Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem," *Technologies (Basel)*, vol. 9, no. 4, Dec. 2021, doi: 10.3390/technologies9040081.
- [13] D. A. Wulandari, R. Rohmat Saedudin, and R. Andreswari, "ANALISIS SENTIMEN MEDIA SOSIAL TWITTER TERHADAP REAKSI MASYARAKAT PADA RUU CIPTA KERJA MENGGUNAKAN METODE KLASIFIKASI ALGORITMA NAIVE BAYES ANALYSIS OF TWITTER SOCIAL MEDIA SENTIMENT ON THE PUBLIC'S REACTION TO THE DRAFTS OF JOB CREATION LAW USING THE CLASSIFICATION METHOD NAIVE BAYES ALGORITHM," *e-Proceeding of Engineering*, vol. 8, no. 5, pp. 9007–9016, 2021.
- [14] M. Hossin and M. N. Sulaiman, "A Review on Evaluation Metrics for Data Classification Evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, no. 2, pp. 01–11, Mar. 2015, doi: 10.5121/ijdkp.2015.5201.
- [15] D. Normawati and S. A. Prayogi, "Implementasi Naive Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021.