

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**KLASIFIKASI SENTIMEN MASYARAKAT DI TWITTER  
TERHADAP ANCAMAN RESESI EKONOMI 2023 DENGAN  
METODE K-NEAREST NEIGHBOR**

**TUGAS AKHIR**

Disusun Sebagai Salah Satu Syarat  
Untuk Memperoleh Gelar Sarjana Teknik  
Pada Jurusan Teknik Informatika

Oleh

**DIMAS FERARIZKI**

**NIM. 11950113420**



UIN SUSKA RIAU

UIN SUSKA RIAU

**FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU**

**PEKANBARU**

**2023**

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**LEMBAR PERSETUJUAN**

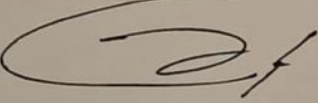

**KLASIFIKASI SENTIMEN MASYARAKAT DI TWITTER  
TERHADAP ANCAMAN RESESI EKONOMI 2023 DENGAN  
METODE K-NEAREST NEIGHBOR**

**TUGAS AKHIR**

Oleh

**DIMAS FERARIZKI**  
NIM. 11950113420

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir  
di Pekanbaru, pada tanggal 7 Desember 2023

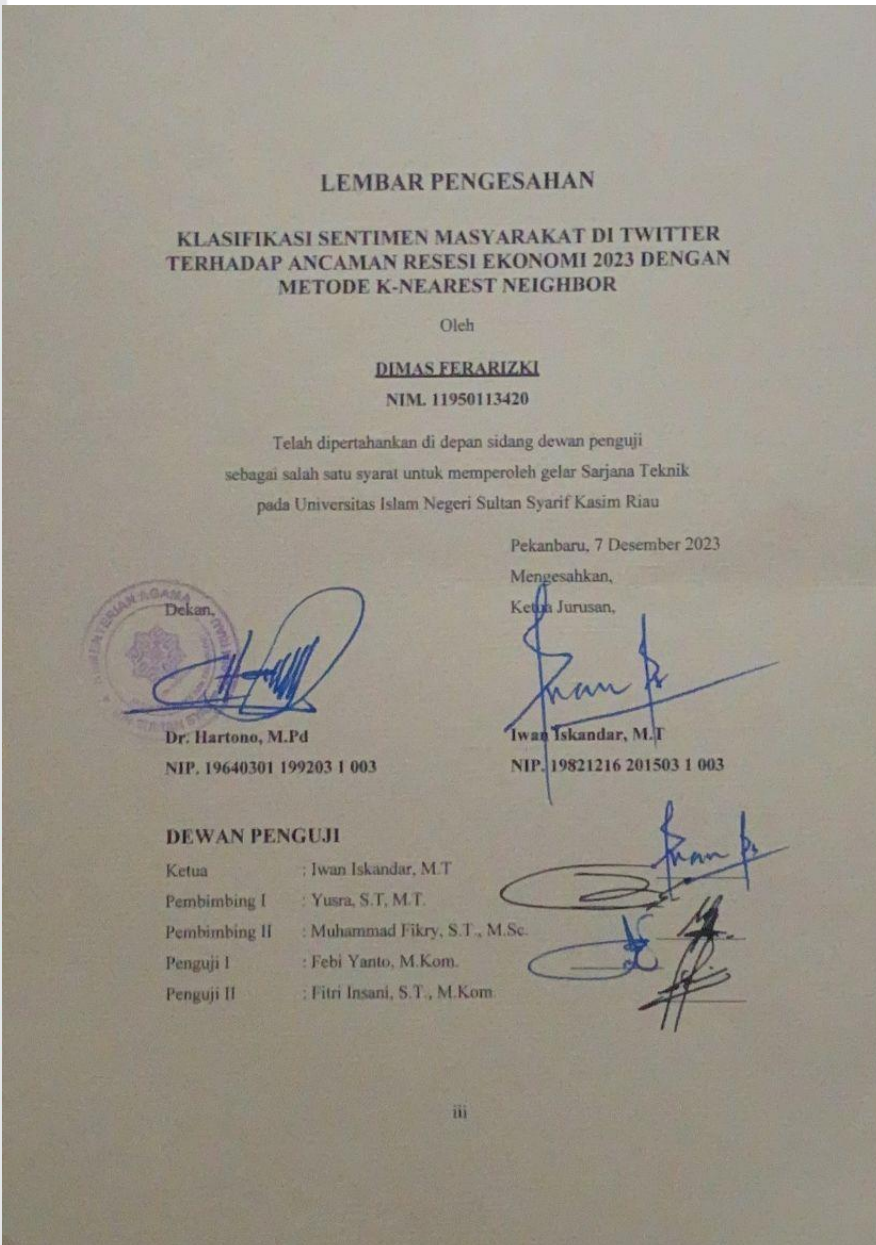
Pembimbing I,	Pembimbing II,
	
<b>Yusra, S.T, M.T.</b> NIP. 198401232015032001	<b>Muhammad Fikry, S.T, M.Sc.</b> NIP. 198010182007101002

ii



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**SURAT PERNYATAAN**

Saya yang bertandatangan di bawah ini:

Nama : Dimas Ferarizki  
NIM : 11950113420  
Tempat, Tgl. Lahir : Pekanbaru, 19 Oktober 2001  
Fakultas : Sains dan Teknologi  
Prodi : Teknik Informatika  
Judul Jurnal :

**Klasifikasi Sentimen Masyarakat Di Twitter Terhadap Ancaman Resesi Ekonomi Tahun 2023 Dengan Metode K-Nearest Neighbor**

Menyatakan dengan sebenar – benarnya bahwa:

1. Penulisan jurnal dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu jurnal saya ini, saya nyatakan bebas dari plagiat.
4. Apabila dikemudian hari terbukti terdapat plagiat dalam penulisan jurnal saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang – undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 19 Desember 2023  
Yang membuat pernyataan



**Dimas Ferarizki**  
NIM. 11950113420

## Klasifikasi Sentimen Masyarakat di Twitter Terhadap Ancaman Resesi Ekonomi 2023 dengan Metode K-Nearest Neighbor

Dimas Ferarizki\*, Yusra, Muhammad Fikry, Febi Yanto, Fitri Insani

Sains dan Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: <sup>1,\*</sup>11950113420@students.uin-suska.ac.id, <sup>2</sup>yusra@uin-suska.ac.id, <sup>3</sup>muhammad.fikry@uin-suska.ac.id,

<sup>4</sup>febiyanto@uin-suska.ac.id, <sup>5</sup>fitri.insani@uin-suska.ac.id

Email penulis Korespondensi: 11950113420@students.uin-suska.ac.id

**Abstrak**—Resesi adalah terjadinya penurunan dalam aktivitas ekonomi yang secara menyeluruh, hal ini dianggap sebagai fase penurunan ekonomi yang signifikan dan berkelanjutan dari berbagai sektor dan indikator ekonomi. Ancaman resesi pada tahun 2023 telah menjadi bahan pembicaraan di banyak negara, termasuk di Indonesia. Hal ini terjadi karena di Indonesia terancam sebagai negara yang terdampak dari resesi akibat dari melemahnya aktifitas ekonomi pada sektor riil. Penelitian klasifikasi sentimen ini bertujuan untuk menganalisis pendapat serta opini masyarakat terkait isu berita resesi tahun 2023 yang disampaikan melalui platform media sosial twitter dan memahami apakah pendapat-pendapat tersebut termasuk dalam kategori sentimen positif atau negatif. Selain itu, penelitian ini juga bertujuan untuk mengukur tingkat akurasi dalam mengklasifikasikan sentimen-sentimen tersebut ke dalam kelas yang sesuai. Ada beberapa proses utama yang dimulai dari pengumpulan data kemudian pelabelan data secara manual, *text processing*, pembobotan fitur (TF-IDF), seleksi fitur *thresholding* dan klasifikasi metode *K-Nearest Neighbor*. Berdasarkan hasil klasifikasi menggunakan model pengujian dari total 1000 data komentar yang terbagi antara 596 data kelas positif dan 404 data kelas negatif terhadap ancaman resesi tahun 2023, diperoleh hasil akurasi tertinggi sebesar 85% pada nilai  $k = 3$  dengan menggunakan model perbandingan 90:10 *training* dan *testing* data.

**Kata Kunci:** Resesi; Klasifikasi Sentimen; K-Nearest Neighbor; Twitter; Ekonomi.

**Abstract**—A recession is a decline in overall economic activity, this is considered a phase of significant and sustainable economic decline in various sectors and economic indicators. The threat of a recession in 2023 has become a topic of discussion in many countries, including Indonesia. This happens because Indonesia is threatened as a country affected by a recession due to weakening economic activity in the real sector. This sentiment classification research aims to analyze public opinion and opinion regarding the issue of recession news in 2023 which is conveyed via the social media platform Twitter. This research aims to understand whether these opinions fall into the category of positive sentiment or negative sentiment. Apart from that, this research also aims to measure the level of accuracy in classifying these sentiments into appropriate classes. This research has several main processes starting from data collection then manual data labeling, text processing, feature weighting (TF-IDF), Thresholding feature selection and K-Nearest Neighbor method classification. Based on the classification results using a testing model from a total of 1000 comment data divided between 596 positive class data and 404 negative class Twitter data regarding the threat of recession in 2023, the highest accuracy results were obtained at 85% at a value of  $k = 3$  using the 90:10 comparison model training and testing data.

**Keywords:** Recession; Sentiment Classification; K-Nearest Neighbor; Twitter; Economy.

### 1. PENDAHULUAN

Resesi merupakan ketika pertumbuhan ekonomi riil mengalami pertumbuhan negatif atau produk domestik bruto menyusut dua kuartal dalam satu tahun [1]. Resesi menyebabkan perubahan yang signifikan dalam berbagai bidang yang membuat laju pertumbuhan ekonomi melambat seperti di Indonesia yang termasuk negara berkembang [2]. Indonesia pernah mengalami beberapa ancaman ekonomi seperti krisis ekonomi pada tahun 1998, dampak krisis global pada tahun 2008, serta ancaman resesi ekonomi pada tahun 2020 karena penyebaran Covid-19 [3]. Masa pandemi Covid-19 menghambat laju pertumbuhan ekonomi Indonesia pada tahun 2020 di angka -5.32 % dan setelah pandemi dihadapi dengan ancaman resesi global [4], [5].

Menurut Badan Pusat Statistik meskipun terdapat tekanan inflasi global dan ancaman resesi, ekonomi Indonesia tumbuh dengan cukup baik atau kembali normal menyentuh angka 5,44 % pada Triwulan II tahun 2022 [4]. Dengan pertumbuhan ekonomi di Indonesia terbilang masih cukup terlindungi dari ancaman resesi tetapi risiko resesi ekonomi masih ada dengan tingkat risiko sebesar 3 % [6].

Isu ancaman resesi ekonomi menimbulkan berbagai pendapat atau sudut pandang dari masyarakat, pendapat masyarakat terhadap sebuah topik atau isu dapat dilihat melalui media sosial yang banyak digunakan, salah satunya media sosial seperti Twitter. Dikutip dari <https://dataindonesia.id/> di Indonesia jumlah pengguna Twitter pada tahun 2022 mencapai 18,45 juta pengguna. Twitter menjadi indikasi kemampuan masyarakat dalam menyampaikan, menerima, dan menyebarkan pendapat atau informasi kepada pengguna lain sehingga informasi tersebut dapat diketahui oleh masyarakat luas [7]. Banyaknya cuitan dan opini tentang ancaman resesi ekonomi di Indonesia, maka dapat dimanfaatkan untuk memperoleh informasi polaritas sentimen positif, negatif maupun netral terhadap hal tersebut. klasifikasi sentimen adalah satu diantara metode lainnya yang dapat digunakan untuk menganalisa *tweet* atau opini tersebut [8].

Klasifikasi sentimen merupakan suatu ilmu yang mencari dan mengelompokkan emosi ataupun opini masyarakat baik berupa teks maupun vocal [8] Klasifikasi sentimen adalah proses mengidentifikasi perilaku atau pendapat seseorang ketika dihadapi suatu topik atau kejadian dengan mengkategorikan polaritas teks, yaitu apakah teks tersebut mengandung sentimen positif, negatif, atau netral. Salah satu teknik klasifikasi yang paling populer adalah dengan metode *K-Nearest Neighbor* (K-NN) [9]. Metode ini termasuk dalam *lazy learning* yang pengelompokan data masuk ke dalam kelas-kelas



yang telah diprediksi sebelumnya berdasarkan jarak terdekat atau tingkat kesamaannya dengan dataset atau data pelatihan yang ada sebelumnya [9], [10]. Penelitian ini bertujuan untuk mengklasifikasikan sentimen masyarakat terhadap ancaman resesi ekonomi tahun 2023 pada media sosial Twitter dengan menggunakan metode *K-Nearest Neighbor* (K-NN) dalam jumlah yang besar secara otomatis.

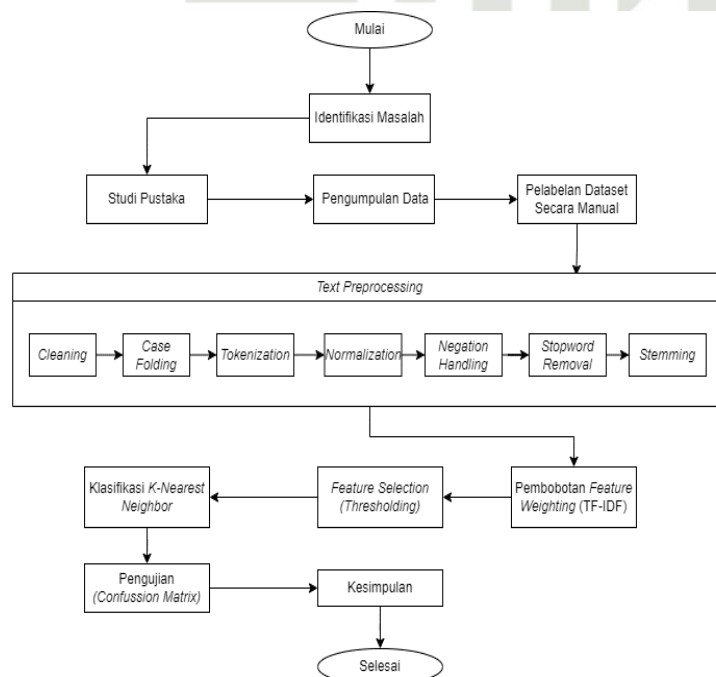
Penelitian sebelumnya oleh R. Gunawan, R. Septiadi, F. Apri Wenando, H. Mukhtar, dan Syahril meneliti tentang mengklasifikasi nilai positif dan negatif dari komentar masyarakat pada pengguna Twitter tentang kebijakan merdeka belajar kampus merdeka. Hasil pengujian yang menggunakan 700 data menunjukkan bahwa metode K-NN dengan perbandingan 90:10 *testing* dan *training* pada  $k=8$  menghasilkan akurasi sebesar 84,28%, *precision* 84,42%, *recall* 84,28%, dan *f-measure* 84,35% [11]. Kemudian penelitian lainnya dilakukan oleh D. Muhidin and A. Wibowo, yang membahas tentang mengeksplorasi perasaan masyarakat tentang “*New Normal*” selama pandemi dengan dua metode yang akan dibandingkan yaitu *Support Vector Machine* dan *K-Nearest Neighbor*. Hasilnya menunjukkan bahwa K-NN lebih efektif dalam menyelesaikan topik sentimen ini yang menghasilkan akurasi sebesar 72,96%, sedangkan SVM hanya mencapai 71,19% [12]

Penggunaan teknik *K-Nearest Neighbor* (K-NN) dalam penelitian sebelumnya dengan pendekatan *Ensemble* untuk mengategorikan sentimen terkait Kurikulum 2013 di platform Twitter, efek dari nilai  $k$  dan kombinasi fitur sangat berpengaruh terhadap akurasi klasifikasi sentimen ditentukan. Hasilnya menunjukkan bahwa nilai  $k$  optimal untuk K-NN adalah 5, dengan akurasi 96% [13]. Kemudian penelitian [14] menguji efektivitas penerapan metode *Naïve Bayes*, SVM, dan K-NN untuk menganalisis sentimen pada platform Halodoc. Dataset yang digunakan terdiri dari 1000 ulasan positif dan negatif dalam bahasa Inggris yang diambil dari *Google Play Store*. Hasil dari penelitian ini menunjukkan bahwa K-NN memberikan hasil terbaik dan paling akurat untuk klasifikasi sentimen, dengan akurasi tertinggi 95,00% dan nilai AUC = 0,985. Penelitian [15] mengevaluasi klasifikasi sentimen negatif dan positif dengan menerapkan metode *K-Nearest Neighbor* (K-NN) pada topik Pembatasan Sosial Berskala Besar (PSBB) di Twitter. Dalam penelitian ini, K-NN digunakan untuk mengevaluasi nilai akurasi dari 3 jenis perhitungan jarak yaitu *Cosinus Similarity*, *Euclidean Distance*, dan *Manhattan Distance*. Hasil penelitian menunjukkan bahwa K-NN dengan jarak *Cosine Similarity* mendapatkan nilai *accuracy* 82% pada  $k = 3$ , K-NN dengan *Euclidean Distance* dan menghasilkan akurasi sebesar 81% pada  $k=11$ , dan K-NN dengan jarak *Manhattan Distance* yang menghasilkan akurasi 80% pada  $k = 5,7,9,11$  dan 13.

Dengan beragam perbedaan dari penelitian sebelumnya, penelitian kali ini mencoba untuk menggunakan metode K-NN dengan kasus yang berbeda. Adapun data yang digunakan berjumlah 1000 data *tweet* atau komentar yang diambil dari Twitter berdasarkan kata kunci “resesi”. Pada penelitian ini akan dilakukan beberapa tahapan untuk klasifikasi terhadap sentimen masyarakat di Twitter terhadap ancaman resesi ekonomi tahun 2023. Dimulai dari tahapan *preprocessing* sampai tahap klasifikasi sentimen, *feature weighting* (TF-IDF) dan *feature selection* (*Thresholding*) dapat menentukan sentimen positif atau negatif terhadap ancaman resesi ekonomi tahun 2023.

## 2. METODOLOGI PENELITIAN

Dalam penelitian ini, terdapat sebuah kerangka kerja penelitian yang melibatkan beberapa langkah. Langkah-langkah penelitian ini dapat dirinci seperti yang ditunjukkan dalam Gambar 1.



Gambar 1. Metodologi Penelitian



## 2.1 Identifikasi Masalah

Langkah awal yang penting dalam proses penelitian adalah menentukan identifikasi dari suatu masalah, karena permasalahan yang terdefinisi dengan baik akan membimbing seluruh penelitian dan memastikan tujuan penelitian dapat tercapai.

### 2.1.1 Studi Pustaka

Studi pustaka sangat identik dengan proses pencarian dan pengumpulan literatur yang berhubungan dengan topik penelitian. Tujuannya adalah untuk mengumpulkan informasi, teori, temuan penelitian terdahulu, dan pandangan yang telah ada dalam literatur terkait.

### 2.1.2 Pengumpulan Data

Langkah awal dalam metodologi penelitian ini adalah pengumpulan data. Data yang diambil bersumber dari platform media sosial Twitter dengan menggunakan kata kunci "resesi". Total data cuitan *tweet* yang berhasil dikumpulkan berjumlah 3000 data. Data tersebut dikumpulkan terhitung dari tanggal 21 Oktober sampai 6 Desember 2022, namun untuk penelitian kali ini hanya menggunakan 1000 data.

### 2.1.3 Pelabelan Dataset

Tahap dari tahap ini adalah untuk menentukan kategori sentimen positif atau negatif dalam data yang telah terkumpul. Proses pelabelan ini dilakukan dengan tujuan mengatasi tantangan dalam mengklasifikasikan data *tweet* sehingga dapat dikelompokkan dengan akurat sesuai dengan kontennya. Tujuan utama dari tahap ini adalah untuk mengidentifikasi apakah sentimen yang terkandung dalam data yang telah terkumpul adalah positif atau negatif.

### 2.2 Text Preprocessing

Tahapan ini sangat berpengaruh dalam proses klasifikasi, dimana dataset akan disaring dan disiapkan untuk menggabungkan kata-kata yang mirip sehingga dapat mengurangi jumlah kata dalam kumpulan dokumen. Hal ini bertujuan untuk mempermudah analisis dan pengolahan data lebih lanjut.

## 2.6 Pembobotan Feature Weighting (TF-IDF)

Pembobotan TF-IDF (*Term Frequency Inverse Document Frequency*) digunakan untuk mengidentifikasi kata-kata kunci yang paling relevan dalam sebuah dokumen [16]. Semakin tinggi nilai TF-IDF suatu kata dalam dokumen, semakin penting kata tersebut dalam konteks dokumen tersebut. Kemudian TF-IDF berfungsi mengolah teks untuk menentukan bobot pentingnya sebuah kata dalam suatu atau kumpulan dokumen. Hasil bobot tersebut digunakan untuk menghitung jarak terpendeknya dengan menggunakan rumus *Euclidean Distance*. Dalam tahap pembobotan kata, langkah-langkah berikut dijalankan :

- Setiap kata pada seluruh dokumen akan dihitung frekuensi kemunculannya.
- Hitung nilai IDF menggunakan rumus :

$$IDF = \log \frac{d}{df} \quad (1)$$

Keterangan:

- $IDF$  = *Inverse Document Frequency*  
= Jumlah dokumen secara keseluruhan  
 $df$  = Jumlah dokumen yang mengandung kata-kata tertentu (*term*)

- Kemudian semua kata pada dokumen dihitung bobotnya dengan menggunakan rumus :

$$W_{dt} = tf_t \times IDF_t \quad (2)$$

Keterangan :

- $W_{dt}$  = Bobot kata "t" dalam dokumen "d."  
 $tf_t$  = Frekuensi kemunculan kata "t" dalam dokumen "d" (*Term Frequency*).  
 $IDF_t$  = Bobot IDF yang telah dihitung sebelumnya

- Hasilnya digunakan sebagai patokan untuk *vector* sebelum perhitungan jarak dilakukan dengan *Euclidean Distance* :

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3)$$

Keterangan :

- $d_{ij}$  = Jarak antar i dan j  
 $x_i$  = Kata ke i di dokumen x  
 $y_i$  = Kata ke i di dokumen y

## 2.7 Seleksi Fitur (*Thresholding*)

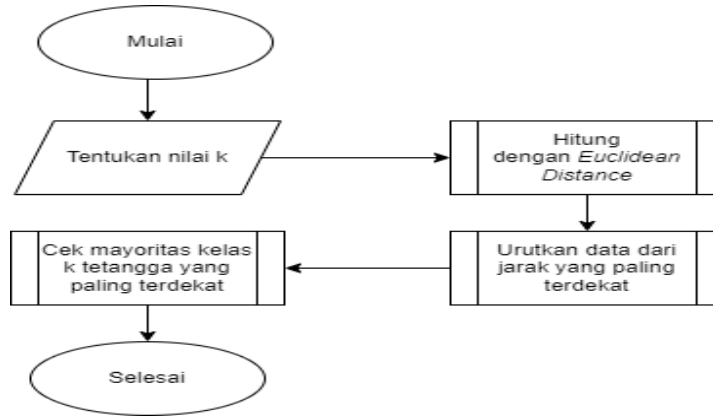
2. Diarangi mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Langkah ini bertujuan untuk mengurangi jumlah dari variabel yang dimasukkan sampai tersisa variabel yang paling bernilai. Dalam penelitian ini, seleksi fitur dilakukan dengan menggunakan nilai ambang batas yang relevan untuk memilih kategori fitur yang optimal [17]. Evaluasi pengaruh jumlah *term* yang dipilih pada proses klasifikasi merupakan faktor dari uji seleksi fitur.

**Klasifikasi K-Nearest Neighbor**

Metode ini adalah salah satu metode yang digunakan untuk melakukan klasifikasi. Metode ini dengan mengidentifikasi jarak terdekat antara sampel pelatihan dan instans permintaan untuk menemukan tetangga terdekat [11]. Kelas yang paling sering muncul di antara tetangga-tetangga tersebut akan menjadi hasil klasifikasi. Tahapan-tahapan metode K-NN dapat dilihat pada Gambar 2.



Gambar 2. Klasifikasi K-Nearest Neighbor

**2.4 Pengujian (Confusion Matrix)**

*Confusion matrix* digunakan bertujuan untuk memahami dan mengevaluasi performa dari metode klasifikasi pada penelitian ini. *Confusion matrix* memperhitungkan *accuracy*, *precision*, *recall*, dan *f1 score* untuk mengevaluasi performa model klasifikasi yang dibuat. Tabel 1 berikut adalah penjelasan dari pengujian *confusion matrix*.

Tabel 1. Confusion Matrix

Predicted Class	Positive	Negative
Positive	TP (True Positive)	FP (False Positive)
Negative	FN (False Negative)	TN (True Negative)

Parameter yang digunakan adalah *accuracy*, *precision*, *recall* dan *f1 score*.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \tag{4}$$

$$Precision (P) = \frac{TP}{TP+FP} \times 100\% \tag{5}$$

$$Recall (R) = \frac{TP}{TP+FN} \times 100\% \tag{6}$$

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision+Recall} \times 100\% \tag{6}$$

**2.10 Kesimpulan**

Puncak dari hasil penelitian terletak pada bagian kesimpulan. Sebagai alternatif, rekomendasi yang penting berkaitan dengan studi ini disertakan dalam saran, dengan tujuan untuk memfasilitasi peluang pengembangan yang lebih maju di masa depan dan meraih hasil yang lebih efektif.

**3. HASIL DAN PEMBAHASAN**

**3.1 Pengumpulan Data**

Proses pengumpulan data terbatas pada topik dan pandangan masyarakat tentang ancaman resesi yang hanya terdapat di platform media sosial Twitter. Metode yang digunakan adalah *crawling* Twitter, menggunakan fungsi *Application Interface (API)* yang Twitter sediakan. Hasil data yang terkstraksi berjumlah 3000 data namun penelitian kali ini menggunakan 1000 data.





A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Dateime	Tweet Id	Location	Username	Text										
2	2022-12-0	1,60E+32	Jakarta	ivooxid	b'Ketakutan Resesi, Wall Street Lanjutkan Penurunan https://t.co/HFr195qJ0U'										
3	2022-12-0	1,60E+32		Ferdinand	b'Ketahanan Pangan Desa menjadi Kekuatan Indonesia menghadapi Resesi global....'										
4	2022-12-0	1,60E+32	Jakarta, In	ArisB5758	b'2023 semoga ekonomi Indonesia tetap tumbuh dan tahan terhadapa terpaan angin resesi dunia. Kemandirian serta Kedaulatan										
5	2022-12-0	1,60E+32		SPedalang	b'Caffe type UKM ramping dan kecil justru penyelamat ekonomi saat terjadi resesi. https://t.co/Ox1mg7w8u1'										
6	2022-12-0	1,60E+32		farisElsbat	b'Yang bilang Indonesia 2023 bebas resesi ga sadar kali ya Januari tanggal merahnya Hari minggu. https://t.co/h9QOUTglOy'										
7	2022-12-0	1,60E+32		beritaekor	b'Indonesia Kebal Resesi, Tren IPO Propektif 2023 https://t.co/b0oUNpDaIT'										
8	2022-12-0	1,60E+32	My Cozy	Cantisontol	b'@CNNIndonesia tanda2 resesi ? \n\nIndonesia itu kekayaan alamnya banyak, kayak gas yang lagi dibutuhkan, ada triliunan kut										
9	2022-12-0	1,60E+32		SPedalang	b'Sebentar lagi kita memasuki tahun 2023. Jangan sampai menjadi tahun kalabendu. Karena diramalkan antar anak bangsa ter										
10	2022-12-0	1,60E+32	Medan Ko	yunanking	b'AMERIKA SERIKAT RESESI TAHUN AWAL 2023 BAHAYA UNTUK INDONESIA. JADI HARUS BANYAK-BANYAK TAHAN EKSPOR KI										
11	2022-12-0	1,60E+32		PDI_	Perjuangan meski dunia mengalami resesi ekonomi.insya allah indonesia sudah mempunyai perisai. https://t.co/u952										
12	2022-12-0	1,60E+32	DKI Jakart	adibapray	b'Menurut LaNyalla, peluang tersebut juga akan mampu menjaga pertumbuhan dan stabilitas ekonomi nasional. Otomatis der										
13	2022-12-0	1,60E+32	DKI Jakart	adibapray	b'Namun, Ketua DPD RI, AA LaNyalla Mahmud Mattalitti, menilai ancaman resesi merupakan tantangan sekaligus peluang bes										
14	2022-12-0	1,60E+32		DKI Jakart	adibapray b'Ketua DPD RI Nilai Resesi Dunia Bisa Jadi Peluang bagi Indonesia. #LaNyalla #Nyalaindonesia #BeraniBangkit #DPDRI\n\nhttt										
15	2022-12-0	1,60E+32	South-sule	hariadyrig	b'Indonesia gelap 2023? Biasanya disetiap resesi akan lahir orang miskin baru dan orang kaya baru\x0\x9f\x98\x81 dan orang										
16	2022-12-0	1,60E+32		syakilahsy	b'Bank Rakyat Indonesia (Persero) Tbk menegaskan komitmennya untk terus mendorong pertumbuhan ekonomi domestik di t										
17	2022-12-0	1,60E+32	DKI Jakart	derasdotic	b'Bank Indonesia (BI) memprediksi pertumbuhan ekonomi Indonesia mampu mencapai 4,7% sampai 5,5% pada tahun 2024. G										
18	2022-12-0	1,60E+32	INDONESIA,	mediatata	b'Cara Jejak Bumi Indonesia (JBI) Hadapi Ancaman Resesi Ekonomi Global 2023\n\nhttps://t.co/Bs8xjzYch7 https://t.co/XqCeEaC										
19	2022-12-0	1,60E+32	Indonesia	Muammaib	b'Seorang @erickthohir aja optimis kok kalo Indonesia bisa lewati resesi di 2023 #BangkitBersamaET https://t.co/ppd1Y4fjR'										
20	2022-12-0	1,60E+32	Solo. Cent	soloosodo	b'Pengamat: Indonesia Lavak Optimistis di Tengah Ancaman Resesi Ekonomi 2023 https://t.co/atv15FGbMl'										

Gambar 3. Pengumpulan Data

Gambar 3 menunjukkan ilustrasi dataset yang akan digunakan setelah proses pengumpulan data dilakukan. Sebelum digunakan data akan diberi label terlebih dahulu.

**Pelabelan Data**

Hasil data yang telah dikumpulkan kemudian dilakukan pelabelan data yang dibagi menjadi kelas positif dan negatif. Pelabelan divalidasi secara manual dengan bantuan validator seorang lulusan sarjana bahasa yaitu Maiherawati, S.Pd.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Dateime	Tweet Id	Location	Username	Text										sentimen
2	2022-12-0	1,60E+32	Jakarta	ivooxid	b'Ketakutan Resesi, Wall Street Lanjutkan Penurunan https://t.co/HFr195qJ0U'										negatif
3	2022-12-0	1,60E+32		Ferdinand	b'Ketahanan Pangan Desa menjadi Kekuatan Indonesia menghadapi Resesi global....'										positif
4	2022-12-0	1,60E+32	Jakarta, In	ArisB5758	b'2023 semoga ekonomi Indonesia tetap tumbuh dan tahan terhadapa terpaan angin resesi dunia positif										positif
5	2022-12-0	1,60E+32		SPedalang	b'Caffe type UKM ramping dan kecil justru penyelamat ekonomi saat terjadi resesi. https://t.co/Ox1mg7w8u1'										positif
6	2022-12-0	1,60E+32		farisElsbat	b'Yang bilang Indonesia 2023 bebas resesi ga sadar kali ya Januari tanggal merahnya Hari mingg positif										positif
7	2022-12-0	1,60E+32		beritaekor	b'Indonesia Kebal Resesi, Tren IPO Propektif 2023 https://t.co/b0oUNpDaIT'										positif
8	2022-12-0	1,60E+32	My Cozy	Cantisontol	b'@CNNIndonesia tanda2 resesi ? \n\nIndonesia itu kekayaan alamnya banyak, kayak gas yang negatif										negatif
9	2022-12-0	1,60E+32		SPedalang	b'Sebentar lagi kita memasuki tahun 2023. Jangan sampai menjadi tahun kalabendu. Karena dir negatif										negatif
10	2022-12-0	1,60E+32	Medan Ko	yunanking	b'AMERIKA SERIKAT RESESI TAHUN AWAL 2023 BAHAYA UNTUK INDONESIA. JADI HARUS BAN negatif										negatif
11	2022-12-0	1,60E+32		PDI_	Perjuangan meski dunia mengalami resesi ekonomi.insya allah indonesia sudah mempi positif										positif
12	2022-12-0	1,60E+32	DKI Jakart	adibapray	b'Menurut LaNyalla, peluang tersebut juga akan mampu menjaga pertumbuhan dan stabilitas e positif										positif
13	2022-12-0	1,60E+32	DKI Jakart	adibapray	b'Namun, Ketua DPD RI, AA LaNyalla Mahmud Mattalitti, menilai ancaman resesi merupakan ta negatif										negatif
14	2022-12-0	1,60E+32	DKI Jakart	adibapray	b'Ketua DPD RI Nilai Resesi Dunia Bisa Jadi Peluang bagi Indonesia. #LaNyalla #Nyalaindonesia i positif										positif
15	2022-12-0	1,60E+32	South-sule	hariadyrig	b'Indonesia gelap 2023? Biasanya disetiap resesi akan lahir orang miskin baru dan orang kaya b positif										positif
16	2022-12-0	1,60E+32		syakilahsy	b'Bank Rakyat Indonesia (Persero) Tbk menegaskan komitmennya untk terus mendorong pertu positif										positif
17	2022-12-0	1,60E+32	DKI Jakart	derasdotic	b'Bank Indonesia (BI) memprediksi pertumbuhan ekonomi Indonesia mampu mencapai 4,7% sa positif										positif
18	2022-12-0	1,60E+32	INDONESIA,	mediatata	b'Cara Jejak Bumi Indonesia (JBI) Hadapi Ancaman Resesi Ekonomi Global 2023\n\nhttps://t.co/B negatf										negatif
19	2022-12-0	1,60E+32	Indonesia	Muammaib	b'Seorang @erickthohir aja optimis kok kalo Indonesia bisa lewati resesi di 2023 #BangkitBersa negatf										negatif
20	2022-12-0	1,60E+32	Solo. Cent	soloosodo	b'Pengamat: Indonesia Lavak Optimistis di Tengah Ancaman Resesi Ekonomi 2023 https://t.co/ negatf										negatif

Gambar 4. Pelabelan Data

Berdasarkan gambar 4 dapat dilihat hasil data yang sudah diberi label dengan dua kelas yaitu positif dan negatif. Setelah data selesai diberikan label, data akan dimasukkan ke model klasifikasi untuk diolah ke proses selanjutnya.

**3.3. Text Preprocessing**

Berikut contoh hasil dari tahapan *text preprocessing*.

**a. Cleaning**

Tahapan awal dari *text preprocessing* ini menyiapkan data untuk dianalisa dengan cara menghapus dan memodifikasi data yang tidak benar serta tidak berkaitan dengan informasi, tahapan ini bertujuan untuk meminimalisir *noise* dalam proses klasifikasi. Contoh atribut yang akan dibersihkan adalah *mention*, tanda baca, *Uniform Resource Locator (URL)*, dan *emoticon* [18].

Tabel 2. Tahapan *Cleaning*

Data Asli	Setelah <i>Cleaning</i>
b'kalau belum? coba bantu UMKM, buat UMKM. penyelamat Indonesia dari jurang resesi itu karena UMKM kita dan kita yg membelinya. inget, utamakan beli produk lokal dulu ya, pertimbangkan baik ya drpd beli impor walau murah di keranjang kuning.'	kalau belum coba bantu UMKM buat UMKM penyelamat Indonesia dari jurang resesi itu karena UMKM kita dan kita membelinya inget utamakan beli produk lokal dulu ya pertimbangkan baik drpd beli impor walau murah keranjang kuning

Proses *Cleaning* dapat dilihat pada tabel 2, atribut yang dihapus ditandai dengan karakter yang ditekankan. Berdasarkan proses pada tabel 2 dapat dilihat bahwasannya tanda baca seperti tanda tanya dan koma dihilangkan, karena tidak memiliki makna atau tidak berkaitan dengan informasi.

**b. Case Folding**

2. Diarangi mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Tahap ini mengubah seluruh kata dalam dokumen menjadi huruf kecil (*lowercase*). Proses *Case Folding* dapat dilihat pada tabel 3 yang menunjukkan kata “UMKM” diubah menjadi huruf kecil “umkm”.

**Tabel 3.** Tahapan *Case Folding*

Data Setelah <i>Cleaning</i>	Data Setelah <i>Case Folding</i>
kalau belum coba bantu <b>UMKM</b> buat <b>UMKM</b> penyelamat indonesia dari jurang resesi itu karena <b>UMKM</b> kita dan kita membelinya inget utamakan beli produk lokal dulu ya pertimbangkan baik drpd beli impor walau murah keranjang kuning	kalau belum coba bantu <b>umkm</b> buat <b>umkm</b> penyelamat indonesia dari jurang resesi itu karena <b>umkm</b> kita dan kita membelinya inget utamakan beli produk lokal dulu ya pertimbangkan baik drpd beli impor walau murah keranjang kuning

Tahapan *tokenizing* bertujuan untuk memotong *string input* atau memisahkan kalimat berdasarkan kalimat kata yang membentuk sebuah dokumen [19]. Ilustrasi dari proses *tokenizing* dapat dilihat pada tabel 4.

**Tabel 4.** Tahapan *Tokenizing*

Data Setelah <i>Case Folding</i>	Data Setelah <i>Tokenizing</i>
kalau belum coba bantu umkm buat umkm penyelamat indonesia dari jurang resesi itu karena umkm kita dan kita membelinya inget utamakan beli produk lokal dulu pertimbangkan baik drpd beli impor walau murah keranjang kuning	['kalau', 'belum', 'coba', 'bantu', 'umkm', 'buat', 'umkm', 'penyelamat', 'indonesia', 'dari', 'jurang', 'resesi', 'itu', 'karena', 'umkm', 'kita', 'dan', 'kita', 'membelinya', 'inget', 'utamakan', 'beli', 'produk', 'lokal', 'dulu', 'pertimbangkan', 'baik', 'drpd', 'beli', 'impor', 'walau', 'murah', 'keranjang', 'kuning']

Berdasarkan pada tabel 4 dapat dilihat perbedaan proses sebelum dan sesudah pemisahan kata yang terjadi pada tahapan *tokenizing*. Kumpulan kata dalam sebuah dokumen dipotong satu persatu, namun masih membentuk sebuah dokumen.

**Normalisasi**

Dalam penelitian ini, kamus normalisasi dibuat dan digunakan sebagai referensi untuk menyesuaikan kesalahan ejaan atau kata yang tidak baku kemudian diubah menjadi bentuk kata yang baku. Tabel 5 menjelaskan perubahan kata sebelum dan sesudah dilakukan proses normalisasi ditandai dengan kata yang ditebalkan.

**Tabel 5.** Tahapan Normalisasi

Data Setelah <i>Tokenizing</i>	Data Setelah Normalisasi
['kalau', 'belum', 'coba', 'bantu', 'umkm', 'buat', 'umkm', 'penyelamat', 'indonesia', 'dari', 'jurang', 'resesi', 'itu', 'karena', 'umkm', 'kita', 'dan', 'kita', 'membelinya', 'inget', 'utamakan', 'beli', 'produk', 'lokal', 'dulu', 'pertimbangkan', 'baik', 'drpd', 'beli', 'impor', 'walau', 'murah', 'keranjang', 'kuning']	['kalau', 'belum', 'coba', 'bantu', 'umkm', 'buat', 'umkm', 'penyelamat', 'indonesia', 'dari', 'jurang', 'resesi', 'itu', 'karena', 'umkm', 'kita', 'dan', 'kita', 'membelinya', 'ingat', 'utamakan', 'beli', 'produk', 'lokal', 'dulu', 'pertimbangkan', 'baik', 'daripada', 'beli', 'impor', 'walau', 'murah', 'keranjang', 'kuning']

Berdasarkan tabel 5 contoh hasil yang didapatkan adalah kata “inget” pada dataset diubah ke dalam bentuk bahasa baku yaitu “ingat”. Kata tersebut diubah dengan merujuk sesuai dengan kamus yang telah dibuat.

**Negation Handling**

Proses penanganan negasi mengidentifikasi kata-kata dengan sifat negatif, kemudian mengubah kata tersebut menjadi kata yang lainnya tetapi memiliki makna yang sama. Proses *negation handling* dapat dilihat pada tabel 6.

**Tabel 6.** Tahapan *Negation Handling*

Data Setelah Normalisasi	Data Setelah <i>Negation handling</i>
['kalau', 'belum', 'coba', 'bantu', 'umkm', 'buat', 'umkm', 'penyelamat', 'indonesia', 'dari', 'jurang', 'resesi', 'itu', 'karena', 'umkm', 'kita', 'dan', 'kita', 'membelinya', 'ingat', 'utamakan', 'beli', 'produk', 'lokal', 'dulu', 'pertimbangkan', 'baik', 'daripada', 'beli', 'impor', 'walau', 'murah', 'keranjang', 'kuning']	['kalau', 'masih', 'asing', 'bantu', 'umkm', 'buat', 'umkm', 'penyelamat', 'indonesia', 'dari', 'jurang', 'resesi', 'itu', 'karena', 'umkm', 'kita', 'dan', 'kita', 'membelinya', 'ingat', 'utamakan', 'beli', 'produk', 'lokal', 'dulu', 'pertimbangkan', 'baik', 'daripada', 'beli', 'impor', 'walau', 'murah', 'keranjang', 'kuning']

Berdasarkan hasil yang didapatkan dari proses *negation handling* pada tabel 6 kata “belum coba” diubah menjadi kata “masih asing”.

**Stopword Removal**

Proses ini [20] melakukan penghilangan kata penghubung yang dipandang tidak bermakna dalam klasifikasi sentimen dengan focus pada upaya perhitungan pada kata-kata yang jauh lebih penting atau bermakna. Tabel 7

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, pertuisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

c. Diarangkan untuk diperjualbelikan dan dipinjamkan atau disebarluaskan.

d. Tidak diperbolehkan untuk dimodifikasi, diperjualbelikan, atau dipersembahkan.

e. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

f. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

g. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

h. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

i. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

j. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

k. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

l. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

m. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

n. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

o. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

p. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

q. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

r. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

s. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

t. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

u. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

v. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

w. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

x. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

y. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

z. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

aa. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

ab. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

ac. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.

ad. Tidak diperbolehkan untuk diperjualbelikan, dipersembahkan, atau untuk propaganda.



menunjukkan proses dari *stopword removal*, tulisan yang ditinggalkan adalah kata yang dihilangkan.

Tabel 7. Tahapan *Stopword Removal*

Data Setelah <i>Negation Handling</i>	Data Setelah <i>Stopword Removal</i>
'Kalau', 'masih', 'asing', 'bantu', 'umkm', 'buat', 'umkm', 'indonesia', 'dari', 'jurang', 'resesi', 'itu', 'karena', 'membelinya', 'ingat', 'utamakan', 'beli', 'produk', 'lokal', 'dulu', 'pertimbangkan', 'baik', 'daripada', 'beli', 'murah', 'keranjang', 'kuning']	['asing', 'bantu', 'umkm', 'umkm', 'penyelamat', 'indonesia', 'jurang', 'resesi', 'umkm', 'membelinya', 'ingat', 'utamakan', 'beli', 'produk', 'lokal', 'pertimbangkan', 'beli', 'impor', 'murah', 'keranjang', 'kuning']

Hasil yang didapatkan dari contoh proses *stopword removal* pada tabel 7 adalah kata seperti “kalau” dan “masih” dihilangkan dari sebuah dokumen.

*Stemming* adalah suatu proses yang melibatkan penghapusan afiks imbuhan dari kata atau transformasi kata menjadi bentuk dasarnya [16]. Tabel 8 menunjukkan proses dari *stemming*, tulisan yang ditinggalkan adalah kata yang dihilangkan dari sebuah dokumen.

Tabel 8. Tahapan *Stemming*

Data Setelah <i>Stopword Removal</i>	Data Setelah <i>Stemming</i>
'asing', 'bantu', 'umkm', 'umkm', 'penyelamat', 'indonesia', 'jurang', 'resesi', 'umkm', 'membelinya', 'ingat', 'utamakan', 'beli', 'produk', 'lokal', 'pertimbangkan', 'beli', 'impor', 'murah', 'keranjang', 'kuning']	['asing', 'bantu', 'umkm', 'umkm', 'selamat', 'indonesia', 'jurang', 'resesi', 'umkm', 'beli', 'ingat', 'utama', 'beli', 'produk', 'lokal', 'timbang', 'beli', 'impor', 'murah', 'keranjang', 'kuning']

Berdasarkan hasil dari proses *stemming* yang dapat dilihat pada tabel 8, kata seperti “penyelamat” mengalami penghapusan afiks imbuhan sehingga menjadi bentuk dasarnya yaitu “selamat”.

3.4 Pembobotan TF-IDF

Setelah semua proses *text preprocessing* selesai dilakukan, dataset yang diolah akan memasuki proses pembobotan TF-IDF. Proses ini bertujuan untuk menghitung nilai bobot dari sebuah kata dalam sebuah dokumen. TF atau *Term Frequency* adalah frekuensi munculnya sebuah suatu *term* dalam suatu dokumen dan IDF atau *Inverse Document Frequency* adalah bobot sebuah dokumen dimana suatu *term* muncul [20]. Berikut ilustrasi proses dari pembobotan TF-IDF dapat dilihat pada gambar 5.

	000	01	07	080818	10	100	100000	11	118	12	...	yoy	yukkk	yup	zamrud	zombie	zona	zulhas	zulrifli	Kelas	
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	NEGATIF
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	NEGATIF
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	NEGATIF
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	NEGATIF
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	NEGATIF

Gambar 5. Tampilan TF-IDF pada Google Collab

Berdasarkan hasil yang ditampilkan pada gambar 5, kolom merepresentasikan sebuah kata dan baris adalah bobotnya. Gambar 5 merupakan proses TF-IDF yang telah diimplementasikan pada Google Collab.

3.5 Feature Selection (Thresholding)

Proses ini bertujuan untuk memilih fitur-fitur terbaik, output yang dihasilkan berupa informasi indeks fitur yang relevan dan akan menjadi masukan untuk pengolahan data. Pada penelitian ini percobaan akan dilakukan dengan menggunakan nilai ambang batas *threshold* sebesar 0,001. Hasil dari proses ini dapat dilihat pada gambar 6 yang sudah diimplementasikan di Google Colab.

Hasil Feature Selection		2022	2023	72	airlangga	akibat	alami	allah	amin	ancam	angka	...	tumbuh	turun	uang	umkm	untung	upaya	usaha	warga	was	Kelas
0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.0	0.000000	0.0	0.0	0.289005	0.0	0.0	NEGATIF
1	0.0	0.113314	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.0	0.000000	0.0	0.0	0.180367	0.0	0.0	NEGATIF
2	0.0	0.083745	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.0	NEGATIF
3	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.0	NEGATIF
4	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.0	NEGATIF
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
995	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.133098	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.0	POSITIF
996	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.260163	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.0	POSITIF
997	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.0	0.232331	0.0	0.0	0.265195	0.0	0.0	POSITIF
998	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.0	POSITIF
999	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.0	POSITIF

Seleksi Fitur: 186 fitur  
 Fitur: ['2022', '2023', '72', 'airlangga', 'akibat', 'alami', 'allah', 'amin', 'ancam', 'angka', 'antisipasi', 'atas', 'badai', 'baik', 'banding', 'bangsa', 'bangun',

Gambar 6. Tampilan Feature Selection pada Google Colab



Berdasarkan hasil yang dapat dilihat pada gambar 6, kata yang ditampilkan adalah kata yang memenuhi ambang batas yang telah ditentukan. Contohnya kata “ancam”, “antisipasi”, dan “bangsa”.

**3.3 Klasifikasi K-Nearest Neighbor**

Perhitungan klasifikasi menggunakan K-NN dilakukan dengan menentukan nilai K, pada penelitian ini nilai K yang digunakan yaitu K = 3,5,7,9,11,13, dan 15. Sistem bekerja efektif jika hasil klasifikasi data sebagai positif atau negatif dapat mengidentifikasi dan mengategorikan data yang diuji masuk ke dalam kelas yang sama dengan kelas pada data

Hasil Klasifikasi	2022	2023	72	airlangga	akibat	alami	allah	amin	ancam	angka	...	turun	uang	umkm	untung	upaya	usaha	warga	was	Kelas Target	Hasil Klasifikasi	
0	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	POSITIF	POSITIF
1	0.0	0.273194	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	POSITIF	POSITIF
2	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	POSITIF	POSITIF
3	0.0	0.000000	0.0	0.000000	0.000000	0.180773	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	NEGATIF	NEGATIF
4	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	NEGATIF	NEGATIF
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
95	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	POSITIF	POSITIF
96	0.0	0.000000	0.0	0.271076	0.000000	0.000000	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	POSITIF	POSITIF
97	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.197669	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	POSITIF	POSITIF
98	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	NEGATIF	NEGATIF
99	0.0	0.000000	0.0	0.000000	0.218612	0.000000	0.0	0.0	0.134963	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	NEGATIF	NEGATIF

Gambar 7. Tampilan Klasifikasi K-NN pada Google Colab

Hasil pada proses ini dapat dilihat pada gambar 7 yang menunjukkan model dapat bekerja dengan baik karena klasifikasi masuk ke dalam kelas targetnya. Kelas target merupakan dataset yang sudah diberi label secara manual dan hasil klasifikasi adalah hasil berupa label dari model yang dibuat ketika proses klasifikasi telah selesai dilakukan.

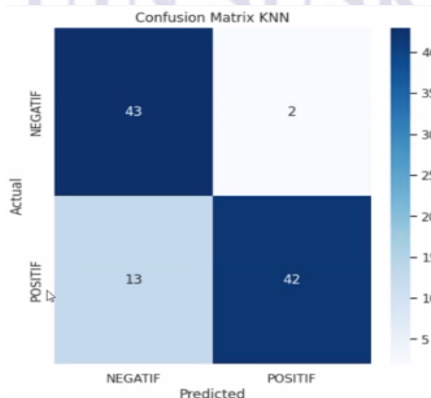
**Pengujian**

Setelah proses pembobotan kata, seleksi fitur, dan klasifikasi dengan metode telah selesai dilakukan, maka tahapan selanjutnya adalah tahapan pengujian untuk mencari dan mengevaluasi nilai *accuracy*, *precision*, *recall*, dan *f1 score*. Pada penelitian ini pengujian dilakukan dengan menggunakan *confusion matrix*, adapun nilai K = 3,5,7,9,11,13,15 sebagai jumlah tetangga terdekat kemudian untuk perbandingan yang digunakan yaitu 60:40, 70:30, 80:20, dan 90:10. Pengujian yang telah dilakukan dapat dilihat pada tabel 9.

Tabel 9. Pengujian Confusion Matrix

Nilai K	Accuracy			
	60:40	70:30	80:20	90:10
3	74.5	78.3	79.0	85.0
5	68.5	71.0	75.5	80.0
7	68.0	70.0	72.5	77.0
9	65.0	68.7	71.0	77.0
11	61.8	64.3	67.5	75.0
13	58.8	62.0	65.0	73.0
15	55.2	59.7	64.5	68.0

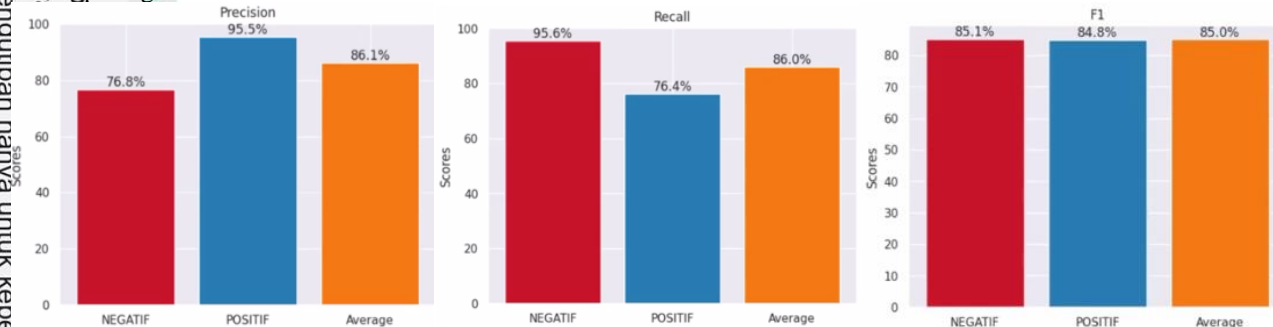
Dari pengujian yang sudah dilakukan, hasil akurasi tertinggi terletak diangka 85,0% dengan rasio 90:10 data latih dan uji pada nilai K=3. Hasil *confusion matrix* dengan akurasi tertinggi dapat dilihat pada gambar 8.



Gambar 8. Tampilan Hasil Confusion Matrix



Gambar diatas menjelaskan hasil dari pengujian model klasifikasi yang telah dilakukan, angka yang terdapat pada gambar menunjukkan seberapa baik model untuk menentukan kelas aktual masuk ke dalam kelas prediksi pada perbandingan data 90:10 dengan nilai K = 3. Kemudian hasil evaluasi lainnya untuk *precision*, *recall*, dan *f1 score* dapat dilihat pada gambar 9.



Gambar 9. Tampilan Hasil *Precision*, *Recall*, dan *F1 Score*

Berdasarkan gambar 9 dapat dilihat bahwa terdapat nilai *precision*, *recall* dan *f1-1 score*, hasil ini didapatkan dari perhitungan pengujian *confusion matrix*. Adapun hasil perhitungan yang didapatkan adalah nilai *accuracy* dengan *macro average* 86,1%, *recall* 86,0% dan *f1 score* 85,0%. Hasil ini menunjukkan bahwa model klasifikasi dapat bekerja dengan

#### 4. KESIMPULAN

Dasarakan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa metode yang digunakan terbilang efektif untuk melakukan klasifikasi sentimen terkait ancaman resesi ekonomi tahun 2023. Hasil pengujian dengan variasi perbandingan data dan nilai K yang berbeda menunjukkan bahwa kinerja terbaik ditemukan pada pengujian dengan nilai K=3 serta perbandingan 90:10 data *training* dan *testing* yang memiliki akurasi sebesar 85,0%. Hasil tersebut juga diikuti dengan nilai *precision* sebesar 86,1%, *recall* sebesar 86,0%, dan *f1 score* sebesar 85,0%. Pentingnya mencatat bahwa pengujian dengan jumlah data uji yang lebih sedikit cenderung menghasilkan akurasi yang lebih tinggi, namun penambahan data uji dapat mengakibatkan penurunan akurasi. Oleh karena itu, penelitian selanjutnya dapat dikembangkan dengan menggunakan dataset yang lebih besar dan menguji metode *feature weighting* dan *feature selection* alternatif.

#### REFERENCES

- S. Blandina, A. Noor Fitriani, dan W. Septiyani, "Strategi Menghindarkan Indonesia dari Ancaman Resesi Ekonomi di Masa Pandemi," *Efektor*, vol. 7, no. 2, hlm. 181–190, Des 2020, doi: 10.29407/e.v7i2.15043.
- S. Itmam, "Resesi Ekonomi dan Implikasinya dari Perspektif Hukum Bisnis," *Equilibrium: Jurnal Ekonomi Syariah*, vol. 9, no. 1, hlm. 1–9, Jun 2021, doi: 10.21043/equilibrium.v9i1.10012.
- A. Adhar, "Penataan Ancaman Ekonomi Sebagai Bagian Dari Keadaan Bahaya Di Indonesia," *Dialogia Iuridica: Jurnal Hukum Bisnis dan Investasi*, vol. 11, hlm. 33, Nov 2020.
- Badan Pusat Statistik, "Berita Resmi Statistik," 2022, Diakses: 11 Desember 2022. [Daring]. Tersedia pada: [https://www.bps.go.id/website/materi\\_ind/materiBrsInd-20220805094500](https://www.bps.go.id/website/materi_ind/materiBrsInd-20220805094500)
- A. Sandra, A. Syach, V. Aulya, dan D. Kustiawati, "Studi Literatur : Mempersiapkan Investasi untuk Hadapi Isu Resesi Ekonomi di Indonesia," *Jurnal Pendidikan dan Konseling*, vol. 4, no. 6, hlm. 1–9, Des 2022, doi: <https://doi.org/10.31004/jpdk.v4i6.9609>.
- Suaidah dan Marliyah, "Upaya Keuangan Syariah Terhadap Ancaman Resesi Global," *Edunomika*, vol. 07, no. 01, hlm. 1–9, Nov 2022.
- A. P. Rodrigues dan N. N. Chiplunkar, "A New Big Data Approach For Topic Classification And Sentiment Analysis Of Twitter Data," *Evol Intell*, vol. 15, no. 2, hlm. 877–887, Jun 2022, doi: 10.1007/s12065-019-00236-3.
- M. Cinda dan D. P. Rini, "Metode Klasifikasi Pada Sentimen Analisis," *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, hlm. 66–70, 2019.
- I. A. Angreni, S. A. Adisasmita, M. I. Ramli, dan S. Hamid, "Pengaruh Nilai K Pada Metode K-Nearest Neighbor (KNN) Terhadap Tingkat Akurasi Identifikasi Kerusakan Jalan," *Rekayasa Sipil*, vol. 7, no. 2, hlm. 63, Jan 2019, doi: 10.22441/rjs.2018.v07.i2.01.
- L. M. Sinaga, Sawaluddin, dan S. Suwilo, "Analysis of classification and Naïve Bayes algorithm k-nearest neighbor in data mining," *IOP Conf Ser Mater Sci Eng*, vol. 725, no. 1, Jan 2020, doi: 10.1088/1757-899X/725/1/012106.
- R. Gunawan, R. Septiadi, F. Apri Wenando, H. Mukhtar, dan Syahril, "K-Nearest Neighbor (KNN) untuk Menganalisis Sentimen terhadap Kebijakan Merdeka Belajar Kampus Merdeka pada Komentar Twitter," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 3, no. 2, hlm. 152–158, Agu 2022, doi: 10.37859/coscitech.v3i2.3841.
- D. Muhiidin dan A. Wibowo, "Perbandingan Kinerja Algoritma Support Vector Machine Dan K-Nearest Neighbor Terhadap Analisis Sentimen Kebijakan New Normal," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 5, hlm. 1–7, Des 2020.
- M. R. Irfan, M. A. Fauzi, T. Tibyani, dan N. D. Mentari, "Twitter Sentiment Analysis on 2013 Curriculum Using Ensemble Features and K-Nearest Neighbor," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 6, hlm. 5409, Des 2018, doi: 10.11591/ijece.v8i6.pp5409-5414.



- [14] E. Indrayuni, A. Nurhadi, dan D. A. Kristiyanti, "Implementasi Algoritma Naive Bayes, Support Vector Machine, dan K-Nearest Neighbor untuk Analisa Sentimen Aplikasi Halodoc," *Faktor Exacta*, vol. 14, no. 2, hlm. 64, Agu 2021, doi: 10.30998/faktorexacta.v14i2.9697.
- [15] Romli, S. Prameswari R, dan A. Z. Kamalia, "Sentiment Analysis about Large-Scale Social Restrictions in Social Media Twitter Using Algoritma K-Nearest Neighbor," *Jurnal Online Informatika*, vol. 6, no. 1, hlm. 96, Jun 2021, doi: 10.15575/join.v6i1.670.
- [16] Kosasih dan A. Alberto, "Analisis Sentimen Produk Permainan Menggunakan Metode TF-IDF Dan Algoritma K-Nearest Neighbor," *Jurnal Nasional Informatika dan Teknologi Jaringan*, vol. 6, no. 1, 2021, doi: 10.30743/infotekjar.v6i1.3893.
- [17] M. Pravina, K. N. Sani, H. D. Harfianto, T. A. Pratama, A. Fahrina, dan Y. Ruldeviyani, "Sentiment Analysis Of Delivery Service Opinions On Twitter Documents Using K-Nearest Neighbor," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 2, hlm. 996–1012, 2022, [Daring]. Tersedia pada: <http://jurnal.mdp.ac.id>
- [18] Yoga Pratama, Y. Umaidah, dan A. Voutama, "Analisis Sentimen Media Sosial Twitter Dengan Algoritma K-Nearest Neighbor Dan Seleksi Fitur Chi-Square (Kasus Omnibus Law Cipta Kerja)," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 5, no. 2, hlm. 897–910, 2021.
- [19] K. Pali dan E. Umar, "Analisis Sentimen Komentar Pengunjung Terhadap Tempat Wisata Danau Weekuri Menggunakan Metode Naive Bayes Classifier Dan K-Nearest Neighbor," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 4, hlm. 2309, Okt 2022, doi: 10.30865/mib.v6i4.4635.
- [20] A. Deviyanto dan M. R. Didik Wahyudi, "Penerapan Analisis Sentimen Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor," *Jurnal Informatika Sunan Kalijaga*, vol. 3, no. 1, hlm. 1–13, 2018.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

