

Uji Akurasi Penggunaan Metode KNN dalam Analisis Sentimen Kenaikan Harga BBM pada Media *Twitter*

Andrean Jonathan Arifin^{1*}, Adi Nugroho²

Teknik Informatika, Universitas Kristen Satya Wacana, Salatiga, Indonesia

*e-mail *Corresponding Author*: spensabara06@gmail.com

Abstract

The government in an effort to run the government is inseparable from the policies that must be made and taken. One policy alone can attract diverse sentiments from society. Based on this, this study was made to analyze public sentiment towards government policies, especially government policies regarding fuel price hikes. In this study, the analysis process using the K-Nearest Neighbor algorithm classifies tweets from Twitter into two categories, namely positive and negative. The research stages started from crawling data, data preprocessing, labeling, classification using the KNN algorithm, and evaluation. With an accuracy of 94.33% in classifying data. With the results of this research, it is hoped that it will make it easier for the government to see people's responses and sentiments towards the fuel increase policy so that the government can produce better policies by incorporating what the people have given.

Keywords: *K-Nearest Neighbor Algorithm; Rise in fuel prices; Sentiment Analysis; Twitter.*

Abstrak

Pemerintah dalam upaya menjalankan pemerintahan tidak terlepas dengan kebijakan-kebijakan yang harus dibuat dan diambil. Dari satu kebijakan saja sudah bisa menarik sentimen yang beragam dari masyarakat. Dengan berdasarkan hal itu penelitian ini dibuat untuk menganalisa sentimen masyarakat terhadap kebijakan pemerintah khususnya pada kebijakan pemerintah mengenai Kenaikan BBM. Pada penelitian ini proses analisa menggunakan algoritme *K-Nearest Neighbor* mengklasifikasikan tweets dari twitter menjadi dua kategori yaitu positif dan negatif. Tahapan penelitian dimulai dari *crawling* data, data *preprocessing*, pelabelan, klasifikasi menggunakan algoritme KNN, dan Evaluasi. Dengan hasil akurasi 94,33% dalam mengklasifikasikan data. Dengan hasil penelitian ini diharapkan dapat mempermudah pemerintah dalam melihat tanggapan dan sentimen masyarakat terhadap kebijakan kenaikan BBM sehingga pemerintah dapat menghasilkan kebijakan yang lebih baik bersama dengan masukkan-masukkan yang diberikan rakyat.

Kata kunci: *Algoritme K-Nearest Neighbor; Kenaikan Harga BBM; Analisa Sentimen; Twitter.*

1. Pendahuluan

Indonesia merupakan peringkat keempat di dunia dalam jumlah penduduk tertinggi dengan jumlah penduduk yang mencapai lebih dari 200 juta jiwa. Dengan banyaknya jumlah penduduk di Indonesia hal itu berbanding lurus dengan tingkat mobilitas di Indonesia yang berakibat tingginya angka penggunaan kendaraan bermotor. Sebagian besar kendaraan di Indonesia masih menggunakan bahan bakar minyak sebagai bahan bakar supaya kendaraan tersebut dapat bergerak, alhasil kebutuhan bahan bakar minyak Indonesia sangatlah besar [1][2].

Bahan bakar minyak dapat mempengaruhi banyak sektor maka dari itu pemerintah Indonesia memiliki hak dalam mengatur harga bbm. Pemerintah memberikan subsidi yang cukup besar ke dalam bahan bakar minyak supaya harga bbm tidak terlalu tinggi dan mobilitas di Indonesia dapat berjalan lancar.

Pada September 2022 harga bahan bakar minyak terus meningkat hal itu disebabkan adanya perang antara Russia dan Ukraina. Russia merupakan importir minyak yang cukup besar, Russia menghentikan kegiatan ekspor minyak yang mengakibatkan ketersediaan minyak di pasar dunia menjadi berkurang. Menurut *Cnbcindonesia.com* pada 1 Juni 2022 "Invasi Rusia ke Ukraina pada akhir Februari telah membuat pasar energi terguncang" [3]. Dengan naiknya harga minyak

mentah maka produk olahan dari minyak bumi dipastikan akan meningkat salah satunya adalah bahan bakar minyak.

Dalam menyikapi terus meningkatnya harga BBM pemerintah Indonesia membuat sebuah kebijakan. Tercatat mulai Sabtu (3/9/2022) pukul 14.30. Presiden Jokowi dalam jumpa pers di Istana Merdeka menjelaskan bahwa pemerintah saat ini harus mengambil keputusan yang sulit untuk menyaingi tingginya harga. Pemerintah mengambil keputusan melakukan penyesuaian harga BBM dan mengalihkan subsidi BBM kedalam bentuk lain [4].

Akibat dari kebijakan yang dimana pemerintah Indonesia akan mengurangi subsidi BBM, hal itu akan mempengaruhi mulai dari ekonomi, bahan pokok dan banyak sektor lainnya yang akan terkena imbasnya, tentu saja permasalahan ini akan menjadi polemik yang cukup besar [5] [6]. Selain banyak kerusuhan kerusuhan dan demonstrasi yang terjadi dampak dari polemik tersebut dapat dilihat di sosial media, banyak orang membicarakan dan memperdebatkan isu kenaikan BBM khususnya pada aplikasi *Twitter*.

Penelitian ini melakukan klasifikasi terhadap sentimen masyarakat menggunakan metode *K-Nearest Neighbor* dengan judul "Uji Akurasi Penggunaan Metode KNN dalam Analisis Sentimen Kenaikan Harga BBM pada *Twitter*" yang dimana dengan metode KNN kita menentukan titik pusat dan mengelompokkan data yang ada di sekitar titik pusat menjadi satu kelompok. Klasifikasi dilakukan dengan membagi data kedalam dua kelompok berdasarkan sentimen yaitu sentimen positif dan negatif. Selanjutnya dilakukan analisa yang membuat kita dapat melihat bagaimana reaksi masyarakat mengenai isu tersebut dan tujuan dari penelitian ini dapat menghasilkan klasifikasi sebagai bahan pertimbangan lanjutan dalam menentukan kebijakan lanjutan atau sebagai pertimbangan lain untuk kebijakan yang ada.

2. Tinjauan Pustaka

Penelitian yang dilakukan Ulfa Kurniasih dan Akrim Teguh Suseno dalam penelitian mereka dengan judul "Analisis Sentimen Terhadap Bantuan Subsidi Upah (BSU) pada Kenaikan Harga Bahan Bakar Minyak (BBM)". Mereka melakukan penelitian terhadap sentimen masyarakat mengenai tindakan lanjutan dari kebijakan pemerintah pemberian subsidi bantuan pada kenaikan BBM. Metode yang digunakan dalam penelitian yang dilakukan adalah *naïve bayes*. Data yang digunakan adalah 795 tweet pada masing-masing kata kunci BBM dan kata kunci BSU. Data tersebut dibagi menjadi 2 yaitu data latih sebesar 263 dan 532 untuk data uji. Metode yang digunakan adalah klasifikasi dengan algoritme *Naïve Bayes*. Hasil analisis menunjukkan bahwa pada kata kunci BBM sentimen positif yang didapatkan yaitu 28,2% dan pada sentimen negatif yaitu 71,8%. Pada kata kunci BSU, sentimen positif sebesar 65,2% dan sentimen negatif 34,8%. Pada tingkat akurasi dengan metode ini mendapatkan hasil 82,64% dan presisi sebesar 92,89%. Oleh sebab itu dapat disimpulkan hasil sentimen masyarakat terhadap Bantuan Subsidi Upah (BSU) mendapatkan respon positif sedangkan pada kenaikan harga Bahan Bakar Minyak (BBM) mendapatkan respon negatif [7].

Penelitian yang dilakukan Muhammad Ghifari, Reva Ragam Santika, Sejati Waluyo dalam penelitian yang berjudul "Analisis Sentimen Twitter Terhadap Kenaikan Bahan Bakar Minyak Menggunakan Algoritme *Naïve Bayes*". Penelitian ini melakukan analisa menggunakan *naïve bayes* terhadap opini masyarakat dari aplikasi twitter yang mengklasifikasikan opini menjadi dua yaitu positif dan negatif. Klasifikasi dilakukan dengan menggunakan metode algoritme *naïve bayes* sebagai pengujian. Hasil dari penelitian ini adalah dari perbandingan data latih dan uji yaitu 90:10 dan mendapatkan nilai akurasi tertinggi 81,818% dibandingkan dengan rasio lainnya. Perkiraan yang diperoleh dari pengelompokan lebih besar untuk label positif. Selain melakukan klasifikasi untuk menggambarkan tampilan, penelitian ini juga mengevaluasi kinerja model yang dihasilkan [8].

Penelitian yang dilakukan oleh Arief Asro'i dan Hery Februriyanti dengan judul "Analisis Sentimen Pengguna *Twitter* terhadap Perpanjangan PPKM Menggunakan Metode *K-Nearest Neighbor*". Mereka melakukan analisa menggunakan algoritme *K-Nearest Neighbor* untuk menentukan sentimen pengguna twitter terhadap *tweet* mengenai kebijakan perpanjangan PPKM, dengan bantuan library *Scikit-learn* yang populer di kalangan Data Scientist. Algoritme tersebut diterapkan ke 6408 data *tweet* dengan kata kunci "PPKM" yang dikumpulkan pada 1 Juli 2021 - 31 Desember 2021. Hasil training model membuktikan bahwa skor akurasi 69,5% dalam mengklasifikasikan *tweet* dengan benar [9].

Penelitian yang dilakukan oleh Rifna Savira, Achmad Solichin, Imelda, Mohammad Syafrullah dengan judul "Analisis Sentimen Pada Twitter Terhadap Kenaikan BBM 2022 Dengan

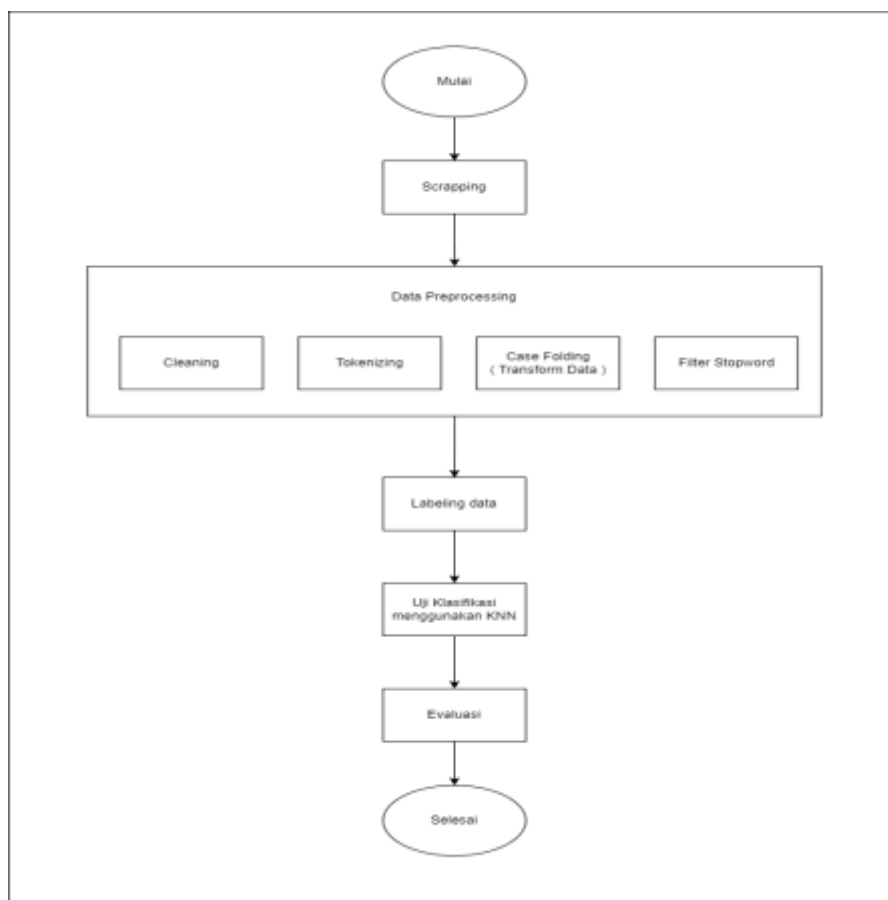
Lexicon dan *Support Vector Machine*". Mereka melakukan analisis mengenai kebijakan pemerintah dalam menghadapi kenaikan BBM pada tahun 2022. Analisa dilakukan dengan menggunakan Lexicon dan *Support Vector Machine*. Data yang digunakan diambil dari aplikasi twitter dengan jumlah 400 data tweet yang dibagi menjadi 234 sentimen negatif dan 162 sentimen positif. Hasil akhir dari penelitian ini adalah skor akurasi 78% dalam mengklasifikasikan tweet dengan benar. [10]

Perbedaan penelitian ini dengan penelitian [5] [6] [7] adalah penelitian ini mengambil kasus dari polemik mengenai kenaikan BBM di Indonesia, polemik ini terjadi pada tahun 2022 sehingga dapat dikatakan penelitian dengan kasus yang sedang hangat diperbincangkan dan populer pada tahun 2022. Selain itu, kenaikan BBM dapat dikatakan polemik yang cukup besar dan menarik perhatian banyak orang karena dengan naiknya harga BBM maka banyak sektor yang akan ikut terkena imbasnya. Perbedaan penelitian ini dengan penelitian [8] yang sama-sama menggunakan polemik yang sama adalah pada penelitian ini data yang digunakan jauh lebih banyak dengan tingkat akurasi yang lebih tinggi dan menggunakan metode yang berbeda sehingga dapat menjadi bahan pertimbangan. Penelitian ini dilakukan menggunakan tools rapidminer dengan tingkat akurasi sistem dalam mengklasifikasikan data yang lebih tinggi dibandingkan penelitian-penelitian sebelumnya yang mencapai 94%.

3. Metodologi

3.1 Prosedur Penelitian

Permasalahan yang akan dianalisis dan diteliti dalam penelitian ini adalah mengenai kebijakan kenaikan harga BBM di Indonesia dalam upaya menyikapi kenaikan harga minyak dunia. Dengan menerapkan metode klasifikasi KNN yang membedakan sentimen masyarakat menjadi dua kategori yaitu positif dan negatif. Pengolahan data dan proses analisis dilakukan menggunakan tools rapidminer. Data yang digunakan diambil dari twitter dengan menggunakan kata kunci Kenaikan BBM data yang diambil berjumlah 20.000 data dan data sampel yang berjumlah 100 data.



Gambar 1. Workflow Penelitian

Gambar 1 merupakan *workflow* dari sistem klasifikasi menggunakan KNN pada penelitian ini. Dimana data bersumber dari twitter dengan keyword "kenaikan BBM". Lalu data melalui tahapan preprocessing data. Pada bagian ini data diolah membuang data rusak dan tidak sesuai, memecah kalimat menjadi kata per kata, menyamaratakan *uppercase* dan *lowercase*, dan lain sebagainya sehingga data dapat dieksekusi oleh sistem. Langkah selanjutnya adalah data diberikan label dan data yang sudah diberikan dijadikan data training. Selanjutnya adalah uji klasifikasi KNN data training diujikan dalam sistem dan menghasilkan model tertentu, dari model tersebut yang nantinya akan menjadi acuan atau model klasifikasi untuk data yang lainnya.

3.2 Algoritme KNN

Metode yang digunakan pada penelitian ini adalah KNN karena sesuai dengan kebutuhan data yang memiliki dimensi sederhana dan hanya menggunakan 2 parameter yaitu user dan tweets, selain itu KNN mudah diterapkan dan mudah beradaptasi dengan kondisi data.

Cara kerja KNN mencari kelompok k objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing [13]. KNN mengklasifikasikan objek berdasarkan data latih yang jaraknya paling dekat dari objek tersebut. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan jarak euclidean dengan persamaan umum [14] Pada penelitian ini KNN akan mengklasifikasikan data kedalam dua kategori yaitu sentimen positif dan negatif. Perhitungan rumus algoritme KNN dalam mengklasifikasikan data sebagai berikut:

$$d = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \dots\dots\dots (1)$$

Keterangan:

d = jarak

a = data uji/testing

b = sampel data

i = variable data

n = Dimensi data

3.3 Confusion Matrix

Setelah melalui tahapan uji klasifikasi hasil tersebut akan dievaluasi menggunakan confusion matrix. Confusion matrix merupakan sebuah model evaluasi klasifikasi berdasarkan data uji dan seluruh data yang diprediksi dengan proporsi yang tepat [16]. Adapun perhitungan tingkat akurasi pada Confusion Matriks 2 kelas adalah:

Tabel 1 Tabel *Confusion Matrix*

Kelas	Terklasifikasi positif	Terklasifikasi negatif
Positif	TP (True Positive)	FN (False Negative)
Negatif	FP (False Positive)	TN (TrueNegative)

Pada penelitian ini dilakukan beberapa simulasi nilai k pada metode KNN dan melakukan metode KNN untuk mendapatkan metode dengan akurasi terbaik di penelitian ini. Akurasi dihitung dengan *Confusion Matrix*.

Keseluruhan proses akan dilakukan menggunakan Rapidminer. Rapidminer merupakan sebuah tools (aplikasi) yang dapat digunakan untuk melakukan data mining. RapidMiner menawarkan antarmuka drag-and drop yang memungkinkan pengguna untuk melakukan pengambilan data, data *preprocessing*, hingga implementasi algoritme hanya dengan menggunakan operator yang sudah ada [17].

4. Hasil dan Pembahasan

4.1. Scraping Data

Scraping data merupakan proses pengambilan data dari twitter, data yang diambil berupa tweet dan retweet dengan menggunakan keyword tertentu. Pada penelitian ini peneliti

menggunakan bahasa pemrograman python untuk mengambil data. Keyword yang digunakan dalam penelitian ini adalah “Kenaikan BBM”. Data yang diambil meliputi tweet, dan retweet.

4.2. Data Preprocessing

Data preprocessing merupakan tahapan dimana data di proses. Disini peneliti menggunakan tools Rapidminer peneliti memilih dan membuang data rusak dan data tidak sesuai dengan menggunakan operators replace dan filter pada rapidminer. Hal ini bertujuan dengan membuang data yang rusak dan tidak sesuai akan meminimalisir adanya ketidakkonsistenan data dan menghilangkan elemen – elemen yang dapat mengganggu ketika data di proses. Tahapan ini terdiri dari:

1) Cleaning Data

Row No.	Tweets	User	#
77	yg baru dikit	Demata_091	6
78	@JACKOBROTT AJG #PKH2019 #BBM naik? #KRI	hu3H9u0R	11
79	Panas LKPU DPRD Kabupaten Sibar: menerima laporan dari warga terkait adanya potensi penyalah DLT komposisi kenaikan harga BBM	hampas	7
80	@sajenas Demu kenaikan BBM?	chujis	11
81	yg kenapa kok naik pejabat pemerintah? yg nanggung abah abah naik? BBM naik harga2 naik? yg papa? Hah! DPR mada kenaikan gas? dn yg kenapa? rata2 kelakuan oknum bermata sipit? yg dn?	Abah21240	11
82	@okni @KementerianBBM wazpadai kenaikan harga minyak dalam jangka pendek pasihan jaminan pasihan bsm amp antisipasi penyelesaian harga bsm utk menjaga subsidi bsm terkendali https://co22g	ISIR	11
83	Poksek Poltek Lima pulah melaksanakan patroli mobil pada stasi SPBU Pasca Kenaikan Harga BBM di wilayah Halim Poltek Lima pulah https://co22g2N0j	PoltekLima	11
84	!	!	0
85	!	!	0
86	Kenaikan minyak mentah dunia > nilai impor melonjak > APBN terdapat > Pemerintah menuntun subsidi > Harga BBM naik kembali	asasa_endo	11
87	Jadi sebenarnya harga BBM ini mengalami kenaikan dipengaruhi oleh banyak faktor kebutuhan akan tetapi semakin sedikitnya minyak yg tersedia. Sehingga menyebabkan harga minyak mentah di dunia naik. Ditaman	arata_pu8PS	11
88	Pertamina sendiri adalah perusahaan swasta yang menyuplai minyak dan gas ketubana BBM di negara kita. Sebenarnya harga BBM itu naik karena adanya subsidi dari Pemerintah. Tapi kenapa sekarang	arata_pu8PS	11
89	Mungkin masih ada beberapa orang yang belum tau beberapa penyebab kenaikan BBM yang dilakukan oleh salah satu BUMN di Indonesia yaitu Pertamina. Yuk Kita Bahas!	arata_pu8PS	8
90	book2 pakas yg sula manaj + nggunang sula bahas random 8Y-8Y- bsa2 bahas kenaikan bsm	bertubi	11

Gambar 2 Sebelum Cleaning Data

Row No.	Tweets	User	#
39	yg baru dikit	Demata_091	6
40	AJG PKH2019 #BBM naik? #KRI	hu3H9u0R	11
41	Panas LKPU DPRD Kabupaten Sibar: menerima laporan dari warga terkait adanya potensi penyalah DLT komposisi kenaikan harga BBM	hampas	7
42	Demu kenaikan bsm?	chujis	11
43	yg kenapa kok naik pejabat pemerintah? yg nanggung abah abah naik? BBM naik harga2 naik? yg papa? Hah! DPR mada kenaikan gas? dn yg kenapa? rata2 kelakuan oknum bermata sipit? yg dn?	Abah21240	11
44	wazpadai kenaikan harga minyak dalam jangka pendek pasihan jaminan pasihan bsm amp antisipasi penyelesaian harga bsm utk menjaga subsidi bsm terkendali	ISIR	11
45	Kenaikan minyak mentah dunia > nilai impor melonjak > APBN terdapat > Pemerintah menuntun subsidi > Harga BBM naik kembali	asasa_endo	11
46	Jadi sebenarnya harga BBM ini mengalami kenaikan dipengaruhi oleh banyak faktor kebutuhan akan tetapi semakin sedikitnya minyak yg tersedia sehingga menyebabkan harga minyak mentah di dunia naik. Ditaman	arata_pu8PS	11
47	book2 pakas yg sula manaj + nggunang sula bahas random 8Y-8Y- bsa2 bahas kenaikan bsm	bertubi	11

Gambar 3 Setelah Cleaning Data

Pada proses ini dilakukan pengecekan dan pembersihan data dari data yang corrupt, missing value, dan lain sebagainya. Pada proses pengambilan data twitter tidak sedikit data yang rusak atau terdapat value yang kosong dan komponen yang tidak diperlukan maka dari itu diperlukan proses cleaning data untuk menghilangkan elemen – elemen pengganggu yang dapat membuat ketidakkonsistenan data dan mempengaruhi hasil proses. Contoh dari data yang dibuang adalah membuang kata – kata yang tidak diperlukan seperti retweet (RT), mention (@), hashtag (#), url (https), data hilang, dan spasi (trim) yang tidak diperlukan dengan menggunakan replace dan filtering pada rapidminer.

2) Tokenizing

Row No.	Tweets
1	Ratat bawah udah babak baru awal kenaikan BBM
2	Tang dua masa-ha per mesin sapal banhan foto sudah bisa mesin bag? Tapi yang dua bisa mesin bag juga bisa kaya dan hampir mesin dia bisa dihal banhan sehalang dia yang jadi mesin ke
3	Sebagai Kenaikan Kenaikan/BBM di Jalang dan DIT? In Prebidai Pertamina
4	13 tahun Tiana Jakarta tant that jauh awal Rp 1500 ahanya akan naik juga 13 tahun itu sudah mengalami beberapa kali pergantian presiden gubernur DKI dan sudah beberapa kali mengalami kenaikan BBM
5	Seus ketikny dari bekal baka fosilama ojaamungin di dapat proyek besar mode kasa 2A_EuM2R4_Eu8E2orp taun ke opan y ata lg kenaikan harga bsmadahnya kenakan listrik

Gambar 4 Sebelum Proses Tokenizing

Row No.	Teks	BBM	BBM	BBM	BBM	BBM	BBM	BBM	BBM	BBM	BBM	BBM
1	Rakyat banyak udah bener bener abis deh kenaikan BBM	1	1	0	1	1	0	0	1	1	0	0
2	Sangat malu malu gara gara mesin dapat bantuan free sudah tidak mesin lagi tapi yang tidak mesin lagi juga tidak bisa dan hampir mesin dia tidak dapat bantuan sekarang dia yang jadi mesin lagi	1	1	0	1	1	0	0	1	1	0	0
3	Peraga kenaikan konsumsi BBM di jalang dan DIT itu Presiden Perdana	1	1	0	1	1	0	0	1	1	0	0
4	19 tahun Transkribasi ke PDF jadi udah Rp 2000 awalnya jadi jadi Rp 10000 itu sudah mengalami beberapa kali pengisian presiden pemerintah di dan sudah beberapa kali mengalami kenaikan BBM	1	1	0	1	1	0	0	1	1	0	0
5	terus listriknya dari bahan bakar fosilnya aja samain aja dia dapat provikan besar mobil listrik 3'4, 6'6, 8'8, 10'10, 12'12, 14'14, 16'16, 18'18, 20'20, 22'22, 24'24, 26'26, 28'28, 30'30, 32'32, 34'34, 36'36, 38'38, 40'40, 42'42, 44'44, 46'46, 48'48, 50'50, 52'52, 54'54, 56'56, 58'58, 60'60, 62'62, 64'64, 66'66, 68'68, 70'70, 72'72, 74'74, 76'76, 78'78, 80'80, 82'82, 84'84, 86'86, 88'88, 90'90, 92'92, 94'94, 96'96, 98'98, 100'100	1	1	0	1	1	0	0	1	1	0	0

Gambar 5 Sesudah Proses Tokenizing

Pada proses ini penulis memecah kalimat menjadi kata per kata. Hal ini bertujuan supaya mempermudah sistem dalam menganalisa berdasarkan kata sehingga keakuratan sistem lebih baik.

3) Case Folding (Transform Data)

Row No.	Teks
1	Rakyat banyak udah bener bener abis deh kenaikan BBM
2	Sangat malu malu gara gara mesin dapat bantuan free sudah tidak mesin lagi tapi yang tidak mesin lagi juga tidak bisa dan hampir mesin dia tidak dapat bantuan sekarang dia yang jadi mesin lagi
3	Peraga kenaikan konsumsi BBM di jalang dan DIT itu Presiden Perdana
4	19 tahun Transkribasi ke PDF jadi udah Rp 2000 awalnya jadi jadi Rp 10000 itu sudah mengalami beberapa kali pengisian presiden pemerintah di dan sudah beberapa kali mengalami kenaikan BBM
5	terus listriknya dari bahan bakar fosilnya aja samain aja dia dapat provikan besar mobil listrik 3'4, 6'6, 8'8, 10'10, 12'12, 14'14, 16'16, 18'18, 20'20, 22'22, 24'24, 26'26, 28'28, 30'30, 32'32, 34'34, 36'36, 38'38, 40'40, 42'42, 44'44, 46'46, 48'48, 50'50, 52'52, 54'54, 56'56, 58'58, 60'60, 62'62, 64'64, 66'66, 68'68, 70'70, 72'72, 74'74, 76'76, 78'78, 80'80, 82'82, 84'84, 86'86, 88'88, 90'90, 92'92, 94'94, 96'96, 98'98, 100'100

Gambar 10 Sebelum Proses Case Folding

Row No.	Teks
1	rakyat banyak udah bener bener abis deh kenaikan btm
2	sangat malu malu gara gara mesin dapat bantuan free sudah tidak mesin lagi tapi yang tidak mesin lagi juga tidak bisa dan hampir mesin dia tidak dapat bantuan sekarang dia yang jadi mesin lagi
3	peraga kenaikan konsumsi btm di jalang dan dit itu presiden perdana
4	19 tahun transkripsi ke pdf jadi udah rp 2000 awalnya akan jadi rp 10000 itu sudah mengalami beberapa kali pengisian presiden pemerintah di dan sudah beberapa kali mengalami kenaikan btm
5	terus listriknya dari bahan bakar fosilnya aja samain aja dia dapat provikan besar mobil listrik 3'4, 6'6, 8'8, 10'10, 12'12, 14'14, 16'16, 18'18, 20'20, 22'22, 24'24, 26'26, 28'28, 30'30, 32'32, 34'34, 36'36, 38'38, 40'40, 42'42, 44'44, 46'46, 48'48, 50'50, 52'52, 54'54, 56'56, 58'58, 60'60, 62'62, 64'64, 66'66, 68'68, 70'70, 72'72, 74'74, 76'76, 78'78, 80'80, 82'82, 84'84, 86'86, 88'88, 90'90, 92'92, 94'94, 96'96, 98'98, 100'100

Gambar 11 Sesudah Proses Case Folding

Pada proses ini penulis melakukan menyamaratakan antara uppercase dan lowercase huruf. Pengguna twitter dalam menuliskan tweet tidak terlalu memperhatikan uppercase dan lowercase sehingga data yang didapatkan sangatlah beragam. Jadi disini penulis harus melakukan case folding hal ini bertujuan supaya sistem tidak membaca kata yang sama lebih dari satu kali dikarenakan adanya perbedaan uppercase dan lowercase

4) Filter Stopword

Proses ini merupakan proses dimana penulis membuang semua kata hubung seperti kata dan, supaya, tetapi dan lain sebagainya. Tujuannya adalah supaya mempercepat sistem.

4.3. Labeling Data

Row No.	Sentimen	Teks
8	Positif	BERITA Peningkatan Harga Bahan Pokok Annon Peningkatan BBM Kapanas, Jember dan Dandim Cak Palar
32	Positif	1000 Paket Sembako Bagi Warga Terdampak Kenaikan Harga BBM Penderita Diabetes Sebesar 4 TDR
23	Positif	1000 Paket Sembako Bagi Warga Terdampak Kenaikan Harga BBM Penderita Diabetes Sebesar 4 TDR
30	Positif	1480 KPM di Kabupaten Tembung DIT Cempaka Kemahasiswaan
39	Positif	1480 Mahasiswa Kalingan Kota Tembung DIT Cempaka Kemahasiswaan
40	Positif	1000 Paket Sembako Untuk Warga Terdampak Kenaikan Harga BBM
44	Positif	1000 Paket Sembako Untuk Warga Terdampak Kenaikan Harga BBM Rp 44 - Bina ANTARA News
45	Positif	1070 Paket Sembako Untuk Warga Terdampak Kenaikan Harga BBM
59	Positif	200 Paket Sembako Untuk Warga Terdampak Kenaikan Harga BBM Rp 200 - Bina ANTARA News
60	Positif	200 Paket Sembako Untuk Warga Terdampak Kenaikan Harga BBM Rp 200 - Bina ANTARA News
61	Positif	2000 Paket Sembako Untuk Warga Terdampak Kenaikan Harga BBM Rp 2000 - Bina ANTARA News
72	Positif	200 Paket Sembako Untuk Warga Terdampak Kenaikan Harga BBM Rp 200 - Bina ANTARA News
73	Positif	2040 Paket Sembako Untuk Warga Terdampak Kenaikan Harga BBM Rp 2040 - Bina ANTARA News
106	Positif	100 Warga Desa Cempaka Dukung Kegiatan Komunitas Kesehatan Warga
111	Positif	83 Binar Warga Terdampak Kenaikan Harga BBM dan Disalurkan Bantuan Para Pemerintah
114	Positif	7000 Warga Desa Cempaka Dukung Kegiatan Komunitas Kesehatan Warga

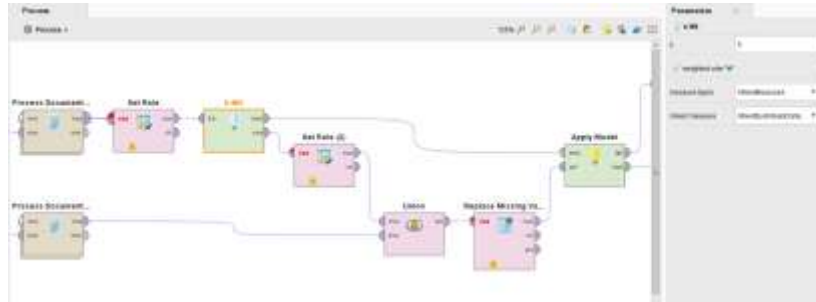
Gambar 12 Tahapan Pemberian Label Sentimen

Tabel 2 Pembagian Data

Kata Kunci	Data awal	Data Akhir	Positif	Negatif
Kenaikan BBM	20000	9253	25	81

Pada proses ini penulis memberikan label dalam bentuk sentimen dengan dua kategori yaitu positif dan negatif. Data yang memiliki label inilah yang nantinya akan menjadi data uji dan model yang akan digunakan sistem untuk mengklasifikasi data lainnya dalam penelitian ini. Pembagian data penelitian ini dapat pada tabel 1. Ada 106 data yang diberikan label sentimen dengan 25 sentimen positif dan 81 sentimen negatif.

4.4. Uji Klasifikasi Menggunakan K-Nearest Neighbor



Gambar 13 Proses dan Operator yang Digunakan

Pada proses ini penulis melakukan training data. Dengan menggunakan Rapidminer data yang sudah diberikan label dimasukkan ke operator KNN dengan jarak k sebesar 5 lalu sistem akan mempelajari data dan menghasilkan model data. Lalu selanjutnya data model akan diaplikasikan dalam klasifikasi data uji pada proses operator *apply module*.

4.5. Evaluasi

Pada proses ini penulis meninjau hasil klasifikasi dengan menggunakan confusion matrix untuk mengetahui seberapa akuratnya sistem dalam mengklasifikasikan data. Jika belum memenuhi target kriteria yang diinginkan penulis maka sistem ini perlu dikaji ulang kembali.

4.6. Pembahasan



Gambar 14 Confusion Matriks

Peneliti membagi data menjadi dua yaitu data training dan data testing. Peneliti menggunakan tools rapidminer untuk menjalankan semua proses penelitian ini. Adapun hasil yang diinginkan peneliti antara lain:

1) Akurasi

Akurasi adalah tingkat keakuratan sistem dalam melakukan prediksi dengan label pada data training. Data training disini menggunakan data yang sudah diberikan label melalui tahap labelling. Adapun proses perhitungan dari tingkat akurasi sebagai berikut:

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP + TN}{TP + FN + FP + TN} \times 100\% \dots\dots\dots(2) \\
 &= \frac{21 + 79}{21 + 4 + 2 + 79} \times 100\% \\
 &= \frac{100}{106} \times 100\% = 94,33\%
 \end{aligned}$$

Keterangan:

TP = *True Positif*

TN = *True Negatif*

FN = *False Negatif*

FP = *False Positif*

2) Presisi

Presisi adalah perbandingan rasio antara prediksi benar positif dengan semua hasil yang diprediksi sistem sebagai positif. Presisi dalam penelitian dihitung menggunakan rapidminer juga. Adapun perhitungan manual presisi sebagai berikut:

$$\begin{aligned} \text{Presisi} &= \frac{TP}{TP + FP} \times 100 \% \dots\dots\dots(3) \\ &= \frac{21}{21 + 2} \times 100 \% \\ &= \frac{21}{23} \times 100 \% = 91,3 \% \end{aligned}$$

Keterangan:

TP = *True Positif*

TN = *True Negatif*

FN = *False Negatif*

FP = *False Positif*

3) Recall

Recall adalah perbandingan rasio antara prediksi positif dengan semua data yang bernilai benar positif. Recall dalam penelitian dihitung menggunakan rapidminer juga. Adapun perhitungan manual recall sebagai berikut

$$\begin{aligned} \text{Recall} &= \frac{TP}{TP + FN} \times 100 \% \dots\dots\dots(3) \\ &= \frac{21}{21 + 4} \times 100 \% \\ &= \frac{21}{25} \times 100 \% = 84\% \end{aligned}$$

Keterangan:

TP = *True Positif*

TN = *True Negatif*

FN = *False Negatif*

FP = *False Positif*

5. Simpulan

Sentimen masyarakat terhadap keputusan pemerintah mengenai kenaikan BBM sangatlah beragam ada pengguna twitter yang memahami alasan kenaikan BBM dan mendukung keputusan pemerintah, tetapi ada juga yang menolak kenaikan BBM. Pada penelitian ini dapat disimpulkan sentimen mengenai kenaikan BBM cenderung negatif dan penggunaan K-Nearest Neighbor cukup baik jika melihat dari hasil dari penelitian ini dengan akurasi 94,33%, presisi 91,3%, recall 84%. Harapannya penelitian ini dapat digunakan sebagai sistem pendukung keputusan bagi pemerintah dalam membuat kebijakan-kebijakan kedepannya. Selain itu dari penelitian ini juga mempermudah pemerintah dalam melihat bagaimana reaksi dan sentimen masyarakat serta mengukur tingkat pemahaman masyarakat dalam memahami kebijakan yang dibuat pemerintah. Dengan begitu pemerintah dapat membuat kebijakan yang lebih bijak dengan mendengarkan suara rakyat lebih baik.

Daftar Referensi

- [1] H. Ghilman Rozy, A. Nuri, "Analisis Dampak Kebijakan Perubahan Publik Harga BBM terhadap Perekonomian Rakyat Indonesia," *Jurnal Ilmu Komputer, Ekonomi dan Manajemen (JIKEM)*, vol. 2, no. 1, pp.1464-1474, 2022
- [2] P. Ayu Nur Irwinesia, "Analisis Framing Berita Demonstrasi Mahasiswa Semarang Terkait Kenaikan Harga Bbm Pada Tv Borobudur", *THE MESSENGER*, vol. IV, no. 1, 2022
- [3] R. Andrianto, "Ini Bukti Harga Minyak Liar karena Perang Rusia-Ukraina," *Cnbcindonesia*, 2022. [Online]. Available: <https://www.cnbcindonesia.com/market/20220601152311-17-343480/ini-bukti-harga-minyak-liar-karena-perang-rusia-ukraina>. [Diakses 01 Juni 2022].
- [4] A. Ramadhan, "Harga BBM Pertalite, Solar, hingga Pertamina Resmi Naik Mulai Hari Ini," *Cnbcindonesia*, 2022. [Online]. Available: <https://www.cnbcindonesia.com/market/20220601152311-17-343480/ini-bukti-harga-minyak-liar-karena-perang-rusia-ukraina>. [Diakses 30 September 2022].
- [5] W. Wardani, S. S. U. Arfah, Z. and P. S. Lubis, "Dampak kenaikan Bahan Bakar Minyak (BBM) Terhadap Inflasi dan Implikasinya Terhadap Makroekonomi di Indonesia," *All Fields of Science J-LAS*, vol. 2, no. 3, pp. 63-70, 2022.
- [6] D. Yuliani, S. Saryono, D. Apriani, M. and M. Ro, "Dampak Kenaikan Harga Bahan Bakar Minyak (BBM) Terhadap Sembilan Bahan Pokok (Sembako) Di Kecamatan Tambun Selatan Dalam Masa Pandemi," *Jurnal Citizenship Virtues*, vol. 2, no. 2, pp. 320-326, 2022.
- [7] K. Ulfa, S. T. Suseno, "Analisis Sentimen Terhadap Bantuan Subsidi Upah (BSU) pada Kenaikan Harga Bahan Bakar Minyak (BBM)", *JURNAL MEDIA INFORMATIKA BUDIDARMA* vol. 6, no. 4, pp. 2335-2340, 2022.
- [8] G. Muhammad, S. R. Ragam, W. Sejati, "Analisis Sentimen Twitter Terhadap Kenaikan Bahan Bakar Minyak Menggunakan Algoritme Naïve Bayes", 2nd Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI), vol. 2, no. 1, April.
- [9] A. Asro'i dan F. Hery, "Analisis Sentimen Pengguna Twitter terhadap Perpanjangan PPKM Menggunakan Metode K-Nearest Neighbor," *Jurnal Khatulistiwa Informatika*, vol. 10, no. 1, pp. 17-24, 2022.
- [10] R. Savira, A. Solichin, I. dan M. Syafrullah, "Analisis Sentimen Pada Twitter Terhadap Kenaikan BBM 2022 Dengan Lexicon dan Support Vector Machine," *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, vol. 2, no. 1, pp. 211 - 218, 2023.
- [11] A. Hermawan, I. Jowensen, J. dan E. , "Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritme Support Vector Machine," *Jurnal Sains dan Teknologi*, vol. 12, no. 2, pp. 129-137, 2023.
- [12] A. Firdaus dan W. I. Firdaus, "Text Mining Dan Pola Algoritme Dalam Penyelesaian Masalah Informasi: (Sebuah Ulasan)," *Jurnal Jupiter*, vol. 13, no. 1, pp. 66 - 78, 2021.
- [13] H. Leiydiana, "Penerapan Algoritme K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor," *Jurnal Penelitian Ilmu Komputer, System Embedded & Logic*, vol. 1, no. 1, pp. 65 - 76, 2013.
- [14] G. Okfalisa, M. I dan N. Reza, "Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification," dalam *Proceedings - 2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering*, 2018.
- [15] Binus, "confusion-matrix," Binus, 2020. [Online]. Available: <https://socs.binus.ac.id/2020/11/01/confusion-matrix/>. [Diakses 18 November 2022].
- [16] M. R. Nahjan, N. Heryana dan A. Voutama, "Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko OJ CELL," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 101 - 104, 2023.
- [17] Z. Efendi and Mustakim, "Text Mining Classification Sebagai Rekomendasi Dosen Pembimbing Tugas Akhir Program Studi Sistem Informasi," in *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI)*, 2017.