



Analisa sentimen terhadap kenaikan BBM di twitter (x) menggunakan naive bayes classifier

Muhammad Abdillah¹, Muhammad Fikry², Yusra³, Alwis Nazir⁴, Fitri Insani⁵

Email: ¹11751100393@students.uin-suka.ac.id, ²muhammad.fikry@uin-suska.ac.id, ³yusra@uin-suska.ac.id,

⁴alwis.nazir@uin-suska.ac.id, ⁵fitri.insani@uin-suska.ac.id

¹²³⁴⁵Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau

Diterima: 05 April 2024 | Direvisi: - | Disetujui: 28 April 2024

©2024 Program Studi Teknik Informatika Fakultas Ilmu Komputer,
Universitas Muhammadiyah Riau, Indonesia

Abstrak

Pada awal September 2022 digemparkan oleh berita naiknya BBM. Pemerintah memutuskan menaikkan harga BBM karena melonjaknya harga minyak dunia. PT Pertamina (Persero) resmi menaikkan harga Bahan Bakar Minyak (BBM) pertiga September 2022, jam 14:30 WIB. Sejak keputusan tersebut akan menimbulkan opini dari masyarakat. Masyarakat banyak memberikan tanggapannya melalui sosial media Twitter, baik berupa tanggapan positif maupun tanggapan negatif. Sehingga hal tersebut masyarakat berikan sentimen positif dan negatif. Data yang digunakan sebanyak 3.000 data *tweet* dengan kata kunci "KENAIKAN BBM," data diambil pada 1 November 2022 hingga 1 Desember 2022. Penelitian ini menggunakan metode Naive Bayes Classifier, dilakukan dengan tiga perbandingan dengan *threshold* (0.001 hingga 0.007). Percobaan ini akan dilakukan dengan tiga pengujian data yaitu data opini, data campuran (opini-non opini) dan data seimbang. Berikut hasil pengujian pada data opini mendapatkan akurasi tertinggi sebesar 80% dengan perbandingan 90:10, data campuran mendapatkan akurasi 67.7% dengan perbandingan 70:30 dan data seimbang mendapatkan akurasi 63.6%, dengan perbandingan 90:10.

Kata kunci: Kenaikan BBM, Twitter, Naive Bayes Classifier

Sentiment analysis on the increase of fuel prices on twitter (x) using naive bayes classifier.

Abstract

In early September 2022, there was a shock from the news of the rise in fuel prices. The government decided to increase the price of fuel due to the surge in world oil prices. PT Pertamina (Persero) officially raised the price of Fuel Oil (BBM) one-third of September 2022, at 2:30 PM WIB (Western Indonesia Time). Since the decision, it has sparked opinions from the public. Many people expressed their responses through the social media platform Twitter, both in positive and negative ways. This resulted in both positive and negative sentiments from the public. The data used consisted of 3,000 *tweets* with the keyword "FUEL PRICE INCREASE," collected from November 1, 2022, to December 1, 2022. This research utilized the Naive Bayes Classifier method, conducted with three comparisons using *thresholds* ranging from 0.001 to 0.007. The experiment was conducted with three types of data testing: opinion data, mixed data (opinion-non-opinion), and balanced data. Here are the test results: for opinion data, the highest accuracy obtained was 80% with a ratio of 90:10, for mixed data, the accuracy obtained was 67.7% with a ratio of 70:30, and for balanced data, the accuracy obtained was 63.6% with a ratio of 90:10.

Keywords: Fuel Price Increase, Twitter, Naive Bayes Classifier

1. PENDAHULUAN

Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap suatu masalah atau objek oleh seseorang, apakah cenderung beropini positif atau negatif [8]. Sentimen analisis (*opinion mining*) adalah bidang studi yang menganalisis

opini, sentimen, evaluasi, penilaian, sikap, dan emosi orang-orang terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atributnya [1]. Twitter juga sering digunakan oleh masyarakat untuk mengungkapkan perasaan mereka tentang peristiwa terkini. Masyarakat sering mengungkapkan perasaan mereka, baik senang maupun sedih, di media sosial ini [3]. Sentimen ini menghasilkan sebuah data berupa pendapat masyarakat dalam bentuk teks yang merupakan bentuk ekspresi dari masyarakat di media sosial Twitter.

Dengan berkembangnya informasi semakin pesat dan cepat untuk mendapatkan informasi. Salah satunya media sosial Twitter, yang dapat memberikan tanggapan (*tweet*) disetiap penggunaannya atas sebuah peristiwa yang terjadi. Twitter merupakan salah satu media sosial yang memungkinkan penggunaannya untuk mengirim dan membaca pesan yang berbasis teks. Pengguna Twitter banyak membicarakan isu-isu hangat dari berbagai dimensi, mulai dari *entertainment*, politik, sosial, hingga pemerintahan [4]. Total pengguna Twitter di dunia mencapai 436 juta dan pengguna Twitter di Indonesia mencapai 18.45 juta di Januari 2022 (Sumber: DataIndonesia.id). Pada awal September Tahun 2022, di Indonesia, Twitter diramaikan dengan isu tentang kenaikan BBM. Banyak orang di media sosial, terutama Twitter, memuji atau menyalahkan tindakan pemerintah saat ini, khususnya pada pemerintahan Presiden Joko Widodo [5].

Banyak yang menuliskan pendapatnya mengenai isu tersebut baik berupa opini positif dan negatif. Hal ini menjadi menarik dan penting bagi pihak-pihak tertentu yang ingin mengetahui baik buruknya sentimen atau opini yang diberikan masyarakat Indonesia terhadap kenaikan BBM tersebut. Penelitian ini menggunakan metode algoritma Naive Bayes Classifier dengan mengklasifikasikan sentimen masyarakat terhadap kenaikan BBM di Twitter. Sentimen analisis dapat diselesaikan dengan metode Naive Bayes Classifier [7]. Metode Naive Bayes Classifier digunakan untuk mencari nilai probabilitas tertinggi untuk mengklasifikasi data uji pada kategori yang paling tepat. Kelebihan Naive Bayes Classifier adalah algoritmanya sederhana tetapi memiliki akurasi yang tinggi. Terdapat dua tahapan dalam klasifikasi *tweet*. Tahap pertama adalah pelatihan terhadap *tweet* yang telah diketahui kategorinya (*training*), sedangkan tahap kedua adalah proses klasifikasi *tweet* yang belum diketahui kategorinya (*testing*) [2]. Sebelum diklasifikasikan, data harus dipreprocessing terlebih dahulu. Setelah diklasifikasikan, *tweet* akan memiliki makna positif atau negatif [6]. *Preprocessing* adalah salah satu tahapan menghilangkan masalah yang dapat mengganggu hasil proses data [9].

Metode Naive Bayes Classifier terdiri dari dua tahap dalam proses klasifikasi data teks yaitu tahap pelatihan dan tahap *testing*. Tahap pelatihan yaitu membentuk model dengan melatih model menggunakan data *training*. Tahap kedua yaitu klasifikasi, pertama memperkirakan ketepatan prediksi dari model yang telah dilatih. Model yang telah dibangun digunakan untuk mengklasifikasikan data *training* dan data *testing* ke dalam kelas sentimen positif atau negatif. Pengukuran ketepatan prediksi dilakukan dengan membentuk *confusion matrix* dari hasil prediksi [2]. *Confusion matrix* merupakan sebuah tabel yang memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh sistem (prediksi) dengan hasil klasifikasi yang sebenarnya.

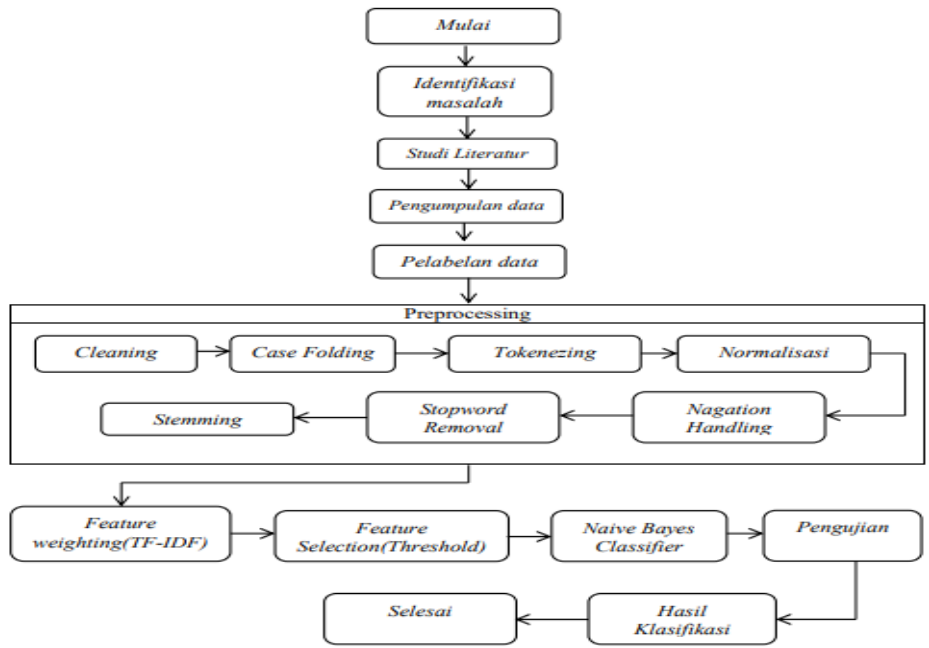
Setelah terbentuk *confusion matrix*, langkah selanjutnya adalah melakukan perhitungan ketepatan klasifikasi. Hasil ketepatan klasifikasi menunjukkan besarnya nilai yang tepat terklasifikasi pada kelas aktual. Berikut merupakan hasil pengukuran ketepatan klasifikasi dari setiap iterasi menggunakan algoritma Naive Bayes Classifier [2]. Penelitian yang dilakukan oleh Khairul Anam dkk. Hasil akurasi yang didapatkan cukup beragam yaitu dari metode Naive Bayes Classifier akurasi 100%, metode K-Nearest Neighbor akurasi 98.25%, dan metode Decision Tree akurasi 62.28%. Dapat disimpulkan Naive Bayes Classifier unggul dari metode K-Nearest Neighbor dan Decision Tree [10].

Data yang digunakan sebanyak 3.000 data *tweet* dengan kata kunci “KENAIKAN BBM,” data diambil pada 1 November 2022 hingga 1 Desember 2022 dan menggunakan *tools* Google Colab untuk melakukan implementasi untuk menguji metode Naive Bayes Classifier serta mengklasifikasi sentimen positif dan negatif masyarakat terhadap kenaikan BBM.

2. METODE PENELITIAN

2.1. Tahap Desain Penelitian

Pada tahapan desain penelitian ini, diawali dengan identifikasi masalah, studi literatur, pengumpulan data di Twitter, pelabelan data menggunakan validator. Selanjutnya akan dilakukan *preprocessing*, pembobotan kata TD-IDF, *feature selection*, klasifikasi Naive Bayes Classifier, pengujian dan hasil klasifikasi. Dapat dilihat pada Gambar 1.



Gambar 1. Tahap Desain Penelitian
Sumber: Hasil pada Gambar 1 (Sofiah, 2023, hal.139).

- a. Identifikasi Masalah
Pada identifikasi masalah ini suatu proses digunakan untuk merumuskan suatu masalah serta menemukan solusi dari sebuah masalah yang didapatkan.
- b. Studi Literatur
Pada studi literatur ini proses pengumpulan referensi disajikan secara mendalam serta dapat memudahkan dalam pengerjaan jurnal pada penelitian ini.
- c. Pengumpulan Data
Pada pengumpulan data ini, diambil melalui media sosial Twitter sebanyak 3.000 data *tweet* dengan kata kunci “KENAIKAN BBM,” data tersebut diambil pada 1 November 2022 hingga 1 Desember 2022.

31	7/11/22 10:55	1,59E+23	Penggemar Mie Ayan	b'@cnindonesia wah enak dong ya jadi komut. gak pusing mikirin kerjaan tp hidup sejahtera negara yang bayarin. gede pula gajinya. giliran
32	7/11/22 10:09 Jkt, ID	1,59E+23	Ayu Ari	b'biasanya dijemputayang naik fortuner, semenjak bbm subsidi dialihkan, dijemput naik beat. gpp si, yang penting masih bisa pelukan wala
33	7/11/22 9:16	1,59E+23	Serigala Subuh	b'@akunrec09153247 @emfajar @detikcom elu paranormal tong?nyng 1 udah jelas2 gagal. masih elu puja2nyng belum coba udah elu kate gi
34	7/11/22 8:27 Kota Bukittinggi, Suma	1,59E+23	jisman usman	b'kenaikan bbm untuk semua masyarakat. dukung penyesuaian harga bbm\n#bantuanbbmuntukakyat\n#bibtbbmtepatasaran\n#nsubsidi
35	7/11/22 7:53	1,59E+23	Slamet Aye	b'@teddus sok tau aja coba kamu keliling ke pelosok kekampung kampung si miskin yg menderita dari efek kenaikan bbm'
36	7/11/22 7:28	1,59E+23	Menolak Panik	b'@cnindonesia pantas disalahkan...kalau nggak mau salah jangan jabat dan jangan beri kebijakan kenaikan bbm...nyusahin aja'
37	7/11/22 6:42	1,59E+23	Yuki Kunnn	b'kenaikan bbm ini untuk semua masyarakat. mari kita bersama-sama mendukung apa yang sudah diprogramkan oleh pemerintah\n\n#sumu
38	7/11/22 6:33	1,59E+23	achm bad	b'@temansenyum @cnindonesia memang harga gak mahal2, tapi efek lainnya, contoh ini klu kita naik bis dari surabaya ke probolinggo dul
39	7/11/22 6:25 Indonesia	1,59E+22	OTVOTATAA&A&A&A	b'@flowerbe4n_ pemerintahan lagi bisnis sama rakyatnya, vaksin covid, kenaikan harga bbm, kendaraan listrik, tv digital, bentar lagi
40	7/11/22 6:11	1,59E+22	Arief Usman	b'@erickthohir contoh yg mesti anda lakukan...ncobalah untuk merangkul semua masyarakat yang hidup dalam keadaan susah setelah kena
41	7/11/22 6:09 DKI Jakarta, Indonesia	1,59E+22	Catatan PakA, A* A* A	b'@teddus kayaknya "rakat" dimonopoli kelompok pro jokowi ya... harusnya para pendukung jokowi, mendukung keluhan masyarakat
42	7/11/22 6:02 Nusa Tenggara Barat, It	1,59E+22	OTDA NTB	b'kenaikan bbm ini untuk semua masyarakat... yuk dukung program pemerintah ini. \n\n#bbmsubsidi\n#bbmbersubsidi\n#bantuanbbmunt
43	7/11/22 5:32	1,59E+23	Herry	b'@temansenyum @cnindonesia klo km org mampu boleh ngomong gak mahal, tapi org dicesa / org miskin hrg segitu mahal goblokkkk.\n
44	7/11/22 5:25	1,59E+23	Generous_Millami	b'pastinya kenaikan harga bbm untuk kebaikan masyarakat ya \nndaerah istimewa yogyakarta\n#bantuanbbmuntukakyat\n#bibtbbmtepatas
45	6/11/22 11:44 Kendal jateng	1,59E+23	rizqianurkry	b'@henrysubiakto wa'alaikum,,, \nknokelah prof makmur untk anda mungkin, tapi untk mayoritas ekonomi lemah masih jauh prof,,, bkn saya
46	6/11/22 9:59	1,59E+22	WidyaaLuna	b'meski terjadi kenaikan, harga bbm subsidi di indonesia masih cukup murah dibandingkan dengan negara tetangga di asean lainnya #bbms
47	6/11/22 9:26	1,59E+23	Ayu Azalia	b'meski harga bbm indonesia mengalami kenaikan tetapi harga bbm indonesia lebih murah dibanding negara tetangga #bbmsubsidiuntukw
48	6/11/22 9:24	1,59E+23	Akhum	b'@bunduloh ujung ujungnya duit... \n(udul)nyng untung siapa..? yg menderita pasti rakyat..? \nnapa belum cukup dengan kenaikan bbm dan akit
49	6/11/22 8:32 DKI Jakarta, Indonesia	1,59E+23	TedGaraAlam	b'di pemerintahan sbg saya termasuk paling keras dim aksi massa menolak kenaikan bbm yg ditimpali dgn bit, hingga hari ini sy menolak bi
50	6/11/22 8:13 Indonesia	1,59E+23	Imelda Sari	b'kenaikan harga bbm jelas berdampak pd sendi2 ekonomi masy pasca didera pandemi covid-19, inflasi &mp
51	6/11/22 5:00 Papua Barat, Indonesia	1,59E+23	A* A, A* A*	b'kebijakan kenaikan bbm upaya pemerintah agar subsidi tepat sasaran. \n#bantuanbbmuntukakyat\n#bibtbbmtepatasaran\n#nsubsidi
52	6/11/22 3:20 Jakarta	1,59E+22	Udit K Hariono	b'@erwinarnada protes kenaikan bbm. \nngiliran bit paling depan \nxf0vx9fvx98vx82\n#memang nature nya gitu'
53	6/11/22 3:08 Indonesia	1,59E+23	AndiA, seiA* A* A, se	b'ini sama aja dengan ga setuju kenaikan bbm tinggal boikot bbm. kan bisa jalan kaki, mudah to. \nhttps://t.co/nmlsgaskqm'
54	6/11/22 1:20	1,59E+23	Alfahmi	b'kurawa bang, bapak tai indonesia gimana kabar sekarang? lagi nikmatin duit hasil kenaikan bbm bukan?
55	6/11/22 0:44 Banten, Indonesia	1,59E+23	hadl	b'@tempodotco setuju banget pa \n#pemerintah tidak tanggap dengan berbagai kesulitan rakyat karena efek kenaikan harga bbm. \n#pejabat

Gambar 2. Tahap Pengumpulan Data
Sumber: Hasil pada Gambar 3 (Tiara Dwi Arsita, 2023, hal.144).

- d. Pelabelan Data
Pelabelan data ialah memberikan label opini positif, negatif dan non opini pada sentimen tertentu di Twitter. Tahap pelabelan menggunakan seorang validator yakni dosen Bahasa Indonesia di Universitas Riau yang bernama ibu Elfina, S.Pd., M.Pd. yang telah melakukan validasi klasifikasi setimen analisis kenaikan BBM pada data sentimen opini positif dan opini negatif.

Tabel 1. Sample Dataset

Dataset	Pelabelan
b'BPJS Klaim Kenaikan Harga BBM Tak Pengaruhi Ekonomi Triwulan III https://t.co/3kmxe09in0'	POSITIF

2.2. Preprocessing

Preprocessing ialah tahapan mengolah data mentah dari sentimen kenaikan BBM di Twitter yang diubah menjadi data terstruktur agar mendapatkan sebuah hasil.

2.3. Cleaning

Cleaning berfungsi proses membersihkan data dari kata-kata yang tidak penting, seperti simbol, URL, *mention* (@), *hashtag* (#), *emoji* dan *emoticon*.

Tabel 2. Proses *Cleaning*

Sebelum <i>Cleaning</i>	Hasil <i>Cleaning</i>
b'BPJS Klaim Kenaikan Harga BBM Tak Pengaruhi Ekonomi Triwulan https://t.co/3kmxe09in0'	BPJS Klaim Kenaikan Harga BBM Tak Pengaruhi Ekonomi Triwulan

Sumber: Hasil dari Tabel 2 (Tiara Dwi Arista, 2023, hal.145)

2.4. Casefolding

Casefolding berfungsi menyamaratakan penggunaan huruf kapital pada sentimen tersebut menjadi huruf kecil.

Tabel 3. Proses *Casefolding*

Sebelum <i>Casefolding</i>	Hasil <i>Casefolding</i>
BPJS Klaim Kenaikan Harga BBM Tak Pengaruhi Ekonomi Triwulan	bpjs klaim kenaikan harga bbm tak pengaruhi ekonomi triwulan

2.5. Tokenizing

Tokenizing ialah sebagai langkah pembagian teks pada dokumen menjadi token atau sebuah kalimat yang dibuat menjadi kata perkata.

Tabel 4. Proses *Tokenizing*

Sebelum <i>Tokenizing</i>	Hasil <i>Tokenizing</i>
bpjs klaim kenaikan harga bbm tak pengaruhi ekonomi triwulan	[bpjs], [klaim], [kenaikan], [harga], [bbm], [tak], [pengaruhi], [ekonomi], [triwulan]

2.6. Normalisasi

Normalisasi ialah langkah mengubah bentuk penulisan masing-masing kata sesuai dengan KBBI/kata baku.

Tabel 5. Proses Normalisasi

Sebelum normalisasi	Hasil normalisasi
bpjs klaim kenaikan harga bbm tak pengaruhi ekonomi triwulan	bpjs klaim kenaikan harga bbm tak pengaruhi ekonomi triwulan

2.7. Nagation Handling

Nagation handling ialah kata-kata negatif yang memengaruhi orientasi sentimen (mengubah kata yang mengandung negatif menjadi arti sebenarnya). Kata-kata yang lain dalam sebuah frase, seperti jangan, bukan, dan tidak.

Tabel 6. Proses *Nagation Handling*

Sebelum <i>Nagation Handling</i>	Hasil <i>Nagation Handling</i>
bpjs klaim kenaikan harga bbm tak pengaruhi ekonomi triwulan	bpjs klaim kenaikan harga bbm mendukung ekonomi triwulan

2.8. Stopword Removal

Stopword removal ialah langkah untuk menghilangkan kata-kata yang tidak penting, seperti kata hubung (di, ke, dan, dari, akan).

Tabel 7. Proses Stopword Removal

Sebelum Stopword Removal	Hasil Stopword Removal
bpjs klaim kenaikan harga bbm mendukung ekonomi triwulan	bpjs klaim kenaikan harga bbm mendukung ekonomi triwulan

2.9. Stemming

Stemming ialah untuk mengubah kata menjadi bentuk dasar atau menghilangkan imbuhan pada awalan dan akhiran pada sebuah kata menjadi kata dasar.

Tabel 8. Proses Stemming

Sebelum Stemming	Hasil Stemming
bpjs klaim kenaikan harga bbm mendukung ekonomi triwulan	bpjs klaim kenaikan harga bbm dukung ekonomi triwulan

2.10. Feature Weight

Setelah melakukan preprocessing, selanjutnya setiap kata akan diberikan nilai bobot menggunakan TF-IDF. Langkah-langkah melakukan pembobotan kata:

TF (Term Frequency (frekuensi istilah)). Menghitung frekuensi kata pada setiap dokumen

IDF (Inverse Document Frequency (frekuensi berkas terbalik)). Menghitung IDF menggunakan rumus:

$$IDF = \log \frac{d}{df}$$

Keterangan:

IDF = Bobot dari IDF

d = Jumlah seluruh dokumen

df = Jumlah frekuensi kemunculan

Hasil dari feature weight menggunakan tools Google Colab, dapat dilihat pada Gambar 3.

	000	09	10	100	1001	100k	100km	100persen	100rb	1075	...	ywdh	zainal	zaim	zaman	zefanya	zerr	zholimi	zico	zona	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	positif
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	positif
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 3. Hasil Perhitungan TF-IDF
Sumber: Hasil pada Gambar 2 (Sofiah, 2023, hal.143)

2.11. Feature Selection

Setelah dilakukan pembobotan kata menggunakan TF-IDF, selanjutnya menghilangkan fitur yang tidak relevan/redundan agar lebih mudah ditafsirkan. Pada feature selection ini menggunakan threshold 0.004 menggunakan tools Google Colab, hasil dari feature selection dapat dilihat pada Gambar 4.

Hasil Feature Selection												
	harga	jalan	kali	mahal	masyarakat	naik	nya	rakyat	rokok	sisi	turun	Kelas
0	0.000000	0.000000	0.134878	0.000000	0.0	0.099976	0.0	0.000000	0.0	0.106369	0.0	negatif
1	0.000000	0.172308	0.000000	0.000000	0.0	0.172115	0.0	0.000000	0.0	0.000000	0.0	positif
2	0.000000	0.000000	0.080959	0.000000	0.0	0.120019	0.0	0.14272	0.0	0.000000	0.0	negatif
3	0.000000	0.000000	0.264213	0.227554	0.0	0.195844	0.0	0.000000	0.0	0.000000	0.0	positif
4	0.074174	0.000000	0.175090	0.000000	0.0	0.086522	0.0	0.000000	0.0	0.000000	0.0	negatif
...
1495	0.207315	0.080700	0.163126	0.000000	0.0	0.000000	0.0	0.000000	0.0	0.085764	0.0	positif
1496	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.0	0.000000	0.0	0.000000	0.0	negatif
1497	0.000000	0.000000	0.000000	0.000000	0.0	0.086541	0.0	0.000000	0.0	0.000000	0.0	negatif
1498	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.0	0.000000	0.0	0.000000	0.0	positif
1499	0.000000	0.000000	0.374036	0.000000	0.0	0.000000	0.0	0.000000	0.0	0.000000	0.0	positif

Gambar 4. Hasil Proses *Feature Selection*

2.12. Klasifikasi Naive Bayes Classifier

Pada tahapan ini akan dilakukan pengelompokan menggunakan metode Naive Bayes Classifier. Selanjutnya dilakukan *modeling* Naive Bayes Classifier untuk menghasilkan nilai *accuracy*, *precision*, *recall* dan *f1-Score* menggunakan *tools* Google Colab, hasil dari metode Naive Bayes Classifier dapat dilihat pada Gambar 5.

Hasil Klasifikasi												
	harga	jalan	kali	mahal	masyarakat	naik	nya	rakyat	rokok	sisi	turun	Kelas Target
0	0.000000	0.000000	0.000000	0.000000	0.0	0.080039	0.000000	0.190357	0.0	0.000000	0.0	negatif
1	0.000000	0.312041	0.210252	0.000000	0.0	0.155846	0.000000	0.000000	0.0	0.000000	0.0	positif
2	0.000000	0.096453	0.194970	0.000000	0.0	0.000000	0.000000	0.229137	0.0	0.000000	0.0	negatif
3	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.154192	0.0	negatif
4	0.000000	0.301796	0.101674	0.175135	0.0	0.000000	0.000000	0.000000	0.0	0.000000	0.0	negatif
...
145	0.169043	0.098703	0.066505	0.000000	0.0	0.197185	0.000000	0.000000	0.0	0.000000	0.0	positif
146	0.000000	0.000000	0.222207	0.000000	0.0	0.000000	0.160533	0.391722	0.0	0.000000	0.0	negatif
147	0.000000	0.000000	0.000000	0.000000	0.0	0.282801	0.000000	0.000000	0.0	0.300886	0.0	negatif
148	0.000000	0.000000	0.071392	0.000000	0.0	0.105836	0.154731	0.000000	0.0	0.000000	0.0	negatif
149	0.000000	0.264714	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.000000	0.0	negatif

Gambar 5. Hasil Klasifikasi Naive Bayes Classifier

3. HASIL DAN PEMBAHASAN

3.1. Pengujian

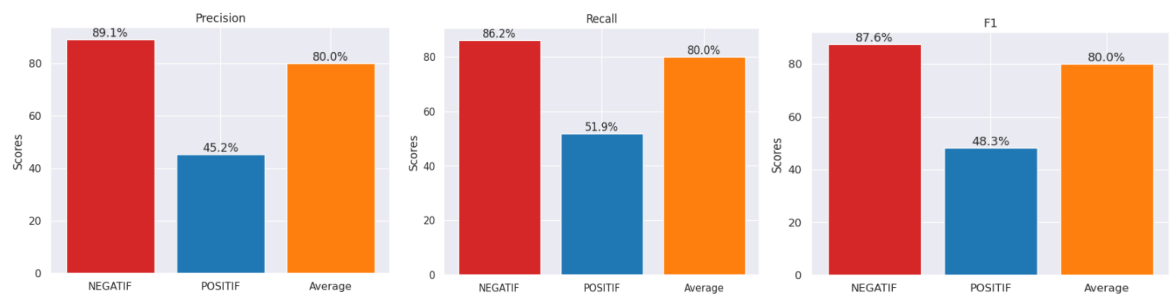
Pada pengujian akan dilakukan uji data *testing* dan *training*. Data ini menggunakan *threshold* 0.001 hingga 0.007 dengan menggunakan *confusion matrix*. data ini akan diuji ada tiga yaitu data opini, data campuran (opini-non opini) dan data seimbang. Pengujian dilakukan menggunakan perbandingan 70:30, 80:20, dan 90:10.

3.2. Pengujian Data Opini

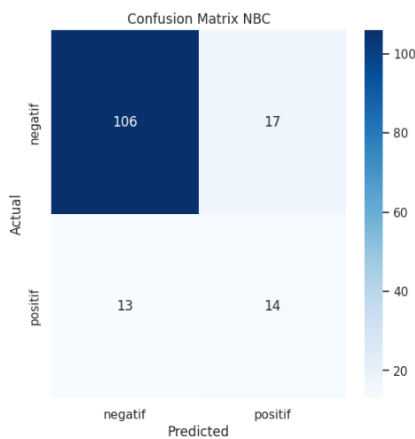
Pada pengujian ini menggunakan 1500 data opini yang terdiri dari 330 kelas positif dan 1170 kelas negatif. Hasil diperlihatkan pada tabel 9.

Tabel 9. Hasil Akurasi Data Opini			
Threshold	Accuracy		
	70:30	80:20	90:10
0.001	48.2%	51.7%	48.7%
0.002	48.7%	48.7%	42%
0.003	74.2%	76.3%	78.7%
0.004	77.8%	76.3%	80%
0.005	71.3%	71%	71.3%
0.006	67.1%	69.7%	70.7%
0.007	78.7%	76.3%	79.3%

Dari pengujian data opini, akurasi yang tertinggi pada perbandingan 90:10 dengan *accuracy* 80% , *precision* 80%, *recall* 80%, *f1-score* 80%, hasil perbandingannya dapat dilihat pada Gambar 6.



Gambar 6. Grafik Hasil Pengujian Data Opini



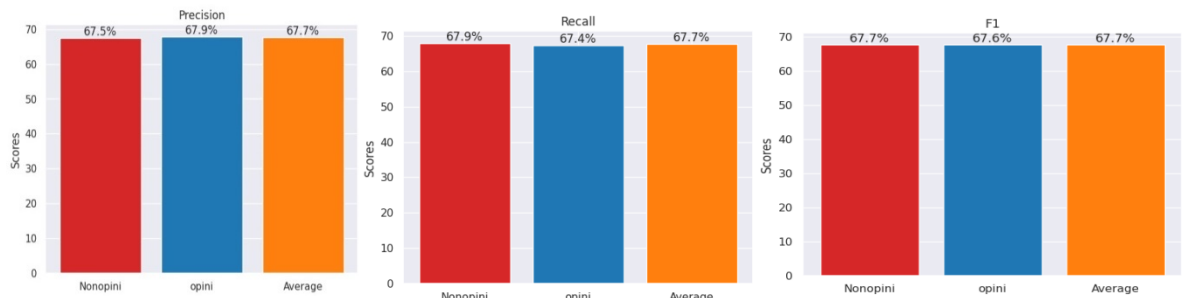
Gambar 7. Hasil Confusion Matrix Data Opini

3.3. Pengujian Data Campuran (opini-non opini)

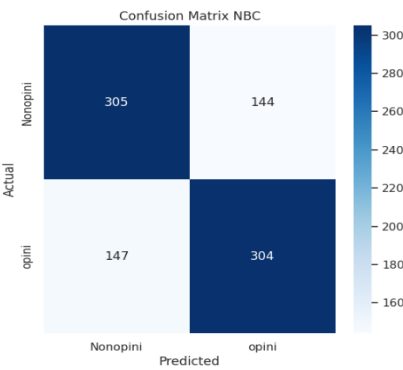
Pada pengujian ini menggunakan 3000 data campuran yang terdiri dari kelas opini dan non opini. Hasil diperlihatkan pada tabel 10.

Threshold	Accuracy		
	70:30	80:20	90:10
0.001	64.0%	61.8%	60.3%
0.002	66.1%	63.0%	62.7%
0.003	65.1%	63.5%	63.0%
0.004	63.3%	59.7%	57.3%
0.005	67.7%	65.0%	63.0%
0.006	54.4%	56.0%	56.3%
0.007	56.3%	54.7%	53.3%

Dari pengujian data campuran, akurasi yang tertinggi pada perbandingan 70:30 dengan accuracy 67.7%, precision 67.7%, recall 67.7%, f1-score 67.7%, hasil perbandingannya dapat dilihat pada Gambar 8.



Gambar 8. Grafik Hasil Pengujian Data Campuran



Gambar 9. Hasil *Confusion Matrix* Data Campuran

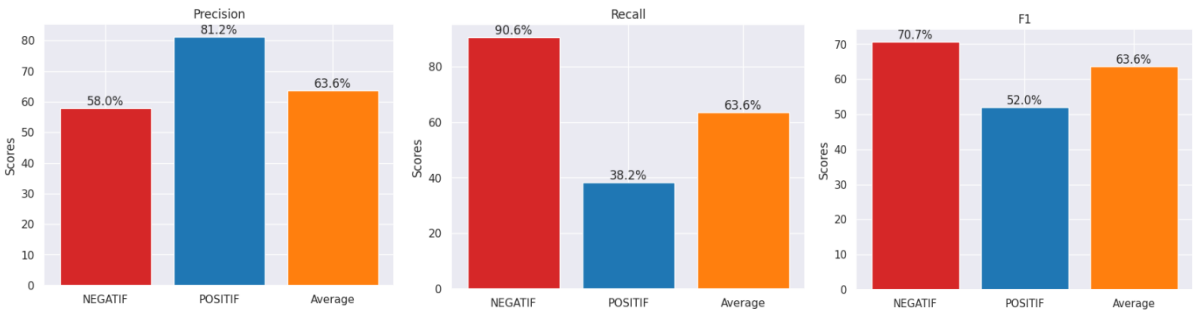
3.4 Pengujian Data Seimbang

Pada pengujian ini menggunakan 660 data yang terdiri dari 330 kelas positif dan negatif. Hasil diperlihatkan pada tabel 11.

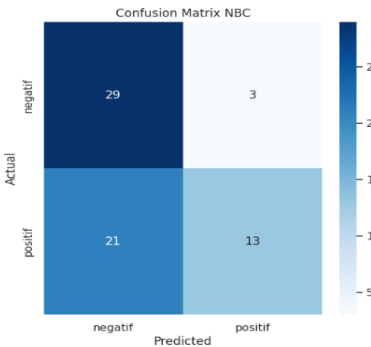
Tabel 11. Hasil Akurasi Data Seimbang

Threshold	Accuracy		
	70:30	80:20	90:10
0.001	61.6%	59.8%	54.5%
0.002	63.1%	62.1%	63.6%
0.003	61.1%	61.4%	57.6%
0.004	61.1%	58.3%	51.5%
0.005	59.6%	56.1%	50%
0.006	62.1%	59.8%	56.1%
0.007	59.1%	59.1%	56.1%

Dari pengujian data seimbang, akurasi yang tertinggi pada perbandingan 90:10 dengan *accuracy* 63.6% ,*precision* 63.6%,*recall* 63.6%,*f1-score* 63.6%, hasil perbandingannya dapat dilihat pada Gambar 10.



Gambar 10. Grafik Hasil Pengujian Data Seimbang



Gambar 11. Hasil *Confusion Matrix* Data Seimbang

4. KESIMPULAN

Kesimpulan berdasarkan hasil pengujian yang didapat pada penelitian ini dengan metode Naive Bayes Classifier dapat disimpulkan bahwa metode Naive Bayes Classifier pada studi kasus kenaikan BBM di Twitter. Dengan melakukan tiga pengujian data opini, data campuran, dan data seimbang. Akurasi data opini tertinggi yaitu didapatkan sebesar 80% pada perbandingan 90:10. Akurasi data campuran tertinggi yaitu didapatkan sebesar 67.7% pada perbandingan 70:30, dan Akurasi data seimbang tertinggi yaitu didapatkan sebesar 63.6% pada perbandingan 90:10. Oleh karena itu Naive Bayes Classifier memperoleh pengujian data opini dengan akurasi terbaik.

DAFTAR PUSTAKA

- [1] H. Sujadi, S. Fajar, and C. Roni, "Analisis Sentimen Pengguna Media Sosial Twitter Terhadap Wabah Covid-19 Dengan Metode Naive Bayes Classifier Dan Support Vector Machine," *INFOTECH J.*, vol. 8, no. 1, pp. 22–27, 2022, doi: 10.31949/infotech.v8i1.1883.
- [2] E. W. Novianti, and W. Wibowo, "Analisis Sentimen Pengguna Twitter terhadap Program Kartu Prakerja di Tengah," *J. Sains dan senin ITS*, vol. 11, no. 1, pp. 136–142, 2022.
- [3] M. W. A. Putra, Susanti, Erlin, and Herwin, "Analisis Sentimen Dompot Elektronik Pada Twitter Menggunakan Metode Naive Bayes Classifier," *IT J. Res. Dev.*, vol. 5, no. 1, pp. 72–86, 2020, doi: 10.25299/itjrd.2020.vol5(1).5159.
- [4] R. Nooraeni, A. B. Safiruddin, A. F. Afifah, K. D. Agung, and N. N. Rosyad, "Analisis Sentimen Publik terhadap Sistem Zonasi Sekolah Menggunakan Data Twitter dengan Metode Naive Bayes Classification," *Fakt. Exacta*, vol. 12, no. 4, p. 315, 2020, doi: 10.30998/faktorexacta.v12i4.5205.
- [5] S. Suryono and E. Taufiq Luthfi, "Analisis sentimen pada Twitter dengan menggunakan metode Naive Bayes Classifier," *Jnanaloka*, pp. 81–86, 2021, doi: 10.36802/jnanaloka.2020.v1-no2-81-86.
- [6] B. M. Pintoko and K. M. L., "Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naive Bayes Classifier," *e-Proceeding Eng.*, vol. 5, no. 3, pp. 8121–8130, 2018.
- [7] A. Pramono, R. Indriati, and A. Nugroho, "Sentiment Analysis Tokoh Politik Pada Twitter," *Semin. Nas. Inov. Teknol.*, pp. 195–200, 2017, [Online]. Available: <https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/403/317>
- [8] Z. Saharuna, P. Negeri, U. Pandang, M. Nur, and Y. Utomo, "Analisis Sentimen pada Twitter terhadap Pelayanan Pemerintah Kota Makassar Eddy Tungadi Politeknik Negeri Ujung Pandang," *Proceeding SNTET*, pp. 87–93, 2019, [Online]. Available: <https://dev.Twitter.com>
- [9] R. Apriani *et al.*, "Analisis Sentimen dengan Naive Bayes Terhadap Komentar Aplikasi Tokopedia," *J. Rekayasa Teknol. Nusa Putra*, vol. 6, no. 1, pp. 54–62, 2019, [Online]. Available: <https://rekayasa.nusaputra.ac.id/article/view/86>
- [10] M. K. Anam, B. N. Pikir, and M. B. Firdaus, "Penerapan Naive Bayes Classifier, K-Nearest Neighbor (KNN) dan Decision Tree untuk Menganalisis Sentimen pada Interaksi Netizen danPemeritah," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 1, pp. 139–150, 2021, doi: 10.30812/matrik.v21i1.1092.
- [11] F. S. Pattiha and H. Hendry, "Perbandingan Metode K-NN, Naive Bayes, Decision Tree untuk Analisis Sentimen Tweet Twitter Terkait Opini Terhadap PT PAL Indonesia," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 2, p. 506, 2022, doi: 10.30865/jurikom.v9i2.4016.
- [12] M. I. Fikri, T. S. Sabrila, Y. Azhar, and U. M. Malang, "Comparison of the Naive Bayes Method and Support Vector Machine on Twitter Sentiment Analysis," *SMATIKA J. STIKI Inform. J.*, vol. 10, no. 2, pp. 71–76, 2020.
- [13] R. D. Septiana, A. B. Susanto, and T. Tukiyat, "Analisis Sentimen Vaksinasi Covid-19 Pada Twitter Menggunakan Naive Bayes Classifier Dengan Feature Selection Chi-Squared Statistic dan Particle Swarm Optimization," *J. SISKOM-KB (Sistem Komput. dan Kecerdasan Buatan)*, vol. 5, no. 1, pp. 49–56, 2021, doi: 10.47970/siskom-kb.v5i1.228.
- [14] R. Haditira *et al.*, "Analisis Sentimen Pada Steam Review Menggunakan Metode Multinomial Naive Bayes dengan Seleksi Fitur Gini Index Text," *e-Proceeding Eng.*, vol. 9, no. 3, pp. 1793–1799, 2022, [Online]. Available: https://openlibrarypublications.telkomuniversity.ac.id/_index.php/engineering/article/view/17982
- [15] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *J. TEKNO KOMPAK*, vol. 15, no. 1, pp. 131–145, 2020.
- [16] M. Syarifuddin, "Analisis Sentimen Opini Publik Terhadap Efek Psbb Pada Twitter Dengan Algoritma Decision Tree,Knn, Dan Naive Bayes," *INTI Nusa Mandiri*, vol. 15, no. 1, pp. 87–94, 2020, doi: 10.33480/inti.v15i1.1433.
- [17] K. V. S. Toy, Y. A. Sari, and I. Cholissodin, "Analisis Sentimen Twitter menggunakan Metode Naive Bayes dengan Relevance Frequency Feature Selection (Studi Kasus: Opini Masyarakat mengenai Kebijakan New Normal)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 11, pp. 5068–5074, 2021, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [18] R. A. Raharjo, I. M. G. Sunarya, and D. G. H. Divayana, "Perbandingan Metode Naive Bayes Classifier Dan Support Vector Machine Pada Kasus Analisis Sentimen Terhadap Data Vaksin Covid-19 Di Twitter," *Elkom J. Elektron. dan Komput.*, vol. 15, no. 2, pp. 456–464, 2022, doi: 10.51903/elkom.v15i2.918.
- [19] R. N. Fahmi, N. Nursyifa, and A. Primajaya, "Analisis Sentimen Pengguna Twitter Terhadap Kasus Penembakan Laskar Fpi Oleh Polri Dengan Metode Naive Bayes Classifier," *JIKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 2, pp. 61–66, 2021, [Online]. Available: <https://ejournal.akakom.ac.id/index.php/jiko/article/view/437/0>
- [20] D. I. Sari, Y. F. Wati, and Widiastuti, "Analisis Sentimen Dan Klasifikasi Tweets Berbahasa Indonesia Terhadap Transportasi Umum Mrt Jakarta Menggunakan Naive Bayes Classifier," *J. Ilm. Inform. Komput.*, vol. 25, no. 1, pp. 64–75, 2020, doi: 10.35760/ik.2020.v25i1.2427.
- [21] 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 Cla," *e-Proceeding Eng.*, vol. 8, no. 5, pp. 9007–9016, 2021.
- [22] F. Wulandari, E. Haerani, M. Fikry, and E. Budianita, "Analisis sentimen larangan penggunaan obat sirup menggunakan algoritma naive bayes classifier," *J. CoSciTech (Computer Sci. Inf. Technol.)*, vol. 4, no. 1, pp. 88–96, 2023, doi: 10.37859/coscitech.v4i1.4781.
- [23] F. A. Wenando, R. Hayami, Soni, A. Fitria, and D. F. Shifana, "Sentimen analisis masyarakat terhadap kasus penembakan brigadir j menggunakan algoritma naive bayes classifier," *J. CoSciTech (Computer Sci. Inf. Technol.)*, vol. 4, no. 2, pp. 484–490, 2023, doi: 10.37859/coscitech.v4i2.5686.