

Analisis Sentimen Terhadap Topik Kenaikan Harga Bahan Bakar Minyak (BBM) pada Media Sosial Twitter

Tiara Danirmala¹, Yusuf Sulisty Nugroho²

l200190176@student.ums.ac.id, yusuf.nugroho@ums.ac.id

Universitas Muhammadiyah Surakarta

Informasi Artikel

Diterima : 8 May 2023

Direview : 20 May 2023

Disetujui : 27 Jun 2023

Kata Kunci

Analisis sentimen, kenaikan BBM, twitter, naïve bayes, topik diskusi

Abstrak

Media sosial *Twitter* menjadi salah satu media yang banyak digunakan masyarakat Indonesia untuk membagikan informasi, pendapat, ataupun sekedar berdiskusi tentang topik yang sedang tren dan beredar luas, misalnya tentang kenaikan harga bahan bakar minyak (BBM). Penelitian ini bertujuan untuk menganalisis sentimen masyarakat tentang kenaikan harga BBM di media sosial *Twitter*. Penelitian ini dilakukan dengan cara menganalisis sentimen *tweet* yang diunggah di *Twitter* untuk mengetahui polaritasnya dengan menerapkan klasifikasi Naïve Bayes. Selain analisis sentimen, data *tweet* juga dicari topik diskusi yang paling banyak dibahas terkait dengan kenaikan harga BBM dengan menggunakan metode *Latent Dirichlet Allocation (LDA)*. Hasil penelitian menunjukkan bahwa sebanyak 53.3% *tweet* memiliki polaritas positif, 31.2% *tweet* berpolaritas negatif, dan 15.5% *tweet* dinyatakan netral. Hasil evaluasi klasifikasi dengan *Naive Bayes* diperoleh nilai akurasi 60%. Sedangkan topik yang banyak dibicarakan masyarakat di *Twitter* secara umum menyatakan penolakannya terhadap kenaikan harga BBM. Hal ini menunjukkan bahwa masyarakat menolak kebijakan tersebut namun dinyatakan dengan komentar secara positif.

Keywords

Sentiment analysis, fuel price increase, twitter, naïve bayes, discussion topics

Abstrak

Social media Twitter is one of the media that is widely used by the Indonesian people to share information, opinions, or just to discuss trending and widely circulated topics, for example the increase in the price of fuel oil (BBM). This study aims to analyze public sentiment about the increase in fuel prices on social media Twitter. This research was conducted by analyzing the sentiments of tweets uploaded on Twitter to find out their polarity by applying the Naïve Bayes classification. In addition to sentiment analysis, tweet data is also searched for the most discussed topics of discussion related to rising fuel prices using the Latent Dirichlet Allocation (LDA) method. The results showed that 53.3% of tweets had positive polarity, 31.2% of tweets had negative polarity, and 15.5% of tweets were declared neutral. The results of the classification evaluation with Naive Bayes obtained an accuracy value of 60%. Meanwhile, the topics that are widely discussed by the public on Twitter generally express their rejection of the increase in fuel prices. This shows that the community rejects the policy but expresses it with positive comments.

A. Pendahuluan

Saat ini, hampir semua kegiatan membutuhkan dukungan fasilitas teknologi informasi, dimana teknologi ini dapat membantu manusia dalam membuat, mengubah, menyimpan, mengomunikasikan, dan menyebarkan informasi dalam berbagai aspek kehidupan sehari-hari dengan mudah dan cepat [1]. Salah satu implementasi teknologi informasi adalah penggunaan media sosial seperti *Twitter*, *Facebook*, *Instagram*, dan sebagainya. Media sosial, seperti *twitter*, memegang peranan penting dalam kegiatan penyaluran informasi dan komunikasi [2]. *Twitter* merupakan salah satu media komunikasi yang populer bagi masyarakat Indonesia, karena jumlah penggunanya yang terus meningkat serta seluruh informasi atau berita yang sedang trending dengan cepat beredar luas.

Di Indonesia, *twitter* menjadi salah satu *platform* media sosial yang aktif digunakan oleh masyarakat [3]. Indonesia menempati peringkat kelima sebagai negara pengguna *twitter* terbesar di dunia [4]. Berdasarkan laporan *We Are Social*, jumlah pengguna *twitter* di Indonesia pada tahun 2022 menunjukkan adanya kenaikan sebesar 31,3%. Dari jumlah pengguna tahun sebelumnya yaitu 14,05 juta menjadi 18,45 juta pada tahun 2022. Jumlah tersebut setara dengan 4,23% dari seluruh pengguna *twitter* di dunia yang berjumlah 436 juta.

Beberapa bulan terakhir, topik yang sedang hangat dan banyak dibicarakan di *twitter* oleh pengguna media sosial di Indonesia adalah tentang kenaikan harga bahan bakar minyak (BBM). Pemerintah resmi mengumumkan kenaikan harga BBM subsidi di Indonesia pada bulan September 2022, yaitu harga *pertalite* dari Rp 7.650 menjadi Rp 10.000 per liter, harga solar dari Rp 5.150 menjadi Rp 6.800 per liter, dan *pertamax* naik dari Rp 12.500 menjadi Rp 14.800 per liter. Kenaikan harga BBM tersebut telah memberikan dampak terhadap roda perekonomian masyarakat di Indonesia, sehingga banyak masyarakat yang mengeluh atas kebijakan tersebut. Padahal tingkat ekonomi masyarakat Indonesia masih belum sepenuhnya pulih dari pandemi, namun sudah dihantam dengan kenaikan harga BBM.

Menyikapi fenomena tersebut, masyarakat banyak yang memberikan berbagai opini terhadap kenaikan harga BBM melalui salah satu media sosial, yaitu *twitter*. Opini tersebut mendeskripsikan tanggapan yang berbeda dari setiap pengguna *twitter*, baik positif maupun negatif. Namun, sentimen dari tanggapan masyarakat Indonesia di *twitter* terhadap kenaikan BBM belum diketahui. Oleh karena itu, untuk mengetahui sentimen dari opini yang disampaikan di *twitter* perlu dilakukan analisis sentimen terkait topik kenaikan harga BBM serta untuk mengetahui topik utama yang dibicarakan terkait topik tersebut. Analisis sentimen memberikan gambaran secara umum tentang persepsi masyarakat dengan mengklasifikasikan opini yang muncul ke dalam kategori positif, negatif dan netral [5].

Media sosial *twitter* telah banyak digunakan dalam berbagai penelitian terdahulu. Misalnya penelitian tentang analisis sentimen terhadap topik viral pada media sosial *twitter* dengan kata kunci Desa Penari [6]. Pada penelitian tersebut, dari 1000 data *tweet* didapatkan *tweet* bernilai positif sejumlah 33 data (3.3%), *tweet* bernilai netral sejumlah 767 data (76.7%) dan *tweet* negatif sejumlah 200 data (20%). Penelitian serupa juga pernah dilakukan untuk menganalisis sentimen masyarakat Indonesia terhadap wacana vaksin *covid-19* pada media sosial *twitter* [7]. Hasil penelitian dengan metode *Latent Dirichlet Allocation* (LDA) menunjukkan bahwa respon positif masyarakat sebesar 30%, sedangkan respon negatif masyarakat sebesar 26%. Pembahasan masyarakat tentang perdebatan vaksin yang dianggap terburu-buru, sertifikasi halal vaksin dan keraguan

masyarakat terhadap kualitas vaksin muncul sebagai topik utama terkait wacana vaksinasi tersebut.

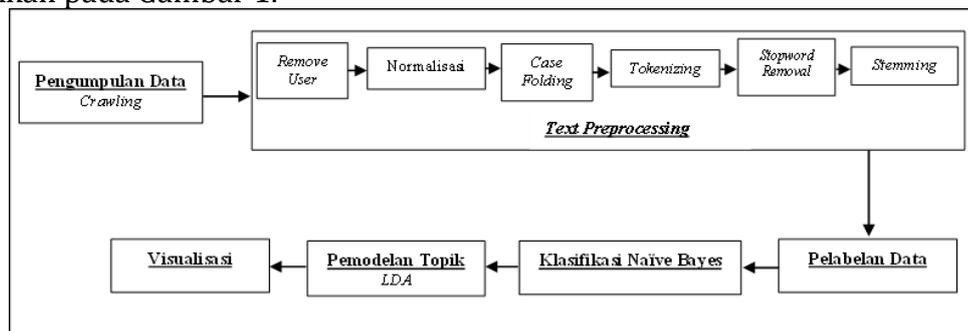
Berdasarkan latar belakang masalah dan kajian beberapa penelitian terdahulu, penelitian ini dilakukan untuk menganalisis sentimen masyarakat terhadap topik kenaikan harga BBM pada media sosial *twitter* serta topik utama yang dibicarakan oleh masyarakat terkait fenomena tersebut dengan metode *Latent Dirichlet Allocation* (LDA). *Latent Dirichlet Allocation* (LDA) merupakan metode *unsupervised learning* yang berfungsi meringkas data, mengelompokkan data ke dalam beberapa kelas, dan memproses data dalam jumlah besar [8]. Penelitian dilakukan menggunakan dataset dari *tweet* berbahasa Indonesia yang diunggah pada media sosial *twitter* dan diekstraksi menggunakan kata kunci “BBMNaik” dan “KenaikanBBM” yang diunggah oleh pengguna mulai tanggal 04 September 2022 sampai dengan tanggal 24 September 2022. Data tersebut diolah dan diklasifikasikan berdasarkan sentimen positif, netral, dan negatif, sehingga dapat diperoleh kesimpulan berdasarkan analisis sentimen tersebut.

Hasil penelitian menunjukkan bahwa sebanyak 53.3% *tweet* dikategorikan memiliki polaritas positif, 31.2% *tweet* berpolaritas negatif, dan 15.5% *tweet* dinyatakan netral. Hasil evaluasi klasifikasi dengan *Naive Bayes* diperoleh nilai akurasi 60%. Sedangkan topik yang paling banyak dibicarakan adalah tentang “perintah”. Dari hasil analisis sentimennya dapat disimpulkan bahwa pada umumnya masyarakat menentang kebijakan pemerintah menaikkan harga BBM, tetapi masyarakat bersikap dan memberikan komentar pada *Twitter* atas kebijakan tersebut secara positif.

B. Metode Penelitian

Pada penelitian ini dilakukan analisis sentimen terhadap *tweet* pengguna tentang kenaikan harga bahan bakar minyak (BBM) di media sosial *twitter*. Analisis sentimen dilakukan untuk menentukan persepsi atau subjektivitas pengguna terhadap topik kenaikan BBM dan permasalahannya [7]. Hal tersebut dapat membantu untuk memahami tanggapan masyarakat terhadap topik kenaikan BBM dan membantu dalam menentukan apakah fenomena tersebut positif, negatif, atau netral [9].

Penelitian ini dilakukan dengan melalui 6 tahap utama, seperti yang ditunjukkan pada Gambar 1.



Gambar 1. Enam tahap utama yang dilakukan dalam penelitian ini

Tahap-tahap penelitian yang ditunjukkan dalam Gambar 1 dapat dijelaskan sebagai berikut:

1. Pengumpulan Data

Pada tahap pengumpulan data terdapat proses *crawling* yaitu proses dimana dilakukan pengambilan data dari media sosial *Twitter* untuk digunakan dalam

penelitian. Data diperoleh dengan melakukan pengambilan data *tweet* masyarakat pada platform *Twitter* dengan memanfaatkan modul *snsrape*. Data yang digunakan merupakan komentar-komentar atau opini masyarakat pada *Twitter* tentang kebijakan kenaikan harga BBM dengan kata kunci “BBMNaik” dan “KenaikanBBM”, yang diunggah oleh pengguna mulai tanggal 04 September 2022 sampai dengan tanggal 24 September 2022. Data yang diambil memiliki atribut *datetime*, *tweet id*, *username*, dan *text*. *Datetime* merupakan data yang berisi tanggal dan waktu *tweet* diunggah. *Tweet id* merupakan kode unik yang dimiliki setiap akun di *twitter*, dan *username* merupakan data yang berisi username dari akun yang mengunggah *tweet* pada *twitter*. Sedangkan *text* merupakan data yang berisi *tweet* atau komentar yang diunggah oleh akun pada *twitter*. Data yang digunakan untuk klasifikasi sentimen adalah data dari atribut *text*.

2. Text Preprocessing

Pada tahap *preprocessing* terdapat beberapa proses dengan tujuan agar data yang diperoleh lebih terstruktur. Adapun beberapa tahap *preprocessing* diantaranya, *remove user*, normalisasi, *case folding*, *tokenizing*, *stopword removal*, serta *stemming*. Pada proses pertama adalah *remove user* yang dilakukan untuk menghapus nama user (@pengguna) dari *tweet* [10]. Normalisasi merupakan proses cleaning atau pembersihan data *tweet* dari karakter dan tanda baca yang tidak diperlukan, seperti URL, tag (#), tanda baca seperti titik (.), Koma (,) dan tanda baca lainnya [11]. *Case folding* adalah proses konversi huruf menjadi huruf kecil, menghapus kata berhenti dan semua karakter non-alfabet [12]. *Tokenizing* yaitu tahap penguraian atau memisahkan kata-kata dalam sebuah paragraf, kalimat atau halaman ke dalam kata tunggal [13]. *Stopword removal*, teknik yang bertujuan untuk menghapus kata didalam dokumen dataset yang dianggap tidak memberikan deskripsi sesuatu dalam bahasa Indonesia [14]. Sedangkan *stemming* merupakan proses yang dilakukan pada tahap terakhir untuk mengubah kata berimbuhan ke dalam bentuk kata dasar [15].

3. Pelabelan Data

Tahap *labeling* merupakan proses pemberian label pada data yang sudah diambil atau didapatkan dari proses *crawling* [3]. Data *tweet* diseleksi dan diberi label secara otomatis menggunakan *library vader lexicon* sehingga data terklasifikasi dan terkategori sesuai dengan label yang diberikan. Ada 3 jenis label yang digunakan untuk mengklasifikasi sentimen pengguna yaitu positif, negatif dan netral.

4. Klasifikasi Naïve Bayes

Tahap ini mengklasifikasi data hasil ekstraksi data dengan menerapkan algoritma *Naïve Bayes*. *Naïve Bayes* merupakan teknik klasifikasi berbasis probabilistik sederhana yang berdasarkan teorema bayes dengan asumsi independensi yang kuat [1]. Tujuan dari pengujian menggunakan algoritma *Naïve Bayes* yaitu mengklasifikasi data sebagai sentimen positif atau negatif secara otomatis [16].

5. Pemodelan Topik

Tahap pemodelan topik dilakukan dengan metode *Latent Dirichlet Allocation* (LDA). LDA digunakan untuk meringkas, mengelompokkan, mengolah data yang besar dan menghasilkan daftar topik yang berbobot untuk setiap dokumen [7].

Metode LDA dapat menangkap topik utama yang dibicarakan masyarakat terkait topik kenaikan harga BBM pada media sosial *Twitter*.

6. Visualisasi

Hasil analisis yang diperoleh kemudian dilakukan visualisasi dengan menggunakan *pie-chart* dan *word cloud*.

C. Hasil dan Pembahasan

1. Pengumpulan Data

Pengumpulan data *Twitter* dilakukan dengan menggunakan modul *snsrape*. Data *tweet* yang diambil adalah *tweet* yang diunggah oleh pengguna pada tanggal 04 September 2022 sampai dengan 24 September 2022 dengan kata kunci “BBMNaik” dan “KenaikanBBM”. Ekstraksi data dapat menghasilkan total sebanyak 5.015 *tweet* dalam Bahasa Indonesia, seperti yang ditunjukkan dalam Tabel 1.

Tabel 1. Data *tweet* pengguna terkait dengan kenaikan harga BBM

Hashtag	Jumlah data mentah <i>tweet</i>	Jumlah <i>tweet</i> setelah menghilangkan duplikasi
#BBMNaik	2.237 (44.6%)	2.133 (48.3%)
#KenaikanBBM	2.778 (55.4%)	2.286 (51.7%)
Total Tweet	5.015 (100 %)	4.419 00%)

2. Text Preprocessing

Proses selanjutnya yaitu *text preprocessing* data terhadap 4.419 *tweet* setelah menghilangkan duplikasi. Tujuan dari tahap ini agar data lebih terstruktur dan bersih sebelum melakukan tahap proses selanjutnya. Tahap ini dilakukan secara terpisah dan dilakukan secara berurutan mulai dari *remove user*, normalisasi, *case folding*, *tokenizing*, *stopword removal*, dan terakhir *stemming*.

- Remove User* merupakan tahap yang digunakan untuk menghapus *username* pada data *tweet*, seperti yang dicontohkan pada Tabel 2 yaitu menghapus user pengguna @kontanNews.

Tabel 2. Contoh *Remove user* pada salah satu *tweet*

<i>Tweet</i>	<i>Remove User</i>
b'Sejumlah Perusahaan Otobus (PO) terpantau mulai mengerek harga jual tiketnya sebagai imbas dari kenaikan harga BBM @kontanNews https://t.co/SB4ll0vZSb'	b'Sejumlah Perusahaan Otobus (PO) terpantau mulai mengerek harga jual tiketnya sebagai imbas dari kenaikan harga BBM https://t.co/SB4ll0vZSb'

- Normalisasi atau *cleaning* bertujuan untuk membersihkan fitur-fitur yang tidak diperlukan seperti URL, tag, maupun tanda baca. Contoh normalisasi ditunjukkan pada Tabel 3 yaitu membersihkan URL https://t.co/SB4ll0vZSb.

Tabel 3. Contoh normalisasi terhadap *tweet* hasil dari tahap *remove user*

<i>Remove User</i>	Normalisasi
b'Sejumlah Perusahaan Otobus (PO)	Sejumlah Perusahaan Otobus

terpantau mulai mengerek harga jual tiketnya sebagai imbas dari kenaikan harga BBM @kontanNews https://t.co/SB4ll0vZSb	terpantau mulai mengerek harga jual tiketnya sebagai imbas dari kenaikan harga BBM
--	--

- c. *Case Folding* dilakukan untuk mengubah variabel teks menjadi karakter, mengubah huruf besar menjadi huruf kecil [6], seperti yang dicontohkan pada Tabel 4 yaitu mengubah huruf kapital pada kata “Sejumlah”, “Perusahaan”, “Otobus”, “BBM”, menjadi “sejumlah”, “perusahaan”, “otobus”, “bbm”.

Tabel 4. Contoh *Case Folding* pada *tweet* hasil tahap normalisasi

Normalisasi	Case Folding
Sejumlah Perusahaan Otobus terpantau mulai mengerek harga jual tiketnya sebagai imbas dari kenaikan harga BBM	sejumlah perusahaan otobus terpantau mulai mengerek harga jual tiketnya sebagai imbas dari kenaikan harga bbm

- d. *Tokenizing* dilakukan untuk pemisahan kata sehingga menjadi kata yang berdiri sendiri atau satu kata [17], seperti yang dicontohkan pada Tabel 5 yaitu membagi kalimat pada kolom *case folding* menjadi 15 token atau kata tersendiri.

Tabel 5. Contoh penerapan *tokenizing* pada *tweet* hasil tahap *case folding*

Case Folding	Tokenizing
sejumlah perusahaan otobus terpantau mulai mengerek harga jual tiketnya sebagai imbas dari kenaikan harga bbm	['sejumlah', 'perusahaan', 'otobus', 'terpantau', 'mulai', 'mengerek', 'harga', 'jual', 'tiketnya', 'sebagai', 'imbasi', 'dari', 'kenaikan', 'harga', 'bbm']

- e. *Stopword Removal*, dimana proses tersebut membersihkan dokumen dari kata yang tidak diperlukan untuk mengurangi noise [1], seperti yang dicontohkan pada Tabel 6 yaitu menghapus kata “sejumlah”, “mulai”, “sebagai”, “dari”.

Tabel 6. Contoh proses *Stopword removal* pada *tweet* hasil penerapan *tokenizing*

Tokenizing	Stopword Removal
['sejumlah', 'perusahaan', 'otobus', 'terpantau', 'mulai', 'mengerek', 'harga', 'jual', 'tiketnya', 'sebagai', 'imbasi', 'dari', 'kenaikan', 'harga', 'bbm']	['perusahaan', 'otobus', 'terpantau', 'mengerek', 'harga', 'jual', 'tiketnya', 'imbasi', 'kenaikan', 'harga', 'bbm']

- f. *Stemming*, proses pengembalian berbagai bentukan kata menjadi bentukan kata dasar dengan menghilangkan imbuhan [18], seperti yang dicontohkan pada Tabel 7 yaitu mengubah kata “perusahaan” menjadi “usaha”, “terpantau” menjadi “pantau”, “mengerek” menjadi “kerek”, “tiketnya” menjadi “tiket”, “kenaikan” menjadi “naik”.

Tabel 7. Contoh proses *Stemming* pada tweet hasil dari proses *stopword removal*

Stopword Removal	Stemming
['perusahaan', 'otobus', 'terpantau', 'mengerek', 'harga', 'jual', 'tiketnya', 'imbas', 'kenaikan', 'harga', 'bbm']	usaha otobus pantau kerek harga jual tiket imbas naik harga bbm

3. Pelabelan Data

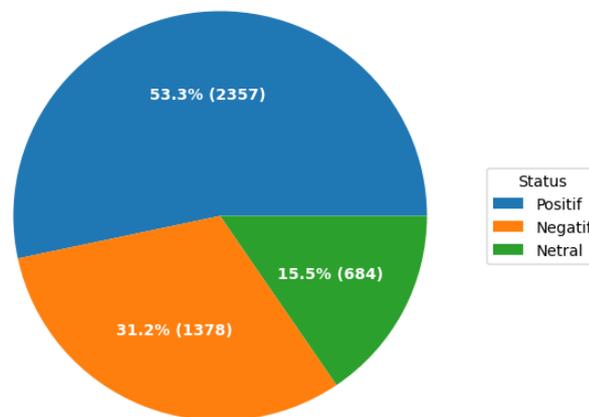
Tahap selanjutnya dalam penelitian ini adalah melakukan pelabelan data secara otomatis pada atribut *text* yang digunakan dalam klasifikasi data. Data yang digunakan pada tahap ini adalah 4,419 data *tweet* yang diperoleh setelah menghilangkan duplikasi. Data dalam atribut *text* diberi label secara otomatis menggunakan *library vader lexicon*. Label yang digunakan terdapat 3 jenis label yaitu positif, negatif, dan netral. Label positif menunjukkan bahwa *tweet* yang diunggah menggambarkan perasaan positif pengunggah dan seringkali menyertakan kata-kata positif, sedangkan label negatif menunjukkan bahwa *tweet* tersebut memiliki sentimen negatif saat membaca. Adapun *tweet* dengan label netral menggambarkan bahwa *tweet* tersebut tidak menunjukkan sentimen positif maupun negatif [19]. Tabel 8 menampilkan contoh *tweet* yang memiliki sentimen positif, negatif, maupun netral.

Tabel 8. Contoh sentimen pada *tweet* yang diunggah di media *Twitter*

Tweet	Text Preprocessing	Label
b'Ditopang indikator ekonomi kenaikan harga BBM dinilai jadi momentum tepat\n\nMereka bilang impian saya terlalu besar, saya bilang mereka berfikir terlalu kecil #DemiWujudkanRIMaju https://t.co/R6FpNTCyFa'	topang indikator ekonomi naik harga bbm nilai jadi momentum tepat mereka bilang impian saya terlalu besar saya bilang mereka berfikir terlalu kecil demiwujudkanrimaju	positif
b'@geloraco Otak taruh di wc,jika pendapatan menurun akibat kenaikan bbm.makan biaya perawatan rutin semakin dipangkas oleh pemilik angkot,lama2 kelamaan bangkrut.angkot ganting sopirnya telangtung.otak jangan disimpan'	otak taruh dapat turun akibat naik bbm makan biaya perawatan rutin pangkas milik angkot bangkrut angkot ganting sopir telangtung otak simpan	negatif

b'Pengusaha Pemerintah Dampak Harga\xc2\xa0BBM https://t.co/HOCfDpLmJJ '	Minta Kendalikan Kenaikan	pengusaha pemerintah naik harga bbm	minta kendali dampak	netral
--	---------------------------------	---	-------------------------	--------

Gambar 2 menjelaskan hasil pelabelan secara otomatis terhadap dataset yang menunjukkan bahwa sebanyak 53,3% data memiliki sentimen positif. Adapun respon negatif pengguna ditunjukkan sebanyak 31,2% dan 15,5% sisanya menunjukkan respon netral. Hasil analisis sentimen ini menunjukkan bahwa secara umum pengguna *twitter* memberikan respon secara positif terhadap kebijakan pemerintah tentang kenaikan harga BBM.



Gambar 2. Hasil proses *labeling* data twitter topik kenaikan harga BBM

4. Klasifikasi Naive Bayes

Tahap berikutnya adalah melakukan klasifikasi data dengan menerapkan algoritma *Naive Bayes*. Algoritma Naive Bayes merupakan pengklasifikasi probabilitas sederhana yang menghitung sekumpulan probabilitas dengan menghitung frekuensi dan kombinasi nilai dalam kumpulan data tertentu [20]. Tahap ini bertujuan untuk menentukan nilai prediksi dan mencari nilai akurasi. Data klasifikasi dibagi menjadi dua bagian yaitu data *training* dan data *testing* [21]. Data *training* ditentukan sebanyak 80% dan data *testing* sebanyak 20% dari jumlah keseluruhan 4.419 data *tweet*.

Tabel 9. Hasil klasifikasi Naive Bayes dengan data training dan data testing

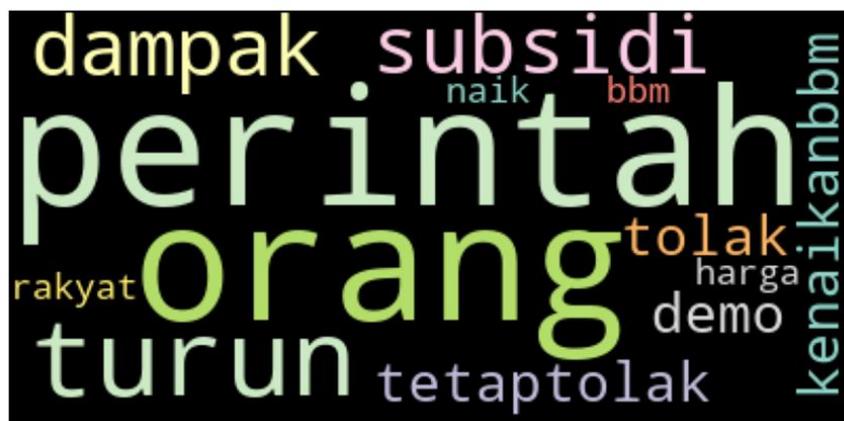
	Precision	Recall	F1-Score
-1	0.61	0.48	0.54
0	0.42	0.20	0.27
1	0.62	0.80	0.70
Accuracy			0.60
Macro avg	0.55	0.50	0.50
Weghted avg	0.58	0.60	0.58

Hasil klasifikasi dengan menerapkan *Naive Bayes*, seperti yang dijelaskan pada Tabel 9, diperoleh nilai akurasi 60%. Nilai *precision* pada kelas positif adalah 62%, kelas negatif 61%, dan nilai *precision* kelas netral 42%. Dari nilai akurasi tersebut menunjukkan bahwa performa klasifikasi ini Hal ini diartikan untuk nilai *precision* ini merupakan nilai akurasi yang benar antara data yang diminta dengan data yang diprediksi, terlihat untuk *precision* kelas positif menghasilkan nilai akurasi yang lebih tinggi dibandingkan dengan kelas negatif dan netral. Sedangkan untuk *recall* atau keberhasilan model dalam menentukan kembali sebuah informasi, memiliki nilai akurasi yang tinggi pada kelas positif yaitu 80%, dibandingkan dengan kelas negatif 48% dan kelas netral 20%.

5. Pemodelan Topik dan Visualisasi

Pada tahap ini data diolah kembali untuk melakukan pemodelan topik dengan metode *Latent Dirichlet Allocation (LDA)*. Metode ini diterapkan untuk mendeteksi topik secara otomatis pada data *tweet* dan mengelompokannya sesuai dengan topik yang sedang dibicarakan. Sebelum melakukan pemodelan, data diekstraksi (pembobotan kata) dengan metode *TF-IDF*. Setelah itu, dilakukan pemodelan topik dengan metode *Latent Dirichlet Allocation (LDA)*.

Hasil pemodelan divisualisasikan dengan menggunakan *Word Cloud*, seperti yang ditunjukkan pada Gambar 4 yang menunjukkan 13 topik yang paling banyak dibicarakan dalam *tweet* dengan kata kunci “BBMNaik” dan “KenaikanBBM”. Topik tersebut adalah “perintah” sebagai topik yang paling banyak dibicarakan dalam *Twitter*. Sedangkan topik lainnya yaitu “orang”, “turun”, “dampak”, “subsidi”, “tetaptolak”, “tolak”, “demo”, “kenaikanbbm”, “rakyat”, “harga”, “bbm”, dan “naik”, dimana topik-topik ini menunjukkan bahwa pada dasarnya masyarakat menolak kenaikan harga BBM.



Gambar 3. Topik data *tweet* tentang kenaikan harga BBM pada twitter

D. Simpulan

Penelitian tentang analisis kenaikan harga bahan bakar minyak (BBM) pada media sosial *twitter* menunjukkan bahwa data yang digunakan sebanyak 4.419 *tweet* dalam bahasa Indonesia dari tanggal 04 September 2022 sampai dengan 24 September 2022 memiliki sentimen positif sebanyak 53.3% (2.357 data), negatif 31.2% (1.378 data), dan netral 15.5% (684 data). Sedangkan pemodelan topik dengan LDA menunjukkan secara umum topik diskusi yang dibicarakan oleh

masyarakat berisi tentang penolakan kenaikan harga BBM. Hasil penelitian ini dapat disimpulkan bahwa meskipun secara umum masyarakat menolak kebijakan pemerintah tentang kenaikan harga BBM, namun masyarakat menyatakan penolakannya dalam *Twitter* secara positif jika dilihat dari hasil analisis sentimennya.

E. Ucapan Terima Kasih

Ucapan terima kasih keada Universitas Muhammadiyah Surakarta yang telah memberikan dukungan terhadap penelitian ini.

F. Referensi

- [1] A. Kusuma and A. Nugroho, "Analisa Sentimen Pada Twitter Terhadap Kenaikan Tarif Dasar Listrik Dengan Metode Naïve Bayes," *J. Ilm. Teknol. Inf. Asia*, vol. 15, no. 2, pp. 137–146, Dec. 2021, doi: 10.32815/JITIKA.V15I2.557.
- [2] A. F. Hidayatullah, "Twitter Sebagai Media Dakwah. Teknoin, 22(1). <https://doi.org/10.20885/teknoin.vol22.iss1.art5>Twitter Sebagai Media Dakwah," *Teknoin*, vol. 22, no. 1, 2016, doi: 10.20885/teknoin.vol22.iss1.art5.
- [3] A. L. Fairuz, R. D. Ramadhani, and N. A. F. Tanjung, "Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 1, no. 1, pp. 42–51, Feb. 2021, doi: 10.20895/DINDA.V1I1.180.
- [4] D. Ngakan, G. Wahyu, M. Putra, R. P. Ridzkyanto, and G. Nurika, "Positive and Negative Sentiments Regarding BPJS-Kesehatan Premium Increase: An Analysis Using Text Data Mining on Twitter," vol. 5, pp. 41–47, 2020.
- [5] R. Cahyani, I. S. Rozas, and N. Yalina, "Analisis Sentimen Pada Media Sosial Twitter Terhadap Tokoh Publik Peserta Pilpres 2019," *MATICS J. Ilmu Komput. dan Teknol. Inf. (Journal Comput. Sci. Inf. Technol.)*, vol. 12, no. 1, pp. 79–86, Apr. 2020, doi: 10.18860/MAT.V12I1.8356.
- [6] R. Arief and K. Imanuel, "ANALISIS SENTIMEN TOPIK VIRAL DESA PENARI PADA MEDIA SOSIAL TWITTER DENGAN METODE LEXICON BASED," *J. Ilm. Matrik*, vol. 21, no. 3, pp. 242–250, Dec. 2019, doi: 10.33557/JURNALMARIK.V21I3.727.
- [7] F. F. Rachman and S. Pramana, "Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter," *Indones. Heal. Inf. Manag. J.*, vol. 8, no. 2, pp. 100–109, Dec. 2020, doi: 10.47007/INOHIM.V8I2.223.
- [8] M. A. N. Febriansyach, F. Rashif, G. I. P. Nirvana, and N. A. Rakhmawati, "Implementasi LDA untuk Pengelompokan Topik Tweet Akun Bot Twitter bertagar #covid-19," *CogITo Smart J.*, vol. 7, no. 1, pp. 170–181, Jun. 2021, doi: 10.31154/COGITO.V7I1.299.170-181.
- [9] R. Patel and K. Passi, "Sentiment Analysis on Twitter Data of World Cup Soccer Tournament Using Machine Learning," *IoT 2020, Vol. 1, Pages 218-239*, vol. 1, no. 2, pp. 218–239, Oct. 2020, doi: 10.3390/IOT1020014.
- [10] S. Yang, "Text Mining of Twitter Data Using a Latent Dirichlet Allocation Topic Model and Sentiment Analysis," *Int. J. Comput. Inf. Eng.*, vol. 12, no. 7, pp. 525–529, 2018.

- [11] W. Yulita *et al.*, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier," *J. Data Min. dan Sist. Inf.*, vol. 2, no. 2, pp. 1–9, Aug. 2021, doi: 10.33365/JDMSI.V2I2.1344.
- [12] H. M. Keerthi Kumar, S. Likhitha, and B. S. Harish, "A Detailed Survey on Topic Modeling for Document and Short Text Data," *Artic. Int. J. Comput. Appl.*, vol. 178, no. 39, pp. 975–8887, 2019, doi: 10.5120/ijca2019919265.
- [13] H. Najjichah, A. Syukur, and H. Subagyo, "PENGARUH TEXT PREPROCESSING DAN KOMBINASINYA PADA PERINGKAS DOKUMEN OTOMATIS TEKS BERBAHASA INDONESIA," *J. Teknol. Inf.*, vol. 15, no. 1, 2019, Accessed: May 04, 2023. [Online]. Available: <http://research>.
- [14] MohYasid, L. Junaedi, and A. Sentimen Maskapai Citilink Naïve Bayes, "ANALISIS SENTIMEN MASKAPAI CITILINK PADA TWITTER DENGAN METODE NAÏVE BAYES," *J. Ilm. Inform.*, vol. 7, no. 02, pp. 82–86, Oct. 2019, doi: 10.33884/JIF.V7I02.1329.
- [15] A. S. Aribowo, H. Basiron, N. S. Herman, and S. Khomsah, "An evaluation of preprocessing steps and tree-based ensemble machine learning for analysing sentiment on Indonesian youtube comments," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 5, pp. 7078–7086, Oct. 2020, doi: 10.30534/IJATCSE/2020/29952020.
- [16] R. Fajar, S. Program, P. Rekayasa, N. Lunak, and R. Bengkalis, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, pp. 50–59, Jun. 2018, doi: 10.35314/ISI.V3I1.335.
- [17] M. A. Abdullah and D. Mahdiana, "Analisis Sentimen terhadap Peluang Kerja di Indonesia selama Masa Pandemi COVID-19 dengan Metode Klasifikasi Naive Bayes," *J. MEDIA Inform. BUDIDARMA*, vol. 6, no. 2, pp. 1234–1237, Apr. 2022, doi: 10.30865/MIB.V6I2.3972.
- [18] K. Kelvin, J. Banjarnahor, E. I. -, M. N. Nababan, and S. H. Sinurat, "Analisis perbandingan sentimen Corona Virus Disease-2019 (Covid19) pada Twitter Menggunakan Metode Logistic Regression Dan Support Vector Machine (SVM)," *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 5, no. 2, pp. 47–52, Feb. 2022, doi: 10.34012/JURNALSISTEMINFORMASIDANILMUKOMPUTER.V5I2.2365.
- [19] Y. S. Nugroho *et al.*, "How are project-specific forums utilized? A study of participation, content, and sentiment in the Eclipse ecosystem," *Empir. Softw. Eng.*, vol. 26, no. 6, 2021, doi: 10.1007/s10664-021-10032-2.
- [20] M. M. Saritas and A. Yasar, "Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification," *Int. J. Intell. Syst. Appl. Eng.*, vol. 7, no. 2, pp. 88–91, Jun. 2019, doi: 10.18201//ijisae.2019252786.
- [21] V. A. Kharde and S. S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques," *Int. J. Comput. Appl.*, vol. 139, no. 11, pp. 975–8887, 2016, Accessed: Nov. 30, 2022. [Online]. Available: <http://ai.stanford>.