

Sentimen Kebijakan Pemerintah terhadap Kepercayaan Masyarakat Menggunakan Metode NLP Rule Based

Evasaria Magdalena Sipayung^{a1}, Bhustomy Hakim^{b2}

^aProgram Studi Informatika, Fakultas Teknologi dan Desain, Universitas Bunda Mulia, Tangerang, Indonesia

Jl. Jalur Sutera Barat kav 7-9, Alam Sutera, Tangerang, Banten 15143

¹evasaria.sipayung@gmail.com

²bhakim@bundamulia.ac.id

Abstrak

Kebijakan pemerintah dipahami dari saran yang akan dicapai atau diatur di mana sasarannya adalah publik dalam pengertian masyarakat dan kepentingan umum maka kebijakan pemerintah dapat dikategorikan sebagai kebijakan publik. Kebijakan yang diterapkan Pemerintah pada 2022, diantaranya adalah tahapan penghapusan bahan bakar minyak jenis Premium dan Pertalite, kenaikan harga gas elpiji nonsubsidi, penerapan kelas standar BPJS, dan penangkapan ikan dibatasi melalui sistem kuota. Pada penelitian ini analisis sentimen kebijakan pemerintah dilakukan dengan menggunakan metode NLP (Natural Language Processing) pada data tweet. Terdapat tiga kebijakan yang menjadi acuan kepercayaan masyarakat yaitu kebijakan "Kenaikan Pertalite", "Kelas BPJS", dan "LPG Non-Subsidi". Sentimen paling negatif didapat saat kebijakan "Kenaikan Pertalite" dimana terdapat 73,4% tweet bermakna negatif yang mengkritisi pemerintah. Kebijakan "Kelas BPJS" juga mendapat sentimen negatif dengan presentasi tweet negatif sebesar 43,4% yang mengeluhkan kebijakan baru kenaikan harga kelas BPJS dan proses alur BPJS. Namun, untuk kebijakan "LPG Non-Subsidi", hasil analisis sentimennya menunjukkan bahwa 52,9% tweet menyetujui kebijakan tersebut bahkan dengan kata-kata asosiasi yang sangat baik. Fenomena sosial media yang disebut "buzzer" karena 35% dari tweet positif tersebut berasal dari akun yang dipertanyakan kredibilitasnya. Model analisis sentimen sudah menunjukkan bahwa model tersebut sangatlah baik dengan akurasi 86,3% yang dapat merepresentasikan bahwa kejadian yang digambarkan pada data sosial media twitter tersebut sudah sesuai dengan kenyataan yang terjadi.

Kata kunci: kebijakan pemerintah, sentimen analisis, NLP, rule based, komentar

Sentiment on Public Trust Using the NLP Rule Based Method

Abstract

Government policy is understood from suggestions that will be achieved or regulated where the target is the public in the sense of society and the public interest, then government policy can be categorized as public policy. The policies implemented by the Government in 2022 include the stages of eliminating Premium and Pertalite fuel oil, increasing the price of non-subsidized LPG gas, the application of BPJS standard classes, and fishing is limited through a quota system. In this research, government policy sentiment analysis was carried out using the NLP (Natural Language Processing) method on tweet data. There are three policies that serve as a reference for public trust, namely the "Pertalite Increase", "BPJS Class" and "Non-Subsidized LPG" policies. The most negative sentiment was obtained during the "Increase in Pertalite" policy where there were 73.4% of negative tweets criticizing the government. The "BPJS Class" policy also received negative sentiment with the presentation of negative tweets of 43.4% complaining about the new policy of increasing BPJS class prices and the BPJS flow process. However, for the "Non-Subsidized LPG" policy, the results of the sentiment analysis showed that 52.9% of tweets agreed with the policy even with very good association words. The social media phenomenon is called a "buzzer" because 35% of positive tweets come from accounts whose credibility is questionable. The sentiment analysis model has shown that the model is very good with an accuracy of 86.3% which can represent that the events depicted in the Twitter social media data are in accordance with the reality that occurred.

Keywords: government policy, sentiment analysis, NLP, rule based, comment

I. PENDAHULUAN

Kebijakan publik sebagai bagian upaya atau tindakan pemerintah berupa pengaturan/keputusan yang dibuat dalam rangka menyelenggarakan tugas-tugas

pemerintahannya. Kebijakan yang diambil pemerintah dapat menimbulkan berbagai respon dan komentar dari masyarakat namun Pemerintah belum menyediakan wadah kepada masyarakat untuk menyampaikan komentar

berkaitan dengan kebijakan pemerintah tersebut. Hal ini menyebabkan masyarakat menyampaikan komentarnya lewat media sosial. Salah satu media sosial yang dapat digunakan untuk memberikan komentar adalah twitter. Kebijakan yang diterapkan Pemerintah pada 2022 di antaranya penghapusan bahan bakar minyak jenis Premium dan Pertalite, dan kenaikan harga gas elpiji nonsubsidi mendapat tanggapan dari masyarakat dalam bentuk komentar yang diberikan melalui media sosial twitter [1]. Respon publik terhadap suatu isu terkini dapat diketahui dari media sosial seperti twitter sebagai sumber informasi dan sumber data alternatif [2]. Salah satu cara untuk mengetahui tanggapan masyarakat dengan cara merangkum opini di media sosial memperlihatkan bahwa metode MNB menggunakan fitur 3-gram memberikan hasil terbaik [3]. Twitter merupakan platform media sosial umum digunakan user untuk berkomunikasi dan menyebarkan informasi berupa tweets. Tweets dapat dijadikan sumber data penting untuk melakukan penelitian Neuro-Linguistic Programming (NLP) seperti analisa sentimen, deteksi polaritas dan prediksi emoji [4]. Semua data yang telah melalui preprocess dan pelabelan kemudian akan dimodelkan menggunakan algoritma *Long Short Term Memory* (LSTM). Model yang terbentuk dapat digunakan untuk mengklasifikasikan sentimen tweet secara otomatis [5].

Oleh karena itu perlu dilakukan analisis sentimen komentar masyarakat melalui media sosial twitter untuk dapat mengetahui kepercayaan masyarakat terhadap kebijakan pemerintah. Sentiment Analysis akan diterapkan dalam percobaan ini untuk mengekstrak pengetahuan tentang bagaimana ulasan positif atau negatif yang terlihat seperti dalam format teks. Ini adalah penambangan data (data mining) kontekstual teks sebagai pengidentifikasian informasi subjektif untuk memahami sentimen orang dari sesuatu sehingga mesin dapat mengklasifikasikan pesan masa depan dan memberikan pernyataan apakah sentimen yang mendasari diberi label (positif atau negatif) [6]. Aplikasi dasar analisis sentimen terletak pada pengumpulan pendapat orang. Pendapat semacam itu adalah pendahulu dari banyak keputusan bisnis. Mirip dengan stop-kata, domain analisis sentimen tergantung pada daftar kata-kata yang menggambarkan pengaruh penulis [7]. Sentimen masyarakat terhadap rencana penerapan PSBB di Kota Ambon berdasarkan komentar di media sosial cukup berimbang antara sentimen positif dan negatif serta didominasi komentar dengan sentimen netral [8]. NBC menyelesaikan Opinion mining dengan Part-of-Speech (POS) Tagging yang digunakan untuk memberikan kelas kata (tag) secara gramatikal ke setiap kata dalam suatu kalimat teks. Hasil proses tersebut kemudian diterapkan rule untuk mengetahui apakah suatu dokumen termasuk opini atau bukan. Kemudian opini diklasifikasikan ke dalam opini positif dan negatif (subproses Opinion orientation) [9].

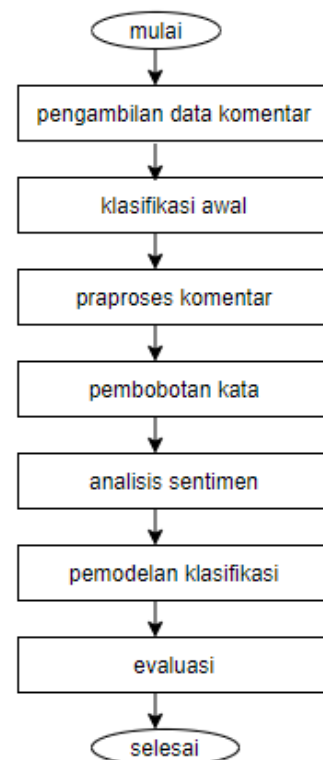
Analisis sentimen kebijakan pemerintah yang dilakukan dalam penelitian ini menggunakan metode NLP (Natural Language Processing) Rule Based. Metode NLP digunakan untuk memberikan kelas kata secara gramatikal ke setiap kata dalam suatu kalimat teks. Hasil proses

tersebut kemudian diterapkan rule untuk mengetahui apakah suatu komentar diklasifikasikan ke dalam positif dan negatif [10].

II. METODE

Penelitian ini mengklasifikasikan data menjadi dua kelas yaitu sentimen positif dan negatif. Data tweet yang digunakan adalah tweet dengan kata kunci dari kebijakan pemerintah yang ada.

Penelitian ini dilakukan melalui beberapa tahap terdiri dari: pengambilan data, klasifikasi awal, praproses data, analisis sentimen, pembobotan kata, analisis sentimen, pemodelan klasifikasi, dan evaluasi. Tahap yang dilakukan pada penelitian ditunjukkan pada Gambar 1.

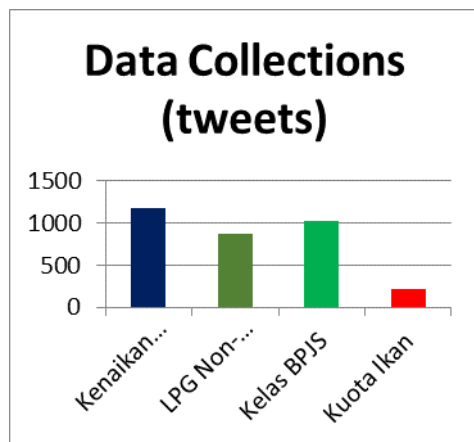


Gambar 1. Tahapan penelitian

A. Pengumpulan Data

Pada penelitian ini, pengambilan data yang dilakukan pada sosial media Twitter yang digunakan masyarakat untuk mengutarakan apa yang dirasakan terkait isu-isu yang sedang trending saat itu. Kebijakan pemerintah yang dijadikan acuan dan diambil datanya antara lain data tweet tentang kebijakan kenaikan pertalite pada range tanggal 29 Oktober – 9 September 2022 dengan keywords “Kenaikan Pertalite”, tweet tentang kebijakan LPG yang menjadi tidak disubsidi lagi pada range tanggal 4-15 Juli 2022 dengan keywords “LPG Non Subsidi”, tweet tentang kebijakan Kelas BPJS yang sudah dinaikan tarifnya pada range tanggal 18-29 Juli 2022 dengan keywords “Kelas BPJS”, dan tweet tentang kebijakan pembatasan kuota penangkapan ikan di laut Indonesia pada range tanggal 12-23 September 2022 dengan keywords “Pembatasan kuota ikan”.

Dengan API Twitter yang dipakai, didapat data-data yang diambil sesuai dengan kebijakan-kebijakan dan range tanggal yang ditentukan. Hasil penarikan data kebijakan ditunjukkan pada Gambar 2.



Gambar 2. Grafik penarikan data twitter per kebijakan

Dari pengambilan data yang dilakukan, terdapat 1176 data tweet yang membahas kebijakan Kenaikan Peralite pada masa kebijakan itu sedang dibahas, 865 data tweet yang membahas kebijakan LPG Non-Subsidi, 1025 data tweet tentang kebijakan Kelas BPJS, dan 213 data tweet yang dibuat oleh masyarakat untuk membicarakan kebijakan pembatasan kuota penangkapan ikan [1]. Pada penelitian ini, data kebijakan Kenaikan Peralite, LPG Non-Subsidi, dan Kelas BPJS sangat mumpuni untuk dipakai lebih lanjut sebagai data analisis sentimen karena data tweet mereka sudah banyak yaitu 800-1200 data. Sedangkan data kebijakan Kuota Ikan sangat kurang dari segi kuantitas untuk diolah sentimennya. Oleh karena itu, penelitian ini membahas 3 (tiga) data kebijakan berbeda.

B. Pengolahan Data

Pada penelitian ini, data yang diambil berupa tweet-tweet yang dibuat tentang tiap kebijakan di media sosial twitter yang masih sangat mentah yang terdiri dari tag, link, dan beberapa mention akun, serta kapitalisasi tulisannya. Kesemuanya itu diparsing dibuat menjadi satu persatu dokumen. Contoh head dari data tweets tentang kebijakan Kenaikan Peralite ditunjukkan pada Tabel 1.

TABEL I
DATA TWEET TENTANG KEBIJAKAN KENAikan PERALITE

No	Username	Tweet
1.	@bangbhust	Sungguh tidak memikirkan rakyatnya yang sedang kesulitan. Peralite ...
2.	@sfwtlh13	RT Kebiasaan rezim dzolim!! Liat sumber https://www.kompas.com/.../
3.	@hannaimelda	Kenaikan harga BBM ini dasarnya apa?!! Pasti ada sebab akibat jadi ..
4.	@JerrynxChndr	Tinyurl.com/bsHbsbeg RT @anDhKaTrp Peralite naik hidup sulit ...
5	@lukalama	Mantep bgt emg subsidi BBM kita dicatut TEROSSS RT @blkment ...

Setelah dilakukan penarikan data, kemudian data tersebut dibersihkan dengan menggunakan normalization dan case folding. Normalization ini dilakukan untuk menghapus tag html seperti
, hashtag, dan tanda baca seperti tanda seru (!), tanda tanya (?), dan lain sebagainya. Sedangkan case folding digunakan untuk membuat semua kapital menjadi huruf kecil sehingga semua menjadi seragam. Hasil dari normalization dan case folding ditunjukkan pada Tabel II.

TABEL III
DATA HASIL NORMALIZATION DAN CASE FOLDING

No.	Username	Tweet
1.	@bangbhust	sungguh tidak memikirkan rakyatnya yang sedang kesulitan. pertalite ...
2.	@sfwtlh13	kebiasaan rezim dzolim. liat sumber ...
3.	@hannaimelda	kenaikan harga bbm ini dasarnya apa. pasti ada sebab akibat jadi ..
4.	@JerrynxChndr	@anDhKaTrp pertalite naik hidup sulit ...
5	@lukalama	mantep bgt emg subsidi bbm kita dicatut terossss @blkment...

Langkah selanjutnya dalam pengolahan data yaitu Tokenization dimana data setiap tweet tersebut akan dipecah per kata yang akan menghasilkan token-token array di setiap datanya. Untuk membuat semua teks tersebut dapat dibaca oleh mesin, mereka telah diekstraksi ke dalam vektor fitur numerik bernama Vectorization. Jadi, semua data diterjemahkan (tokenized) ke dalam kamus biner yang disebut bag of words [11]. Hal ini dilakukan untuk mempermudah langkah selanjutnya yaitu mengeliminasi Stop Words dan melakukan Stemming. Eliminasi Stop Words guna menghapus kata-kata dalam stop words list pada dokumen tiap tweet. Dalam penelitian ini kata-kata stop words list antara lain; dan, tetapi, lalu, juga, RT, dan lain sebagainya. Lalu perlakuan Stemming pada dokumen tiap tweet berarti memotong kata yang berlebihan menjadi kata dasarnya seperti “memperkejakan” menjadi “kerja”, “kenaikan” menjadi “naik” [12]. Pada Tabel III terdapat hasil data yang berubah dari hasil normalization dan case folding menjadi diterapkan tokenization, stop words, dan stemming.

TABEL IIIII
TOKENIZATION, STOP WORDS, DAN STEMMING

No.	Username	Tweet
1.	@bangbhust	{{“sungguh”, “tidak”, “pikir”, “rakyat”, “sulit”, “pertalite”, ... }}
2.	@sfwtlh13	{{“biasa”, “rezim”, “dzolim.”, “liat”, “sumber”, ... }}
3.	@hannaimelda	{{“naik”, “harga”, “bbm”, “dasar”, “pasti”, “sebab”, “akibat”, ... }}
4.	@JerrynxChndr	{{“@anDhKaTrp”, “pertalite”, “naik”, “hidup”, “sulit”, ... }}
5	@lukalama	{{“mantep”, “bgt”, “emg”, “subsidi”, “bbm”, “catut” “terossss”, ... }}

C. Model Natural Language Processing (NLP)

Penelitian ini menggunakan opinion mining untuk mengambil klasifikasi sentimen dari dataset yang ada dengan menggunakan model Natural Language Processing (NLP) Rule Based. Proses ini melakukan k-fold cross validation yang menentukan k bagian untuk membagi data-data tweet tersebut menjadi testing dataset dan training dataset. Dalam NLP dikenal sebagai process pengolahan teks dimana mesin tidakbisa memahami bahasa natural dari manusia maka perlu dibantu dengan adanya naturallanguage processing [13]. Dalam pembentukan sebuah NLP memerlukan sebuah pipelinedimana untuk model mesin yang dibuat dalam paper ini hanya menggunakan tahapan: word tokenization, Stopwords removal, Stemming, Build bag of words, Split datasets, Vectorizer [14]-[16].

Dalam model NLP Rule Based ini, terdapat perhitungan probabilitas yang memiliki persamaan (1)[12].

$$P(Y|X)=P(x|y)P(Y)P(X) \quad (1)$$

Pada persamaan (1) ini, Y adalah kelas yang sangat spesifik dan X menunjukkan data yang berada di kelas yang masih tidak terklasifikasi, serta $P(Y|X)$ merupakan hasil probabilitas hipotesis yang berasal dari kondisi tertentu. Representasi probabilitas sebelumnya dari suatu kelas yang didasari kondisi hipotesis dituliskan dengan $P(Y)$ dan $P(X|Y)$. $P(X)$ menunjukkan probabilitas dari Y. Dari persamaan (1) tersebut, NLP rule based mengembangkannya menjadi probabilitas yang lebih kompleks dalam persamaan (2).

$$P(Y|X_1, X_2, \dots, X_n) = \frac{P(X_1, X_2, \dots, X_n | Y) P(Y)}{P(X_1, X_2, \dots, X_n)} \\ = \frac{P(X_1 | Y) P(X_2 | Y) \dots P(X_n | Y) P(Y)}{P(X_1, X_2, \dots, X_n)} \quad (2)$$

Dalam persaaan (2) tersebut, $P(Y|X_1, X_2, \dots, X_n)$ adalah hasil perhitungan yang didapat dari semua probabilitas posterior oleh nilai X untuk semua nilai di Y, sehingga klasifikasi ini dapat membuat prediksi sentimen yang didasari oleh probabilitas maksimum dari probabilitas posterior laplace yang ditunjukkan pada persamaan (3) berikut dimana c merupakan jumlah nilai pada Y.

$$P(X_i | Y) = \frac{N_{ic} + 1}{N_c + c} \quad (3)$$

D. TD-IDF

Dalam penelitian ini, vektorisasi agar sekumpulan teks dapat dilakukan perhitungan menggunakan TF-IDF yang dilakukan dalam proses NLP untuk menghasilkan analisis sentimen. Term Frequency–Inverse Document Frequency adalah teknik vektorisasi pembobotan nilai untuk dokumen teks yang akan memperhitungkan bahwa kata-kata tersebut dapat dikatakan mengekstraksi informasi yang baik, buruk ataupun tidak berpengaruh sama sekali. Pembobotan kata dilakukan dengan menggunakan algoritme TF-IDF dengan menghitung jumlah frekuensi

kata pada dokumen (Salton dan Buckley 1998; Berger et al. 2000). Tabel IV merupakan hasil TF-IDF dari dataset kebijakan “Kenaikan Peralite”.

TABEL IVV
HASIL TF-IDF DATASET KEBIJAKAN “KENAIKAN PERTALITE”

Lexicon	R1	R2	TF1	TF2	IDF	TFIDF1	TFIDF2
naik	1	1	0.35	0.4	0.48	0.108	0.098
bbm	1	1	0.35	0.4	0.465	0.065	0.088
hina	1	0	0	0.4	0.465	0.02	0.085
akibat	0	1	0	0	0	0	0

E. Data Validation

Metode analisis yang terlatih akan dianalisis dengan menggunakan model evaluasi performa kinerja. Ada empat ukuran berbeda dalam metrik ini yang menentukan kinerja hasilnya [17]:

True Positive (TP): Ulasan positif yang diklasifikasikan sebagai positif dengan benar

True Negative (TN): Ulasan negatif yang diklasifikasikan sebagai negatif dengan benar

False Positive (FP): Ulasan positif yang diklasifikasikan sebagai negatif secara salah.

False Negative (FN): Ulasan negatif yang diklasifikasikan sebagai positif secara salah.

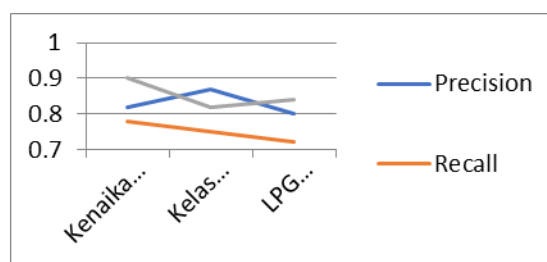
Setelah itu, semua metrik tersebut akan dihitung sebagai skor Precision, Recall, dan F-measures yang mendefinisikan ke dalam rata-rata presisi dan recall tertimbang. Dan ini adalah rumusnya:

$$\text{Precision (P)} = \frac{TP}{TP+FP} \quad (4)$$

$$\text{Recall (R)} = \frac{TN}{TN+FN} \quad (5)$$

$$\text{F-measure (F)} = \frac{2PR}{P+R} \quad (6)$$

Metode Data Validation yang dipakai dalam penelitian ini adalah dengan menggunakan confusion matrix untuk mencari evaluation metric antara lain nilai Precision, Recall, dan F-Measure, serta keakuratan model dengan dataset yang sudah disiapkan. Dengan data True Positive (TP), False Positive (FP), True Negative (TN), dan False Negative (FN) yang didapatkan dari model yang melakukan sentimen ke beberapa dataset twitter tersebut dengan mengimplementasikan training set dan testing set dari tiap-tiap kebijakan. Didapat hasil data untuk menunjukkan Precision, Recall, dan F-Measure yang tergambarkan pada grafik ditunjukkan pada Gambar 3.



Gambar 3. Grafik evaluation metric model analisis sentimen

Untuk evaluation metric pada dataset kebijakan “Kenaikan Peralite”, nilai Precision adalah 0.82, 0.78 untuk nilai Recall, dan 0.9 untuk nilai F-Measure. Model untuk dataset kebijakan “Kelas BPJS” didapatkan nilai Precision 0.87, 0.75 untuk nilai Recall, dan 0.82 untuk nilai F-Measure. Sedangkan untuk dataset kebijakan “LPG Non-Subsidi” didapatkan 0.8 untuk nilai Precision, 0.72 untuk nilai Recall, dan 0.84 untuk nilai F-Measure. Dilihat dari hasil evaluasi data tersebut, model analisis sentimen ini sudah cukup baik untuk diterapkan, karena nilai F-Measure sudah diatas angka 0.80 (bahkan pada dataset kebijakan “Kenaikan Peralite” didapatkan nilai F-Measure tinggi yaitu 0.90) dimana menjadi tanda bahwa model tersebut sudah bisa merepresentasikan hasil animo sesungguhnya. Aakurasi dari model pada penelitian ini sudah memiliki nilai 0.863.

III. HASIL DAN PEMBAHASAN

Pada penelitian ini, data dan model telah diimplementasikan dan berikut ini adalah pembahasan analisis dari hasil yang didapatkan [4]. Sentimen kebijakan pemerintah dilakukan dari tiga kebijakan yaitu “Kenaikan Peralite”, “Kelas BPJS”, dan “LPG Non-Subsidi”. Hasil sentimen kebijakan pemerintah dilakukan dari tiga kebijakan yaitu “Kenaikan Peralite”, “Kelas BPJS”, dan “LPG Non-Subsidi”.

A. Kebijakan “Kenaikan Peralite”

Hasil polarisasi dari sentimen untuk kebijakan “Kenaikan Peralite” ini ditunjukkan pada Gambar 4. Hal ini menunjukkan persebaran nilai polarity yang terkandung dalam kata-kata yang ada dalam teks suatu tweet yang negative sebesar 30 kata, netral 12 kata, dan positif 10 kata yang terasosiasi pada tweet-tweet yang ada tentang kebijakan “Kenaikan Peralite”.

text_word_count	
sentiment	
negative	18.90
neutral	12.21
positive	20.50

Gambar 4. Polarisasi Sentimen kebijakan “Kenaikan Peralite”

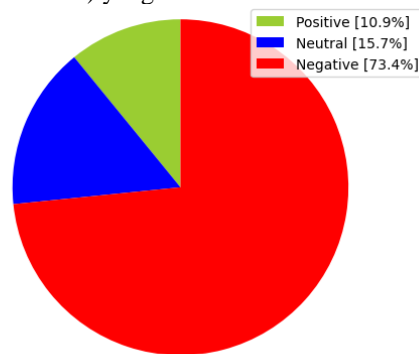
Kata-kata yang berkaitan dengan keyword “Kenaikan Peralite” diasosiasikan dengan gambar Wordcloud yang terdapat pada Gambar 5. Dari data wordcloud didapat bahwa terdapat beberapa kata-kata asosiasi yang bernada ketidaksetujuan dengan kebijakan tersebut seperti “Astaghfirullah”, “hina”, “turun”, “masalah”, dan “akibat”, serta “berdampak”. Menariknya, sentimen pada kebijakan ini juga memperlihatkan bahwa masyarakat membuat golongannya sendiri yang diperkuat dengan asosiasi kata seperti “kalangan”, “rakyat”, “mereka”, dan permintaan “subsidi” dilanjutkan dijalankan. Lalu tentunya sentimen dari kebijakan ini juga menunjukkan ketidaksukaan atau

bentuk aksi mengkritik atas pihak pemerintahan yang saat itu sedang bertanggung jawab. Hal ini didukung dengan beberapa kata asosiasi yang muncul berkaitan dengan “Jagal Cebong” (julukan buruk yang diberikan masyarakat untuk pemerintahan Jokowi), “Jokowi”, dan “dzalim”.



Gambar 5. Wordcloud kebijakan “Kenaikan Peralite”

Pada data tweet tentang kebijakan “Kenaikan Peralite”, didapatkan hasil analisis sentimen yang ditunjukkan pada Gambar 6. Grafik di Gambar 6 menunjukkan bahwa dari 1176 tweet tentang kebijakan “Kenaikan Peralite”, terdapat 73,4% tweet bernada negatif yaitu sekitar 863 tweet menggunakan kata-kata yang bersifat buruk, 15,7% tweet bernada netral yaitu sekitar 185 tweet menggunakan kata-kata yang normal (beberapa diantaranya hanya bersifat informatif seperti untuk berita), dan 10,9% tweet bernada positif dimana 128 tweet menyatakan setuju atas kenaikan harga pertalite tersebut. Hal ini menunjukkan bahwa masyarakat Indonesia banyak tidak setuju dengan pendapat dan narasi tentang kebijakan kenaikan bahan bakar mesin (Pertalite) yang akan dinaikkan.



Gambar 6. Analisis Sentimen kebijakan “Kenaikan Peralite”

B. Kebijakan “Kelas BPJS”

Hasil polarisasi dari sentimen untuk kebijakan “Kelas BPJS” ini ditunjukkan pada Gambar 7. Hal ini menunjukkan persebaran nilai polarity yang terkandung dalam kata-kata yang ada dalam teks suatu tweet yang negatif sebesar 19 kata, netral 12 kata, dan positif 20 kata yang terasosiasi pada tweet-tweet yang ada tentang kebijakan “Kelas BPJS”.

text_word_count	
sentiment	
negative	18.90
neutral	12.21
positive	20.50

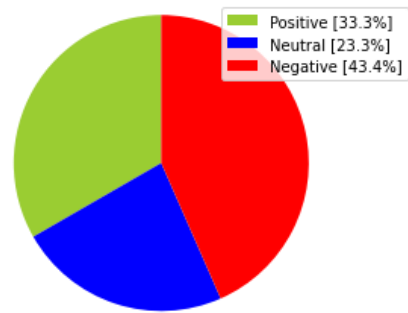
Gambar 7. Polarisasi Sentimen kebijakan “Kelas BPJS”

Wordcloud yang menggambarkan kata-kata asosiasi apa saja yang terkait dengan keyword “Kelas BPJS” ditunjukkan pada Gambar 8. Terdapat beberapa kata-kata asosiasi yang bernada ketidaksetujuan dengan kebijakan tersebut seperti “tihanjing”, “susah”, “rebutan”, “mahal”, dan “yakali”, serta “remuk”. Beberapa kata-kata asosiasi pada sentimen disini juga mengkritisi tentang alur proses BPJS yaitu “pelayanan”, “kursi”, “tertinggi”, dan “bayar”. Namun, kata-kata asosiasi yang bernada positif disini sangat sedikit sehingga dipertanyakan untuk sentimen yang bermakna positif sebesar 33,3% di atas.



Gambar 8. Wordcloud kebijakan “Kelas BPJS”

Pada data tweet tentang kebijakan “Kelas BPSJ”, didapatkan hasil Analisis Sentimen yang ditunjukkan pada Gambar 9. Gambar 9 menunjukkan grafik yang merepresentasikan dari total 1025 tweet, ada 43,4% tweet bermakna negatif yaitu sekitar 445 tweet menggunakan kata-kata yang menunjukkan hal kontra dari kebijakan yang akan dilakukan, 23,3% tweet bermakna netral yaitu sekitar 239 tweet menggunakan kata-kata yang tidak mengandung hal negatif dan positif yang sifatnya penyebaran informasi saja, dan 33,3% tweet bernada positif dimana 341 tweet menyatakan pernyataan pro atas kebijakan Kelas BPJS yang akan dinaikkan harganya. Hal ini menunjukkan bahwa sentimen pada kebijakan “Kelas BPJS” ini mayoritas tidak setuju, namun ternyata banyak juga yang mendukung atas pengimplementasian kenaikan harga Kelas BPJS ini.



Gambar 9. Analisis Sentimen kebijakan “Kelas BPJS”

C. Kebijakan “LPG Non-Subsidi”

Hasil polarisasi dari sentimen untuk kebijakan “LPG Non-Subsidi” ini ditunjukkan pada Gambar 10. Hal ini menunjukkan persebaran nilai polarity yang terkandung dalam kata-kata yang ada dalam teks suatu tweet yang negative sebesar 17 kata, netral 14 kata, dan positif 18 kata yang terasosiasi pada tweet-tweet yang ada tentang kebijakan “LPG Non-Subsidi”.

text_word_count	
sentiment	
negative	17.11
neutral	14.97
positive	18.81

Gambar 10. Polarisasi Sentimen kebijakan “LPG Non-Subsidi”

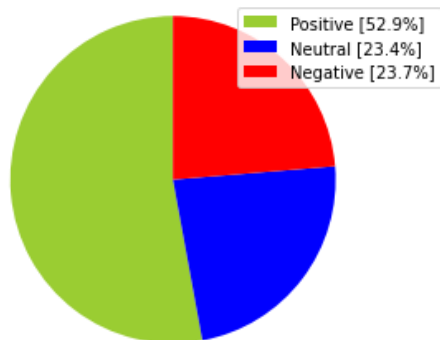
Sentimen tentang kebijakan “LPG Non-Subsidi” menghasilkan kata-kata asosiasi yang terikat dengan keyword “LPG Non-Subsidi”. Hasil tersebut ditunjukkan pada Gambar 11 yang berbentuk wordcloud. Banyak kata-kata dengan konotasi positif yang baik keluar dalam wordcloud tersebut seperti “liburan”, “aman”, “good job”, “alhamdulillah”, “kesiapan”, dan “tenang”. Hal ini sangat berbeda dengan dua pembahasan kebijakan diatas dimana animo masyarakat sangatlah positif dalam menghadapi kebijakan ini.



Gambar 11. Wordcloud kebijakan “LPG Non-Subsidi”

Gambar 12 menunjukkan hasil analisis sentimen untuk kebijakan “LPG Non-Subsidi”. Grafik tersebut menunjukkan presentasi dari 865 tweet tentang kebijakan

“LPG Non-Subsidi” terdapat hanya 23,7% tweet yang berisikan kata negatif yaitu sekitar 205 tweet menggunakan kata-kata tidak setuju dengan kebijakan tersebut, 22,4% tweet berisikan kata netral yaitu sekitar 202 tweet sifatnya sebagai berita saja, dan 52,9% tweet positif dimana 458 tweet menyatakan pernyataan hal kebaikan tentang kebijakan pemberhentian subsidi untuk pembelian gas LPG. Hal ini merepresentasikan bahwa sentimen pada kebijakan “LPG Non-Subsidi” ini banyak yang menyetujui dan mengamininya untuk tetap dilakukan.



Gambar 12. Analisis Sentimen kebijakan “LPG Non-Subsidi”

Penggunaan metode pada studi kasus yang tidak sesuai dapat memberikan hasil yang tidak optimal [18]. Apabila dibandingkan dengan model penelitian sebelumnya terkait dengan analisis sentimen didapatkan perbandingan hasil akurasi yang ditunjukkan pada Tabel V [19][20].

TABEL V
HASIL AKURASI METODE

Kriteria	NLP Rules Based	NBC	KNN
Akurasi	86.30%	63.09%	69.23%
Precision	81.33%	73.86%	79.75%
Recall	83.30%	78.50%	81.25%

Wordcloud yang dihasilkan pada penelitian ini sangat beragam sesuai dengan label yang didapatkannya dimasing-masing kebijakan. Pada penelitian sebelumnya yang mengungkap analisis sentimen kebijakan pemindahan ibu kota, namun untuk kata-kata yang keluar dalam wordcloud hanyalah kata umum yang terkait saja, tidak ada kata dengan nada positif atau negatif yang signifikan.

IV. KESIMPULAN

Analisis Sentimen dilakukan pada data tweet yang sudah dilakukan pengolahan data yaitu scrapping, normalization, tokenization, stop words, dan stemming. Terdapat tiga kebijakan yang menjadi acuan kepercayaan masyarakat yaitu kebijakan “Kenaikan Paltalite”, “Kelas BPJS”, dan “LPG Non-Subsidi”. Sentimen paling negatif didapat saat kebijakan “Kenaikan Paltalite” dimana terdapat 73,4% tweet bermakna negatif yang mengkritisi pemerintah. Kebijakan “Kelas BPJS” juga mendapat

sentimen negatif dengan presentasi tweet negatif sebesar 43,4% yang mengeluhkan kebijakan baru kenaikan harga kelas BPJS dan proses alur BPJS. Namun, untuk kebijakan “LPG Non-Subsidi”, hasil analisis sentimennya menunjukkan bahwa 52,9% tweet menyetujui kebijakan tersebut bahkan dengan kata-kata asosiasi yang sangat baik. Hal ini memberikan fenomena sosial media yang disebut “buzzer” karena 35% dari tweet positif tersebut berasal dari akun yang dipertanyakan kredibilitasnya. Sedangkan untuk model analisis sentimennya sudah menunjukkan bahwa model tersebut sangatlah baik dengan akurasi 86.3% yang dapat merepresentasikan bahwa kejadian yang digambarkan pada data sosial media twitter tersebut sudah sesuai dengan kenyataan yang terjadi.

UCAPAN TERIMA KASIH

Terima kasih kepada Universitas Bunda Mulia khususnya tim Penelitian, Pengembangan, dan Pengabdian kepada Masyarakat (P3M) Universitas Bunda Mulia yang telah membiayai penelitian ini dalam program Hibah Internal UBM Kampus Serpong sampai selesai.

DAFTAR PUSTAKA

- [1] Kementerian Energi dan Sumber Daya Mineral, "Soal Pembatasan BBM, Kebijakan Pemerintah Tidak Akan Memberatkan Rakyat", <https://www.esdm.go.id/en/media-center/news-archives/soal-pembatasan-bbm-kebijakan-pemerintah-tidak-akan-memberatkan-rakyat>, tanggal akses 1 Oktober 2022.
- [2] Ni Putu Gita Naraswati, Ni Putu Gita Naraswati, dkk., "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification", *SISTEMASI: Jurnal Sistem Informasi* Volume 10, Nomor 1, Januari 2021: 222-238, ISSN:2302-8149.
- [3] Muhamad Rizki Fajr, "Analisis Sentimen Publik terhadap Program Kebijakan Pemerintah untuk Mengembangkan Pendidikan Menggunakan Metode Multinomial Naive Bayes", Tesis Magister Ilmu Komputer, 2020.
- [4] N. Chadha, R. C. Gangwar, and R. Bedi, "Current Challenges and Application of Speech Recognition Process using Natural Language Processing: A Survey," *Int. J. Comput. Appl.*, vol. 131, no. 11, pp. 28–31, 2015, doi: 10.5120/ijca2015907471.
- [5] Mohammad Farid Naufal, Selvia Ferdiana Kusuma, "Analisis Sentimen pada Media Sosial Twitter Terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat Berbasis Deep Learning", *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, Vol. 8 No. 1 April 2022.
- [6] Nakov, P., Ritter, A., Rosenthal, S., Sebastiani, F., & Stoyanov, V., "SemEval-2016 task 4: Sentiment analysis in Twitter", 2016.
- [7] Rajput, A., "Natural Language Processing, Sentiment Analysis and Clinical Analytics", 2019.
- [8] Hennie Tuhuteru, "Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine. Information System Development (ISD), Vol. 5 No. 2, Juli 2020.
- [9] Aldi Nurzahputra, Much Aziz Muslim, "Analisis Sentimen pada Opini Mahasiswa Menggunakan Natural Language Processing", *Prosiding Seminar Nasional Ilmu Komputer (SNIK 2016)*.
- [10] Evasaria M. Sipayung, dkk., "Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier", *Jurnal Sistem Informasi (JSI)*, VOL. 8, NO. 1, April 2016.
- [11] Siddhartha, B. S., "An Interpretation of Lemmatization and Stemming in Natural Language Processing", 2021.
- [12] Bhustomy Hakim, "Analisa Sentimen Data Text Preprocessing Pada Data Mining Dengan Menggunakan Machine Learning", *JBASE Universitas Bunda Mulia* Vol. 4 No. 2, pp 16-22, 2021.

- [13] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Naturallanguage processing (almost) from scratch," J. Mach. Learn. Res., vol. 12, pp. 2493–2537, 2011.
- [14] A. Elmessiry, W. O. Cooper, T. F. Catron, J. Karrass, Z. Zhang, and M. P. Singh, "Triaging Patient Complaints: Monte Carlo Cross-Validation of Six Machine LearningClassifiers", JMIR Med. Informatics, vol. 5, no. 3, p. e19, 2017, doi:10.2196/medinform.7140.
- [15] V. Singh and B. Saini, "An Effective Tokenization Algorithm for Information RetrievalSystems," pp. 109–119, 2014, doi: 10.5121/csit.2014.4910
- [16] A. I. Kadhim, Y. N. Cheah, and N. H. Ahamed, "Text Document Preprocessing andDimension Reduction Techniques for Text Document Clustering," Proc. - 2014 4th Int.Conf. Artif. Intell. with Appl. Eng. Technol. ICAIET 2014, pp. 69–73, 2015, doi:10.1109/ICAET.2014.21.
- [17] Flach, P., "Performance Evaluation in Machine Learning: The Good, The Bad, The Ugly and The Way Forward", 2018.
- [18] Angelina Pramana Thenata, "Text Mining Literature Review on Indonesian Social Media", Jurnal Edukasi dan Penelitian Informatika, Vo.l. 7, No. 2, Agustus 2021.
- [19] Paulus Samotana Zalukhu, Teny Handhayani, Manatap Sitorus, "Analisis Sentimen Terhadap Kenaikan BBM di Indonesia pada Media Sosial Twitter Menggunakan Metode Naive Bayes", Jurnal Sistem Informasi dan Teknik Komputer, Vol. 8 No. 1 (2023) ISSN:2502-5899.
- [20] Dedi Pramana, M Afdal, Mustakim, Inggih Permana, "Analisis Sentimen Terhadap Pemindahan Ibu Kota Negara Menggunakan Algoritma Naive Bayes Classifier dan K-Nearest Neighbors", Jurnal Media Informatika Budidarma, Vol. 7 No. 3 (2023).