

Analisa Sentimen Mengenai Kenaikan Harga Bbm Menggunakan Metode NaïveBayes Dan Support Vector Machine

¹Ferdi Rejeki, ²Vina Ayumi

^{1,2}Universitas Mercu Buana, Indonesia

¹ferdirejeki003@gmail.com ; ²vina.ayumi@mercubuana.ac.id;

Article Info

Article history:

Received, 2023-12-29

Revised, 2023-01-21

Accepted, 2023-02-03

Kata Kunci:

Subsidi BBM

Naïve Bayes

Support Vector Machine

AlgoritmaSentimen

ABSTRAK

Kenaikan harga minyak dunia dipengaruhi oleh perang antara Rusia dan Ukraina di Eropa, yang menyebabkan guncangan finansial yang besar terhadap ekonomi dunia. Akibat perang ini, harga minyak naik hampir di seluruh dunia. Kenaikan harga minyak dunia diperkirakan akan berdampak signifikan terhadap kondisi perekonomian Indonesia. Harga minyak yang tinggi ternyata tidak menguntungkan bagi negara manapun, termasuk seluruh dunia di Asia, apalagi bagi negara kita. Kenaikan harga BBM disebabkan tingginya harga pasar dunia untuk minyak mentah dan kelangkaan domestik dibandingkan dengan permintaan. Besaran subsidi BBM yang hampir 520 triliun sangat besar dan menelan anggaran negara, sehingga pemerintah menganggap harga BBM bersubsidi dan nonsubsidi sangat mendesak. Lakukan perhitungan yang sangat cermat dan akurat untuk menghitung besaran kenaikan harga BBM di tanah air. Meningkatnya dampak kenaikan harga BBM secara simultan dan eksponensial akan berdampak pada inflasi harga terhadap fundamental ekonomi makro Indonesia. Pada platform Twitter, sering dijumpai berbagai macam tanggapan masyarakat yang merekalontarkan mengenai Kenaikan Harga BBM di Indonesia tersebut negatif maupun positif. Pada penelitian ini dilakukan analisa sentimen opini masyarakat di Twitter terhadap Kenaikan Harga BBM menggunakan Algoritma Naïve Bayes dan Support Vector Machine dengan menggunakan tools RapidMiner dan Python. Hasil eksperimen menunjukkan bahwa algoritma Support Vector Machine lebih tinggi daripada algoritma Naïve Bayes dengan memberikan nilai akurasi paling tinggi yaitu 90%.

ABSTRACT

The increase in world oil prices was influenced by the war between Russia and Ukraine in Europe, which caused a huge financial shock to the world economy. As a result of this war, oil prices rose almost all over the world. The increase in world oil prices is expected to have a significant impact on the condition of the Indonesian economy. High oil prices are not profitable for any country, including the whole world in Asia, let alone for our country. The increase in fuel prices was due to high world market prices for crude oil and domestic scarcity compared to demand. The amount of fuel subsidies which is almost IDR 520 trillion is very large and swallows up the state budget, so the government considers subsidized and unsubsidized fuel prices to be very urgent. Do a very careful and accurate calculation to calculate the amount of increase in fuel prices in the country. The increasing impact of the simultaneous and exponential increase in fuel prices will have an impact on price inflation on Indonesia's macroeconomic fundamentals. On the Twitter platform, you can often find various kinds of public comments about the increase in fuel prices in Indonesia, both negative and positive. In this study, sentiment analysis of public opinion on Twitter regarding the increase in fuel prices was carried out using the Naïve Bayes Algorithm and Support Vector Machine using RapidMiner and Python tools. The experimental results show that the Support Vector Machine algorithm is higher than the Naïve Bayes algorithm by providing the highest accuracy value of 90%.

Keywords:

Fuel subsidies

Naïve Bayes

Support Vector Machine

Algorithm

Sentiment



Penulis Korespondensi:

Vina Ayumi, S.Kom., M.Kom,
Program Studi Teknik informatika,
Universitas Mercu Buana Indonesia,
Email: vina.ayumi@mercubuana.ac.id

1. PENDAHULUAN

Saat ini media sosial berkembang di Indonesia sebagai alat komunikasi yang sangat populer, masyarakat menggunakan media sosial sebagai salah satu media komunikasi. Salah satunya adalah media sosial Twitter yang digunakan sebagai sarana promosi produk, iklan, kampanye politik, serta sebagai sarana penyampaian kritik, saran, pertanyaan, dan opini publik. Salah satu topik yang ramai di Twitter berkaitan dengan kenaikan harga BBM. Dalam hal ini, masyarakat mengkritisi kenaikan harga BBM. Analisis sentimen adalah teknik penggalian data teks untuk mendapatkan informasi tentang sentimen positif atau negatif. Analisis sentimen atau penambangan opini adalah bagian dari penambangan teks. Bidang ini mempelajari opini, perasaan, evaluasi, sikap, dan emosi orang terhadap entitas seperti produk, layanan, organisasi, individu, isu, topik, peristiwa, dan atributnya.[1]

Situasi di pasar energi sudah suram bahkan sebelum Rusia menginvasi Ukraina, karena harga minyak mentah dunia terus melonjak akibat perang Rusia-Ukraina. Itu karena pasokan minyak global tidak dapat mengimbangi pemulihan permintaan yang kuat ketika pandemi Covid-19 mereda. Sementara itu, konflik Rusia-Ukraina mengancam akan mengganggu ekspor minyak Rusia. Seperti yang kita semua tahu, Rusia adalah produsen minyak terbesar ketiga di dunia. Negara yang dipimpin oleh Vladimir Putin itu menghasilkan sekitar 10 persen pasokan minyak dunia atau sekitar 10,5 juta barel per hari.[2]

Nicke Widyawati, penanggung jawab utama Pertamina, mengatakan alasan kenaikan harga Pertamina untuk menghindari kerugian akibat lonjakan harga minyak mentah dunia akibat konflik bersenjata di Ukraina. PT Pertamina resmi menaikkan harga jual BBM jenis Ron 92 atau Pertamina dari sebelumnya Rp 9.000 per liter menjadi Rp 9.400 per liter. Kenaikan tersebut akan berlaku mulai 1 April 2022. Kenaikan harga minyak mentah dunia di pasar internasional akhir-akhir ini menjadi salah satu penyebab utama kenaikan harga BBM nonsubsidi.[3]

Seperti yang terjadi saat ini, kenaikan harga BBM di Indonesia pasti akan menyebabkan inflasi yang lebih tinggi dan daya beli masyarakat yang rendah, dampaknya akan terkonsentrasi pada mereka yang daya belinya rendah dan masyarakat miskin, termasuk di perkotaan, dan juga mereka yang di pedesaan orang. Secara keseluruhan, dapat dipahami juga bahwa kenaikan harga BBM akan berdampak eksponensial terhadap perekonomian Indonesia.

Produk BBM SPBU Pertamina meningkat, terutama terkait BBM bersubsidi. Sedangkan jenis BBM lainnya telah dibanderol dengan detail, yakni Dexlite CN 51 17.100/liter, Pertamina Dex CN 53 17.400/liter dan Pertamina Turbo RON 98 16.250/liter, hal itu terlihat pada gambar yang dirilis Pertamina pada 3 September. 2022 sebagai berikut [4].



Gambar 1 Harga Bahan Bakar Minyak Per 3 September 2022

Grafik di atas menunjukkan kenaikan harga BBM yang tajam. Sehingga bisa berdampak pada lambatnya pertumbuhan berbagai komoditas lainnya. Pada akhirnya daya beli masyarakat akan menurun, dan pertumbuhan ekonomi akan menurun. 66 Di luar itu, pertumbuhan eksponensial cukup umum terjadi di berbagai bidang. Pertumbuhan eksponensial dalam kehidupan publik dapat memberikan hasil yang mengejutkan bagi orang biasa, karena laju perubahan meningkat secara dramatis dari waktu ke waktu. Jumlah yang awalnya meningkat perlahan secara bertahap meningkat lebih tajam. Selain itu, beberapa angka yang tumbuh secara eksponensial seringkali dinyatakan dalam angka persentase tanpa label indeks, seperti angka inflasi atau suku bunga kredit, dan Badan Pusat Statistik mencatat angka pertumbuhan ekonomi RI sebesar -2,07%.Dibandingkan dengan tahun 2019, realisasi PDB ini turun sekitar 5,02%, terparah sejak krisis tahun 1998, yang mengakibatkan penurunan yang cukup memprihatinkan di hampir semua industri akibat pandemi Covid-19 yang melanda hampir seluruh dunia. pertumbuhan minus 13,16%. Kehidupan manusia mewabah di penghujung tahun 2019, akibat virus yang awalnya berasal dari China dan menyebar ke seluruh dunia, sedemikian rupa sehingga hampir semua wilayah antar negara diblokir. Selama hampir 2 tahun aktivitas dibatasi akibat virus hingga akhirnya para ilmuwan melakukan banyak percobaan hingga menemukan vaksin yang dapat mengatasi penyebaran virus yang telah membunuh jutaan orang di seluruh penjuru dunia. Kepala BPS Suhariyanto mengatakan kontraksi ekonomi Indonesia sejalan dengan pertumbuhan ekonomi di banyak negara, mitra dagang juga mencatatkan pertumbuhan negatif pada kuartal IV 2020. Secara spesifik, Amerika Serikat (AS) -2,5%, Singapura -3,8%, Korea Selatan -1,4%, Hong Kong -3%, dan Uni Eropa -4,8%. Suhariyanto menjelaskan, dari sisi pengeluaran, hampir seluruh komponen mencatatkan angka negatif sepanjang tahun 2020. Konsumsi rumah tangga yang menyumbang 57,66% terhadap PDB mengalami kontraksi sebesar 2,63%. Kemudian, lembaga nirlaba yang melayani rumah tangga (LNPR) mengalami kontraksi sebesar 4,29%, investasi sebesar 4,95%, ekspor sebesar 7,7%, dan impor sebesar 14,71%. "Kecuali konsumsi pemerintah, semua komponen negatif," kata Suhariyanto. Sejarahnya, pada 2014 misalnya, harga BBM premium yang paling banyak dikonsumsi saat itu naik hingga 30% pada November. Inflasi kemudian melonjak menjadi 8,36% (yoy)[5]. Hal yang sama terjadi setahun lalu ketika pemerintah menaikkan harga BBM pada Juni 2013 sehingga menyebabkan inflasi naik menjadi 8,38% (y-o-y). “Dua dampak kenaikan harga BBM tidak hanya pada perekonomian tetapi juga pada aspek sosial masyarakat Indonesia.BBM sangat diperlukan bagi perusahaan untuk beroperasi dan jika harga naik maka biaya produksi akan meningkat hampir di semua sektor dan industri.Oleh karena itu , perusahaan akan meminimalisir biaya operasional seperti menghentikan perekrutan baru hingga pemutusan hubungan kerja (PHK).Kenaikan BBM berpotensi meningkatkan angka pengangguran yang tentunya akan meningkatkan angka kemiskinan di Indonesia.Bahkan per Maret 2022, BPS melaporkan penurunan angka kemiskinan setelah pandemi.67 Selain itu, di sektor bisnis, 17 sektor ekonomi memiliki dampak negatif. Yang terparah dialami oleh sektor

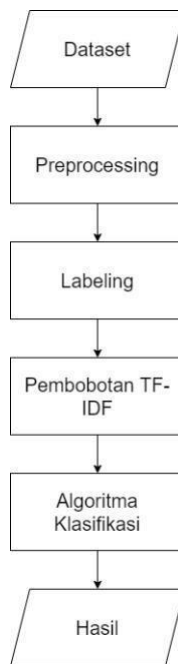
pengangkutan dan penyimpanan sebesar -15,0 persen. Selanjutnya sektor akomodasi dan makanan yang minus 10,22 persen pada 2020. Bahkan, sektor ini tumbuh lagi 5,79 persen pada 2019. menjadi minus 75,03 persen, dan penutupan beberapa hotel dan restoran selama pandemi COVID-19," kata Suhariyanto. . Setelah itu, industri pengolahan minus 2,93 persen, perdagangan minus 3,72 persen, konstruksi minus 3,26 persen, industri pertambangan -1,95 persen. Kemudian sektor jasa lainnya minus ,1 persen, jasa dunia usaha minus 5, persen dan pengadaan listrik dan gas minus 2,3 persen. "Hanya tujuh sektor yang masih tumbuh positif," tambah Suhariyanto. Diantara sektor-sektor yang positif tersebut adalah pertanian, kehutanan dan perikanan dengan peningkatan sebesar 1,75%; jasa keuangan dan asuransi sebesar 3,25%; informasi dan komunikasi meningkat sebesar 10,58%, dimana layanan pendidikan meningkat sebesar 2,63%. Setelah itu, sektor real estate meningkat 2,32%, pelayanan kesehatan dan kegiatan sosial meningkat 11,6% dan penyediaan air bersih meningkat 4,94%. "Layanan medis dan kegiatan sosial meningkat 11,6% karena pencairan pembayaran preferensi Covid-19 untuk staf medis, meningkatkan pendapatan rumah sakit untuk layanan Covid-19," jelas Suhariyanto. Dari sisi kontribusi, lanjutnya, sektor transportasi dan perdagangan menjadi akar penyebab kontraksi ekonomi minus 0,64%. Sedangkan sumber kontraksi lainnya berasal dari manufaktur negatif 0,61%, perdagangan negatif 0,49% dan konstruksi negatif 0,33%..[6]

Dari permasalahan tersebut, penulis berharap model yang telah dibuat untuk mengklasifikasikan sentimen opini publik di Twitter terkait kenaikan BBM dapat memberikan hasil yang baik sehingga dapat diimplementasikan. Penelitian ini akan menggunakan algoritma Naïve Bayes dan support vector machine, dimana algoritma ini adalah Naïve Bayes classifier (NBC) yang merupakan metode klasifikasi yang diturunkan dari teorema Bayes. Metode klasifikasi dengan menggunakan probabilitas dan metode statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu prediksi peluang masa depan berdasarkan pengalaman masa lalu, dikenal dengan teorema Bayes. Menurut teorema Bayes, Naïve Bayes adalah sebuah algoritma yang menggabungkan probabilitas apriori dan bersyarat dalam sebuah formula yang digunakan untuk menghitung probabilitas setiap pengklasifikasi pada probabilitas.

Naïve Bayes Classifier dalam melakukan klasifikasi terdapat dua proses penting yaitu learning (training) dan testing[7]. Sedangkan algoritma baru untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. Penelitian ini berpusat pada klasifikasi biner yang terdiri dua kelas, kelas tersebut adalah Positif atau Negatif.

2. METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif di mana kumpulan data (Dataset) yang dikumpulkan diproses melalui Twitter API menggunakan alat Rapid Miner. Dalam penelitian ini, alat Python dan r didukung dalam preprocessing dan pelabelan di Google Colab. Pada penelitian ini juga dilakukan pengujian yang membandingkan performa classifier Naïve Bayes dan Support Vector Machine serta validasi K-Fold untuk menentukan algoritma mana yang memiliki tingkat akurasi tertinggi. Tahapan penelitian diilustrasikan pada Gambar 2.



Gambar 2 Tahapan Penelitian

Langkah pertama adalah mengumpulkan dataset. Juga, pada langkah selanjutnya, langkah preprocessing dilakukan untuk membersihkan kumpulan data mentah. Setelah data diproses sebelumnya, langkah selanjutnya adalah memberi label untuk kelas positif dan negatif. Selanjutnya, langkah pembobotan TF-IDF dilakukan untuk menetapkan bobot pada setiap kata yang ada dalam dokumen. Kemudian data uji dan data latih dilakukan untuk menghasilkan 3 spilte. Selain itu, metode klasifikasi dilakukan dengan menggunakan algoritma Naïve Bayes, Support Vector Machine.

Sebelum tahap klasifikasi dilakukan, data terlebih dahulu dibagi menjadi data latih dan data uji. Data latih tersebut digunakan untuk proses pembelajaran pada algoritma Naïve Bayes dan Support Vector Machine. Langkah selanjutnya adalah menggunakan data uji untuk proses pengujian model klasifikasi Naïve Bayes dan Support Vector Machine. Pada penelitian ini digunakan algoritma Support Vector Machine karena cocok untuk klasifikasi teks dan cara kerja algoritma yang dapat mengatasi data outlier. Sedangkan algoritma Naïve Bayes cocok untuk data teks karena Naïve Bayes menggunakan metode probabilitas, peluang kata muncul.

A. Dataset

Data yang digunakan merupakan data yang didapat melalui media social Twitter. Total keseluruhan data yang digunakan sebanyak 5313 studi kasus Subsidi BBM. Seluruh data diperoleh dengan dibantu menggunakan tools Rapidminer. Proses pengambilan data dengan memasukkan kata kunci yang relevan dengan topik yang dibahas pada studi kasus, seperti yang ditunjukkan pada Tabel 1. Data yang dikumpulkan disimpan dalam format .xlsx

TABEL I. Sumber Data

No	Query Twitter	Jumlah Data
1	Bahan bakar minyak	456
2	BBM	512
3	Kilang minyak	198
4	Pertalite	548
5	Pertamax	1588

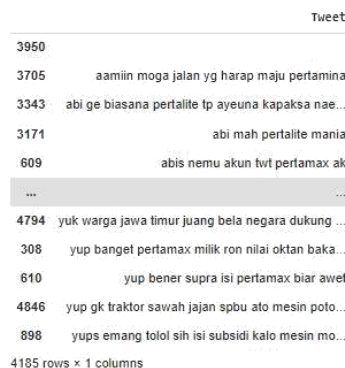
6	Pertamina	506
7	Subsidi BBM	1340
8	Harga BBM	155
Total Data		5313

B. Preprocessing Data

Preprocessing data bertujuan untuk membersihkan dataset dari Data mentah menjadi data siap pakai untuk memudahkan proses klasifikasi positif dan negatif. Data preprocessing adalah teknik penambahan data yang melibatkan konversi data mentah menjadi format yang dapat dimengerti. Dalam penelitian ini, lima langkah yang digunakan dalam tahap preprocessing . Berikut adalah penjelasan dari kelima langkah preprocessing tersebut.

- **Cleansing Data** : Cleansing dalam penelitian ini bertujuan untuk menghapus RT, username, hashtag, angka dan URL pada tweets.
- **Case Folding** : Case Folding dalam penelitian ini bertujuan untuk merubah Teks dari penulisan huruf besar menjadi penulisan huruf kecil.
- **Tokenization** : Tokenization dalam penelitian ini bertujuan untuk memisahkan kata kata dari dalam kalimat.
- **Stopword Removal** : Stopword Removal dalam penelitian ini bertujuan untuk menghapus kata sambung[8].
- **Contoh dari Stopword Removal bahasa Indonesia** adalah sebagai berikut : yang, juga, dari, dia, kami, kamu, aku, saya, ini, itu, atau, dan, pada, dengan, adalah, yaitu, ke, tak, tidak, di, pada, jika, maka, ada, pun, lain, saja, hanya, namun, seperti, kemudian, dll[9].
- **Hasil Stopword Removal** : salam, universitas, islam, internasional, indonesia, (uiii), kembali, membuka, peluang, bagi, kamu, yang, ingin, melanjutkan,
- **Stemming** : Stemming dalam penelitian ini bertujuan untuk membentuk kata kata dasar.
- **Remove Duplicate** : Remove Duplicate dalam penelitian ini bertujuan untuk menghapus data yang memiliki duplikat.

Setelah melakukan tahap Preprocessing, data yang telah diPreprocessing menjadi 4183 untuk data Subsidi BBM. Berikut merupakan hasil data setelah dilakukan tahapan Preprocessing:



The image shows a screenshot of a data table with the following content:

Tweet	
3950	
3705	aamiin moga jalan yg harap maju pertamina
3343	abi ge biasana pertalite tp ayeuna kapaksa nae...
3171	abi mah pertalite mania
609	abis nemu akun twt pertamax ak
...	...
4794	yuk warga jawa timur juang bela negara dukung ...
308	yup banget pertamax milik ron nilai oktan baka...
610	yup bener supra isi pertamax biar awet
4846	yup gk traktor sawah jajan spbu ato mesin poto...
898	yups emang tolong sih isi subsidi kalo mesin mo...

4185 rows x 1 columns

Gambar 3 Preprocessing Subsidi BBM

C. Labeling

Pada tahap ini, data yang telah dibersihkan selama preprocessing akan ditandai sebagai data class. Label kelas data dibagi menjadi dua jenis: kelas positif dan kelas negatif. Proses pelabelan dilakukan menggunakan bahasa R melalui alat Google Colab. Proses penandaan menggunakan bahasa R membutuhkan kamus bahasa Indonesia yang terintegrasi dengan Google Drive. Kamus Bahasa Indonesia yang digunakan terdiri dari dua kamus, Kamus Kata Positif Bahasa Indonesia dan Kamus Kata Negatif Bahasa Indonesia. Pada fase pelabelan, setiap kalimat diberi skor. Setiap kata yang terdeteksi diberi skor untuk mengevaluasi kategori sentimen. Ini memiliki nilai 1 untuk kata-kata positif dan -1 untuk kata-kata negatif. Jika ada kata yang tidak ada dalam kamus positif dan negatif, beri nilai

Scoring dilakukan dengan menghitung nilai total dari setiap kata dalam sebuah kalimat. Jika nilai ≥ 0 maka ditandai sebagai tweet sentimen positif, sebaliknya jika bernilai < 0 maka ditandai sebagai tweet sentimen negatif. Jika hasil penilaian adalah 0, tandai sebagai tweet positif. Kelas data positif akan diberi nilai 1, sedangkan kelas data negatif akan diberi nilai 0. TABEL II. Klasifikasi Labeling Otomatis SVM dan NB

Klasifikasi	Positif	Negatif
Otomatis	3028	1154

Pada penelitian ini juga dilakukan eksperimen dengan menggunakan labeling manual terhadap *dataset* tersebut. Label positif merupakan tweet yang memiliki pujian sedangkan tweet negatif berupa kata-kata kasar dan sarkas.

D. TF-IDF Weighting

Untuk mengetahui seberapa penting sebuah kata mewakili sebuah kalimat, dilakukan pembobotan atau perhitungan. Pemberian skor dalam *TF-IDF* berdasarkan frekuensi munculnya kata dalam dokumen [10].

E. Klasifikasi Algoritma

Text mining adalah salah satu topik penambangan data karena teks juga merupakan data. Namun, karena cakupannya yang sangat luas, istilah penambangan teks dan pengujian data digunakan secara berbeda. Penambangan data menggali pengetahuan dari kumpulan data yang besar dan seringkali terstruktur. Perbedaannya terletak pada tugas-tugas penambangan data seperti klasifikasi, pengelompokan, asosiasi, estimasi atau prediksi, bukan pada karakteristik data yang ditanganinya. Text mining adalah proses penambangan data berupa teks. Sumber data biasanya diperoleh dari dokumen. Tujuannya adalah untuk menemukan kata-kata yang dapat mewakili isi dokumen, sehingga dapat menganalisis keterkaitan antar dokumen. [11].

Analisis sentimen adalah kegiatan menganalisis opini, pendapat, sikap atau emosi seseorang tentang suatu produk, topik atau isu tertentu sehingga dapat diketahui bahwa hal tersebut mencakup sentimen positif dan negatif. [12].

Metode klasifikasi adalah teknik penambangan data yang mengelompokkan data berdasarkan lampiran data dari data sampel. Metode klasifikasi juga merupakan proses mengklasifikasikan dokumen ke dalam satu atau lebih kategori tertentu atau dokumen dari kategori yang sama. [13].

Metode Bayes Bayesian adalah pendekatan statistik untuk penalaran induktif untuk masalah klasifikasi. Kami pertama-tama membahas konsep dasar dan definisi teorema Bayes, dan kemudian menggunakan teorema tersebut untuk klasifikasi dalam penambangan data. [14]. Pengklasifikasi Naive Bayes dalam klasifikasi memiliki dua proses penting: pelatihan dan pengujian.

Support Vector Machines (SVM) adalah metode yang relatif baru untuk membuat prediksi baik dalam klasifikasi maupun regresi. Penelitian ini berfokus pada klasifikasi biner yang terdiri dari dua kelas. Apakah kategorinya positif atau negatif. Konsep klasifikasi menggunakan support vector machine adalah mencari hyperplane terbaik yang berperan sebagai data barrier untuk kedua kelas tersebut. Algoritma SVE bekerja dengan menggunakan ukuran jarak hyperplane alih-alih menemukan titik maksimum untuk memisahkan data secara optimal dari titik terdekat untuk memaksimalkan jarak antar label kelas sesuai dengan ukuran batas keanggotaan kelas. [15].

3. HASIL DAN PEMBAHASAN

Pada bagian ini akan menampilkan hasil semua eksperimen yang telah dilakukan. Pada penelitian ini melakukan eksperimen pada studi kasus yaitu Subsidi BBM. Eksperimen dilakukan dengan menggunakan dua algoritma yaitu Support Vector Machine dan Naïve Bayes. Terdapat tiga skenario yang menjadi pembagian persentase. Pembagian presentasi data merupakan pemisahan data training dan testing berdasarkan persentasenya, contoh 90%:10% berarti 90% merupakan data training dan 10% merupakan data testing. Pada penelitian ini pembagian persentase dibagi menjaditiga skenario eksperimen yaitu eksperimen pertama menggunakan pembagian data 90%:10%, eksperimen kedua menggunakan pembagian data 80%:20% dan eksperimen yang ketiga menggunakan pembagian data 70%:30%.

Hasil dari kinerja model dengan menggunakan metode Labeling otomatis dan pembagian dataatau percentage split 90%:10% bahwa studi kasus Subsidi BBM memiliki nilai akurasi 90% untukalgoritma Support Vector Machine dan 78% untuk algoritma Naïve Bayes.

TABEL III. Support Vector Machine Subsidi BBM Split 90:10 Labeling Otomatis

	Precision	Recall	Accuracy
0	86%	76 %	90%
1	92%	95%	90%

TABEL IV. Naive Bayes Subsidi BBM Split 90:10 Labeling Otomatis

	Precision	Recall	Accuracy
0	57%	73%	78%
1	89%	80%	78%

Hasil dari kinerja model dengan menggunakan metode *Labeling* otomatis dan pembagian dataatau *percentage split* 80%:20% bahwa studi kasus Subsidi BBM memiliki nilai akurasi 90% untukalgoritma Support Vector Machine dan 78% untuk algoritma Naïve Bayes.

TABEL V. Support Vector Machine Subsidi BBM Split 80:20 Labeling Otomatis

	Precision	Recall	Accuracy
0	88%	73%	90%
1	91%	97%	90%

TABEL VI. Naive Bayes Subsidi BBM Split 80:20 Labeling Otomatis

	Precision	Recall	Accuracy
0	56%	73%	78%
1	89%	<u>80%</u>	<u>78%</u>

Hasil dari kinerja model dengan menggunakan metode *Labeling* otomatis dan pembagian data atau *percentage split 70%:30%* bahwa untuk studi kasus Indihome memiliki nilai akurasi sebesar 90% untuk algoritma Support Vector Machine dan 78% untuk algoritma Naïve Bayes.

TABEL VII. Support Vector Machine Subsidi BBM Split 70:30 Labeling Otomatis

	Precision	Recall	Accuracy
0	87%	72%	90%
1	90%	96%	90%

TABEL VIII. Naïve Bayes Subsidi BBM Split 70:30 Labeling Otomatis

	Precision	Recall	Accuracy
0	56%	75%	78%
1	90%	79%	78%

Hasil dari penelitian ini adalah eksperimen dengan menggunakan algoritma Support Vector Machine memiliki nilai akurasi paling tinggi pada tiga skenario eksperimen *percentage split* pada model data yang diuji yaitu Subsidi BBM dengan menggunakan *Labeling* otomatis dan manual. Hasil dari pembahasan tercantum dalam tabel berikut ini.

TABEL IX. Akurasi Labeling Otomatis SVM dan NB

Percobaan	Otomatis	
	SVM	NB
90:10	90%	78%
80:20	90%	78%
70:30	90%	78%

4. KESIMPULAN

Dari hasil penelitian yang dilakukan dengan menggunakan algoritma Naïve Bayes, Support Vector Machine dengan pengujian labeling otomatis dapat disimpulkan bahwa eksperimen dengan menggunakan algoritma Support Vector Machine memiliki tingkat akurasi terbaik pada data yang diuji Subsidi BBM dengan nilai rata-rata akurasi 90% pada eksperimen Labeling otomatis yang paling baik pada penelitian ini.

REFERENSI

- [1] “Implementasi Opinion Mining (Analisis Sentimen) Untuk Ekstraksi Data Opini Publik Pada Perguruan Tinggi - Neliti.” <https://www.neliti.com/publications/69248/implementasi-opinion-mining-analisis-sentimen-untuk-ekstraksi-data-opini-publik> (accessed Jun. 08, 2022).
- [2] “Dampak Perang Rusia-Ukraina, Harga Minyak Dunia Sentuh 107 Dollar AS Per Barrel - Tribunnews.com.” <https://www.tribunnews.com/bisnis/2022/03/02/dampak-perang-rusia-ukraina-harga-minyak-dunia-sentuh-107-dollar-as-per-barrel> (accessed Jun. 06, 2022).
- [3] “Apa Alasan Harga Pertamina Naik? Ini Jawaban Pertamina | KASKUS.” <https://www.kaskus.co.id/thread/624fb84f18223e1eb305ea7d/apa-alasan-harga-pertamax-naik-ini-jawaban-pertamina> (accessed Jun. 06, 2022).
- [4] Welly, “Berikut Harga BBM Terbaru, Peralite Rp10.000 – Pertamina Rp14.500,” *Detik SumSel*, Sep. 03, 2022. <https://detiksumsel.com/berikut-harga-bbm-terbaru-pertalite-rp10-000-pertamax-rp14-500/>
- [5] C. I. Cantika Adinda Putri Noveria, “Kalau Peralite Naik, Ini Dampak Ngeri Bagi Menengah ke Bawah,” *CNBC Indonesia*, Aug. 20, 2022. <https://www.cnbcindonesia.com/news/20220820194435-4-365301/kalau-pertalite-naik-ini-dampak-ngeri-bagi-menengah-ke-bawah> (accessed Dec. 26, 2022).
- [6] CNN Indonesia, “Penyebab Pertumbuhan Ekonomi RI Minus 2,07 Persen pada 2020,” *CNN Indonesia*, Feb. 06, 2021. <https://www.cnnindonesia.com/ekonomi/20210206080350-532-603013/penyebab-pertumbuhan-ekonomi-ri-minus-207-persen-pada-2020> (accessed Dec. 26, 2022).
- [7] M. Rasyida, “Naïve Bayes Classification untuk Penentuan Status Penduduk Miskin,” *Jurnal Informatika Kaputama(JIK)*, vol. 4, no. 2, 2020.
- [8] W. Yulita *et al.*, “Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier,” *JDMSI*, vol. 2, no. 2, pp. 1–9, 2021.
- [9] A. Nugroho Sekolah Tinggi Teknologi Pelita Bangsa Jl Inspeksi Kalimantan, T. Danas, C. Pusat, and K. Bekasi, “Analisis Sentimen Pada Media Sosial Twitter Menggunakan Naive Bayes Classifier Dengan Ekstrasi Fitur N-Gram,” 2018. [Online]. Available: <http://tunasbangsa.ac.id/ejurnal/index.php/jsakti>
- [10] G. A. Buntoro, “Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter,” 2017. [Online]. Available: <https://t.co/jrvaMsgBdH>
- [11] R. Tineges, A. Triayudi, and I. D. Sholihati, “Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM),” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, no. 3, p. 650, Jul. 2020, doi: 10.30865/mib.v4i3.2181.
- [12] H. Annur, “KLASIFIKASI MASYARAKAT MISKIN MENGGUNAKAN METODE NAÏVE BAYES,” 2018.
- [13] H. Derajad Wijaya and S. Dwiasnati, “Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat,” *JURNAL INFORMATIKA*, vol. 7, no. 1, 2020, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/ji>
- [14] D. A. Kristiyanti, Normah, and A. H. Umam, “Prediction of Indonesia Presidential Election Results for the 2019-2024 Period Using Twitter Sentiment Analysis,” Sep. 2019.
- [15] A. M. Pravina, I. Cholissodin, and P. P. Adikara, “Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine (SVM),” 2019. [Online]. Available: <http://j-ptiik.ub.ac.id>