



Analisis Sentimen Data Ulasan Pengguna MyPertamina di Twitter dengan Metode *Machine Learning* dan *Deep Learning*

Andita Widya V. Hutabarat¹, Ni Luh Saddhwi Saraswati Adnyani², Kadarsah Suryadi³

^{1,2,3} Program Studi Teknik Industri, Fakultas Teknologi Industri, Institut Teknologi Bandung
Jalan Ganesha No. 10, Bandung, Jawa Barat, Indonesia, 40132
Email: anditahutabarat05@gmail.com, saddhwi@itb.ac.id, kadarsah@itb.ac.id

Abstract

The increasing number of MyPertamina users has led to an increasing number of reviews related to the use of MyPertamina. Positive reviews can increase user trust and encourage further use, while negative reviews can have detrimental impacts such as reducing the company's image in the eyes of customers. Negative reviews can be an indicator of low customer satisfaction, which companies need to respond quickly. Reviews of MyPertamina fill various social media channels, including Twitter. However, the analysis of user perceptions through social media has not been optimal. Therefore, a better user sentiment mapping is needed. This study was conducted to answer this need by building a text mining model and designing a prototype that can extract and analyze sentiments from tweets related to MyPertamina. This research adopts the CRISP-DM methodology, which consists of business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The data obtained for model development reached 6,920 tweet data. Each data was classified into one of three sentiment categories, namely positive, negative, and neutral. After data preparation, 2,057 data were used for model development. The models tested in this study consist of Support Vector Machine (SVM), Multinomial Naïve Bayes, Gaussian Naïve Bayes, Long Short-Term Memory (LSTM), and Bidirectional Long Short-Term Memory (Bi-LSTM) algorithms. The model that produced the best evaluation score and was selected for prototype development is the SVM model with an accuracy score of 83.74%, weighted precision of 83.96%, weighted recall of 83.74%, and weighted F1-score of 83.72%. The prototype is used for extracting and predicting sentiment for new datasets, which can then be visualized in the form of graphs and word clouds according to the user's needs.

Keywords: *text mining, machine learning, deep learning, sentiment analysis, SVM*

Abstrak

Pengguna MyPertamina semakin banyak sehingga semakin banyak juga ulasan terkait penggunaan MyPertamina. Ulasan positif dapat meningkatkan kepercayaan pengguna dan mendorong penggunaan lebih lanjut, sementara ulasan negatif dapat memiliki dampak yang merugikan seperti menurunnya citra perusahaan di mata konsumen. Ulasan negatif dapat menjadi indikator rendahnya kepuasan pelanggan, yang perlu ditanggapi dengan cepat oleh perusahaan. Ulasan-ulasan perihal MyPertamina membanjiri berbagai kanal media sosial, salah satunya Twitter. Namun, analisis persepsi pengguna dengan media sosial belum maksimal. Oleh karena itu, dibutuhkan pemetaan sentimen pengguna secara lebih baik. Penelitian ini dilakukan untuk menjawab kebutuhan tersebut dengan membangun model *text mining* serta merancang prototipe yang dapat mengekstraksi dan menganalisis sentimen dari *tweet* yang berhubungan dengan MyPertamina. Penelitian ini mengadopsi metodologi CRISP-DM yang terdiri dari pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, hingga *deployment*. Data yang diperoleh untuk pembangunan model mencapai 6.920 data tweet. Masing-masing data diklasifikasikan dalam satu dari tiga kategori sentimen, yakni positif, negatif, dan netral. Setelah persiapan data, sebanyak 2.057 data digunakan untuk pembangunan model. Model yang diuji dalam penelitian ini terdiri dari algoritma *Support Vector Machine* (SVM), *Multinomial Naïve Bayes*, *Gaussian Naïve Bayes*, *Long Short-Term Memory* (LSTM), dan *Bidirectional Long Short-Term Memory* (Bi-LSTM). Model yang menghasilkan nilai evaluasi terbaik dan terpilih dalam pembangunan prototipe adalah model SVM dengan nilai *accuracy* sebesar 83,74%, *weighted precision* sebesar 83,96%, *weighted recall* sebesar 83,74%, dan *weighted F1-score* sebesar 83,72%. Prototipe digunakan untuk ekstraksi dan prediksi sentimen set data baru untuk kemudian divisualisasikan dalam bentuk grafik dan *wordcloud* sesuai dengan kebutuhan pengguna.

Kata kunci: *text mining, machine learning, deep learning, analisis sentimen, SVM*

Pendahuluan

Bahan Bakar Minyak (BBM) merupakan suatu kebutuhan yang penting guna menunjang kegiatan masyarakat di zaman modern. Badan Usaha Milik Negara (BUMN) yang bertugas mengelola penambangan minyak dan gas bumi di Indonesia adalah PT Pertamina (selanjutnya disebut Pertamina). Pertamina terus berkembang secara aktif baik di sektor hulu maupun hilir dalam industri minyak dan gas bumi untuk menjamin ketersediaan sumber energi hingga ke konsumen.

Pertamina mengeluarkan suatu aplikasi digital yang disebut dengan MyPertamina. Aplikasi MyPertamina digunakan untuk memastikan penyaluran BBM bersubsidi, yakni solar dan pertalite, agar lebih efektif dan tepat sasaran dengan mewajibkan kendaraan penerima subsidi untuk mendaftarkan data diri dan kendaraan pada MyPertamina.

Kebijakan Pertamina terkait penjualan BBM bersubsidi jenis solar dan pertalite dengan mewajibkan masyarakat penerima BBM bersubsidi untuk mendaftarkan kendaraan pada aplikasi MyPertamina banyak mendapatkan respons dari pengguna, baik negatif maupun positif. Banyak masyarakat mengeluhkan kendala dalam penggunaan aplikasi MyPertamina, tetapi ada pula masyarakat yang mendukung usaha pemerintah dalam menjamin subsidi tepat sasaran melalui MyPertamina ini. Ulasan-ulasan perihal MyPertamina dapat dilihat pada kolom ulasan di berbagai sosial media, salah satunya yaitu Twitter. Twitter merupakan salah satu media sosial yang banyak digunakan oleh masyarakat Indonesia. Pengguna Twitter di Indonesia mencapai 18,45 juta pada tahun 2022 dan menempatkan Indonesia pada peringkat kelima negara pengguna Twitter terbesar di dunia (Key dkk., 2022).

MyPertamina dapat memanfaatkan *text mining* dalam menganalisis data berupa teks pada *tweet*. Feldman dan Sanger (2007) menyebutkan bahwa *text mining* dapat didefinisikan sebagai suatu proses untuk mencari pengetahuan, mengidentifikasi, dan eksplorasi pola yang menarik dari suatu dokumen yang dapat dilakukan dengan menggunakan beberapa alat analisis. Seiring dengan meningkatnya informasi dalam bentuk teks, maka penelitian untuk mengekstraksi

informasi dari teks pun turut meningkat, hal ini dilakukan dengan tujuan untuk menghasilkan informasi berkualitas tinggi dari teks (Han dkk., 2012). Terdapat beberapa bentuk penerapan *text mining*, salah satunya adalah analisis sentimen.

Analisis sentimen atau *opinion mining* adalah pendekatan komputasional menggunakan teknik NLP untuk mengidentifikasi pendapat, sikap, emosi, dan penilaian dari suatu entitas spesifik yang sifatnya subjektif untuk membangkitkan pengetahuan yang terstruktur dan berguna (Afidah dkk., 2022; Bian dkk., 2022; Bisio dkk., 2017). Analisis sentimen merupakan *tool* yang dapat digunakan untuk mengekstraksi dan menganalisis emosi dan pandangan publik, mendapatkan *insight* bisnis, dan menarik kesimpulan yang lebih baik (Birjali dkk., 2021).

Dengan analisis sentimen, MyPertamina diharapkan dapat terbantu dalam menyortir *tweet* yang masuk dalam kelas sentimen yang berbeda, sehingga dapat dengan mudah menganalisis fitur atau layanan yang mendapatkan banyak respon positif dari masyarakat maupun keluhan atau ulasan negatif dari masyarakat. Ketentuan pemberian label sentimen pada penelitian ini ditunjukkan pada Tabel 1 berikut.

Tabel 1. Penentuan label sentimen

Sentimen	Ketentuan
Positif	Mengandung unsur pujian terhadap MyPertamina, bersifat mengajak atau mempersuasi orang lain untuk menggunakan MyPertamina, atau menyatakan kecenderungan untuk memilih MyPertamina
Netral	Bersifat informatif, berita yang tidak memiliki konotasi negatif maupun positif, pertanyaan tanpa menjurus ke konotasi negatif maupun positif, atau keikutsertaan dalam <i>giveaway</i> dan sebagainya.
Negatif	Mengandung unsur komplain dari konsumen, mengandung unsur kesulitan dalam menggunakan MyPertamina, atau bersifat skeptis terhadap MyPertamina

Terdapat beberapa model yang dapat diimplementasikan untuk analisis sentimen. Pemilihan model terbaik tergantung dari beberapa faktor, seperti ukuran dan kualitas data, tipe sentimen, kompleksitas masalah, serta kebutuhan daya komputasi. Menurut Xu dkk (2018), metode analisis sentimen dapat dibagi menjadi tiga kategori, yakni metode analisis sentimen berdasarkan kamus sentimen

(*sentiment dictionary*), *machine learning*, dan *deep learning*.

Metode berdasarkan *sentiment dictionary* menggunakan kamus untuk mengidentifikasi kata sentimen dalam sebuah teks sehingga mampu menentukan nilai sentimen teks tersebut. Kemudian, berdasarkan aturan perhitungan sentimen, kecenderungan sentimen untuk teks tersebut dapat diperoleh. Metode ini memiliki kemudahan karena tidak perlu pelabelan manual dan mudah untuk diimplementasikan. Akan tetapi, kualitas dari analisis menjadi sangat bergantung pada kamus sentimen yang dimiliki. Jika kamus sentimen tidak memiliki cukup kata-kata sentimen, maka kualitas sentimen analisis menjadi rendah.

Metode analisis sentimen dengan *machine learning* dilakukan dengan menggunakan sekumpulan algoritma yang digunakan untuk melatih set data historis dan membuat prediksi masa depan. Metode analisis sentimen dengan *machine learning* cenderung lebih akurat dibandingkan *sentiment dictionary*, tetapi bergantung pada kualitas dari *corpus* yang dilabel dengan polaritas. *Corpus* adalah sekumpulan teks yang dipilih dan disimpan dalam basis data. Contoh pemanfaatan analisis sentimen dengan *machine learning* adalah pemanfaatan algoritma *Naïve Bayes* (Sholahudin, 2022; Maulana dkk., 2023) dan SVM (Habyba dkk., 2021; Mustasaruddin dkk., 2023).

Sementara itu, metode analisis sentimen dengan *deep learning* memanfaatkan kemampuan untuk mengimitasi struktur dan fungsi dari otak manusia untuk mengidentifikasi sesuatu. *Deep learning* memanfaatkan *multi-layered Artificial Neural Network* (ANN) yang memiliki konsep serupa dengan jaringan neural dalam otak manusia. Jaringan ini mengandung berbagai *nodes* dan *multi-layer* yang berkomunikasi satu dengan lainnya untuk memahami dan menganalisis data yang diperoleh. Model *deep learning* terdiri dari beberapa *layer*. Setiap *layer* mungkin terdiri dari berbagai neuron. Neuron dari setiap *layer* disebut dengan *processing unit*. Peran utama dari neuron-neuron adalah untuk mengalikan *input value* dengan bobot yang diberikan dan mengalkulasikan jumlah dari hasilnya (Alaskar & Saba, 2021).

Pada beberapa penelitian, seperti penelitian yang dilakukan Alghifari (2022), metode *deep learning* menghasilkan akurasi yang lebih tinggi

sebesar 90,97% untuk algoritma Bi-LSTM dan 75,63% untuk algoritma LSTM dibandingkan metode *machine learning* dengan algoritma *Multinomial Naïve Bayes* sebesar 63%, *Logistic Regression* sebesar 67%, dan Linear SVC sebesar 65%. Sementara itu, pada penelitian lain, seperti yang dilakukan oleh Nikmah dkk. (2022), metode *machine learning* menghasilkan akurasi yang lebih tinggi sebesar 86,54% untuk algoritma SVM dan 85,45% untuk algoritma *Naïve Bayes* dibandingkan metode *deep learning* dengan algoritma LSTM sebesar 84,62%.

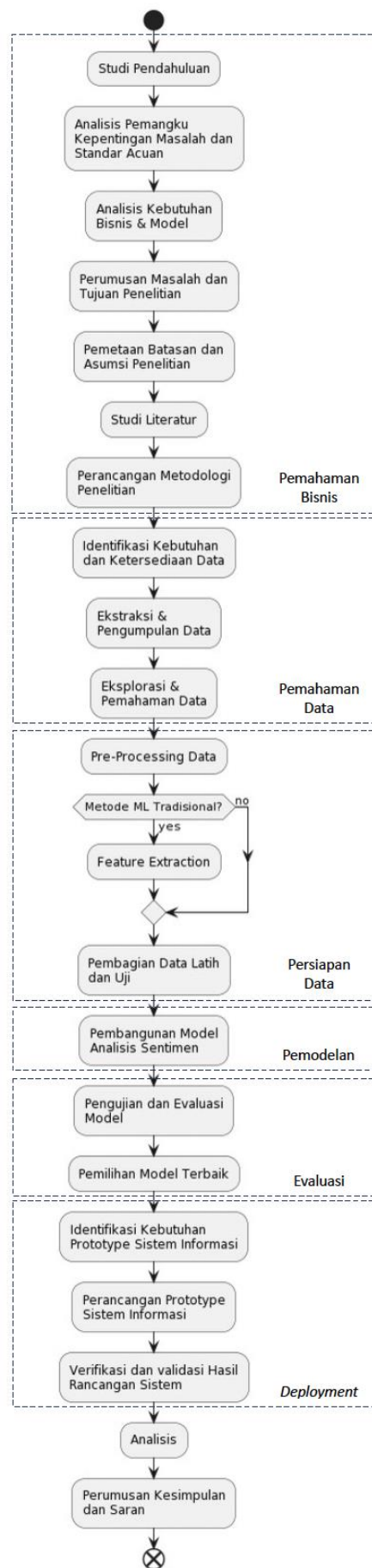
Oleh karena itu, pada penelitian ini, akan dilakukan analisis sentimen MyPertamina dengan sumber data Twitter menggunakan model *machine learning* dan model *deep learning* untuk memperoleh model dengan akurasi terbaik. Model *machine learning* tradisional yang akan digunakan adalah *Support Vector Machine* (SVM), *Multinomial Naïve Bayes*, dan *Gaussian Naïve Bayes*. Model *deep learning* tradisional yang akan digunakan adalah *Long Short-Term Memory* (LSTM) dan *Bidirectional Long Short-Term Memory* (Bi-LSTM).

Terdapat beberapa penelitian terdahulu yang dilakukan untuk menganalisis sentimen terhadap aplikasi MyPertamina, diantaranya yaitu penelitian yang dilakukan oleh Sholahudin (2022) yang menggunakan metode *Naïve Bayes* dengan sumber data Twitter, Mustasaruddin dkk. (2023) yang menggunakan metode SVM dengan sumber data *Google Play Store*, dan Maulana dkk. (2023) yang menggunakan metode *Naïve Bayes Classifier* dengan sumber data *Google Play Store*. Selain pembangunan model, pada penelitian ini juga dibangun suatu prototipe alat bantu ekstraksi data *tweet* dan pemetaan sentimen tiap *tweet*. *Output* dari prototipe disajikan dalam bentuk visualisasi data yang dapat memudahkan pengguna perototipe dalam mengetahui *pain* dan *gain* dari layanan dan produk digital MyPertamina.

Metodologi

Penelitian ini dilakukan dengan langkah-langkah seperti yang diilustrasikan pada Gambar 1. Langkah-langkah tersebut mengadopsi metodologi *Cross Industry Standard Process for Data Mining* (CRISP-DM). Terdapat enam tahap pada metodologi CRISP-DM, yaitu tahap pemahaman bisnis,

pemahaman data, persiapan data, pemodelan, evaluasi, hingga *deployment* (Schröer dkk., 2021).



Gambar 1. Diagram alir metodologi penelitian

Tahap pemahaman bisnis merupakan tahapan awal yang dilakukan pada penelitian ini. Tahap ini terdiri dari studi pendahuluan, analisis pemangku masalah dan standar acuan, analisis kebutuhan bisnis dan model, perumusan masalah dan tujuan penelitian, pemetaan batasan dan asumsi penelitian, studi literatur, serta perancangan metodologi penelitian.

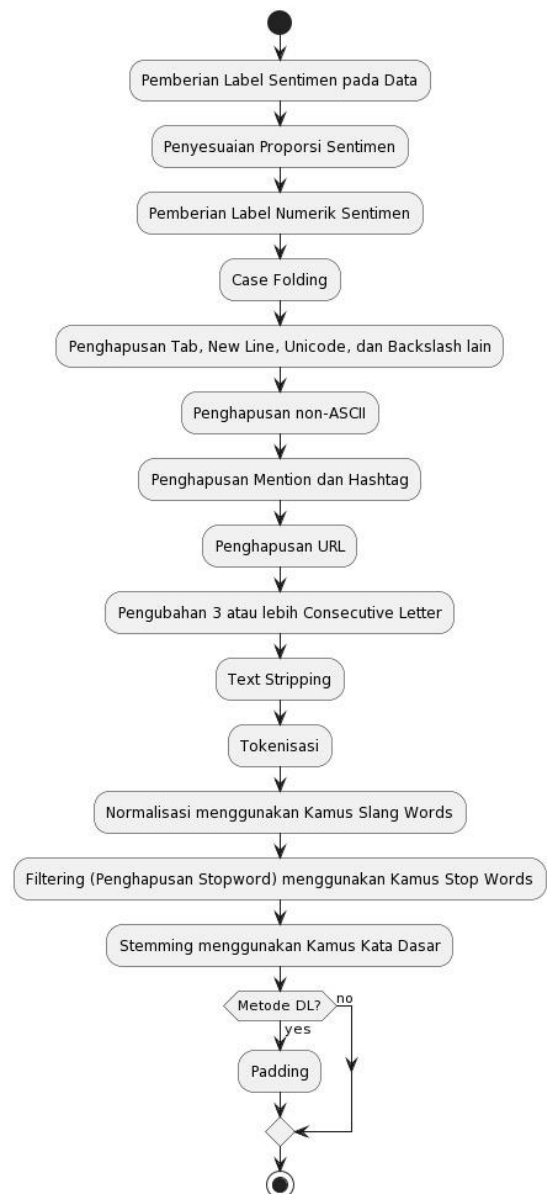
Selanjutnya, tahap pemahaman data merupakan tahapan yang terdiri dari identifikasi kebutuhan dan ketersediaan data, ekstraksi dan pengumpulan data, serta eksplorasi dan pemahaman data sebelum data tersebut disiapkan dan dimodelkan untuk mendapat analisis sentimen sesuai kebutuhan penelitian. Ekstraksi dan pengumpulan data dilakukan dengan *scraping* memanfaatkan *library* Tweepy menggunakan kode Python pada Google Colab. Tweepy adalah salah satu *library* Python yang mudah digunakan untuk mengakses API dari twitter.

Pengumpulan data dilakukan dengan menetapkan *query* yang akan ditarik, mendefinisikan beberapa spesifikasi dalam penarikan *query*, menentukan batas waktu mulai dan waktu selesai dari data *tweet* yang diekstraksi, serta menentukan jumlah *tweet* maksimal dalam sekali ekstraksi. Selain *tweet*, akan ditarik juga informasi terkait akun yang membuat *tweet* tersebut. Data yang berhasil diekstraksi, akan disimpan sementara dalam format *dataframe*. Data yang dikumpulkan pada penelitian ini adalah *tweet* berbahasa Indonesia yang mengandung tulisan “MyPertamina” atau “My Pertamina” (tidak *case sensitive*) selama periode pengamatan 1 Juli 2022 – 31 Januari 2023. Data yang diperoleh untuk mencapai 6.920 data *tweet*.

Tahap persiapan data merupakan tahapan yang terdiri dari *preprocessing data*, *feature extraction*, dan pembagian data latih (*training*) dan data uji (*testing*) untuk dapat mengevaluasi model yang dibuat. Tahap *preprocessing data* terdiri dari beberapa aktivitas, yang ditunjukkan pada Gambar 2.

Normalisasi dilakukan dengan mengubah *slang words* menggunakan kamus *slang words*. *Stop words* dihapus menggunakan kamus *stop word* Bahasa Indonesia dari *library* NLTK. *Stemming* dilakukan dengan memanfaatkan *package* Sastrawi yang sudah diinstalasi dengan kelas yang sudah diimpor sebelumnya. Setelah persiapan data, sebanyak 2.057 data digunakan untuk pembangunan model. Tabel 2

menunjukkan perbandingan data ulasan pengguna sebelum dan sesudah melalui tahap *preprocessing data*.



Gambar 2. Tahap *preprocessing data*

Pada tahap pemodelan, dilakukan pemodelan data dengan menggunakan model *machine learning* tradisional, yang terdiri dari SVM, Multinomial Naïve Bayes, Gaussian Naïve Bayes, serta model *deep learning*, yang terdiri dari LSTM dan Bi-LSTM. Pada pemodelan data dengan model SVM, akan dicari *hyperparameter* yang menghasilkan model SVM optimal. *Hyperparameter* yang akan dikombinasikan pada SVM adalah C, kernel, dan gamma. Pembangunan model akan dilakukan dengan kode Python melalui Google Colab.

Tabel 2. Hasil *preprocessing data*

Sebelum	Sesudah
@CintakarenaNya @pertamina Emang subsidi BBM jadi tepat sasaran sekarang sejak ada myPertamina	subsidi bbm sasar mypertamina

Tahap evaluasi model merupakan tahapan yang dilakukan untuk menguji hasil kinerja model klasifikasi penelitian. Tahap evaluasi terdiri dari pengujian model dan pemilihan model terbaik.

Tahap *deployment* merupakan tahapan pembangunan sistem informasi yang berguna bagi pemangku kepentingan dari model yang telah dibuat dan dievaluasi. Tahap *deployment* terdiri dari identifikasi kebutuhan prototipe sistem informasi, perancangan prototipe sistem, verifikasi, dan validasi hasil rancangan sistem. Prototipe akan dibangun dengan pada Google Colab memanfaatkan Flask-Ngrok.

Hasil dan Diskusi

Terdapat lima model yang digunakan pada penelitian ini. Tiga model di antaranya merupakan model *machine learning* tradisional, yakni SVM, *Multinomial Naïve Bayes*, dan *Gaussian Naïve Bayes*. Di antara berbagai model *machine learning* tradisional, SVM dan *Naïve Bayes* dipilih dalam penelitian ini karena kedua algoritma ini merupakan algoritma yang paling sering digunakan dalam menyelesaikan permasalahan klasifikasi sentimen (Medhat dkk., 2014). Adapun algoritma *Naïve Bayes* dapat dibagi menjadi tiga jenis, yakni *Bernoulli*, *Multinomial*, dan *Gaussian Naïve Bayes*. Karena *Bernoulli Naïve Bayes* hanya dapat dilakukan untuk distribusi biner, algoritma ini tidak digunakan dalam penelitian ini karena penelitian ini berlandaskan pada tiga sentimen, yakni positif, negatif, dan netral.

Sementara itu, dua metode lainnya merupakan metode *deep learning*, yakni *Long Short Term Memory* (LSTM) dan *Bidirectional Long Short Term Memory* (Bi-LSTM). LSTM dan Bi-LSTM merupakan jenis spesifik dari RNN yang memiliki kemampuan lebih dalam menyimpan memori jangka panjang dan memberikan hasil yang lebih akurat.

Pada beberapa penelitian pendahulu, model *deep learning* (DL) sudah dikenal atas kemampuannya dalam mengungguli model *machine learning* (ML) tradisional di berbagai domain, seperti keamanan siber, NLP, bioinformatika, robotik dan kontrol, dll.

(Alzubaidi dkk., 2021). Oleh karena itu, penelitian ini dilakukan untuk membandingkan hasil evaluasi model-model ML tradisional dan DL terhadap klaim tersebut dan melanjutkan proses pembuatan prototipe berdasarkan pertimbangan hasil evaluasi dari setiap metode yang digunakan.

Hasil evaluasi model-model ML tradisional dan DL pada penelitian ini dapat dilihat pada Tabel 3. Evaluasi model pada penelitian ini dilakukan dengan mempertimbangkan kriteria *accuracy* (K1), *weighted precision* (K2), *weighted recall* (K3), dan *weighted F1 score* (K4).

Tabel 3. Hasil evaluasi

Model	K1	K2	K3	K4
SVM	0,8301	0,8348	0,8301	0,8296
SVM Optimal	0,8374	0,8396	0,8374	0,8372
<i>Multinomial Naïve Bayes</i>	0,7840	0,7943	0,7840	0,7835
<i>Gaussian Naïve Bayes</i>	0,7160	0,7202	0,7160	0,7127
LSTM	0,7646	0,7695	0,7646	0,7647
Bi-LSTM	0,8107	0,8151	0,8107	0,8117

Berdasarkan Tabel 3, dapat diketahui bahwa seluruh model yang diuji menghasilkan akurasi di atas 70%, yang menunjukkan bahwa seluruh model masih dapat dikategorikan sebagai model yang baik (Barkved, 2022). Jika dibandingkan antar model *machine learning*, diperoleh bahwa SVM optimal menghasilkan performansi lebih baik dari segi nilai *accuracy*, *precision*, *recall*, maupun *F1-score* dibandingkan model *machine learning* lainnya.

Model SVM optimal merupakan model yang dihasilkan dengan melakukan *hyperparameter tuning* pada model SVM, diperoleh hyperparameter terbaik yakni 'C' = 1, 'gamma' = 'scale', dan 'kernel' = 'linear'. Proses *hyperparameter tuning* ini berhasil meningkatkan nilai *accuracy*, *precision*, *recall*, dan *F1 score* dari model SVM awal.

Sementara itu, jika dibandingkan antar model *deep learning*, model Bi-LSTM menghasilkan performansi lebih baik dari segi nilai *accuracy*, *precision*, *recall*, maupun *F1-score* dibandingkan LSTM. Namun, jika dibandingkan dengan SVM optimal, keempat nilai evaluasi model Bi-LSTM masih lebih rendah, walau tidak signifikan. Dapat terlihat pada Tabel 3 bahwa selisih nilai akurasi SVM optimal dengan Bi-LSTM hanya sebesar 2,67%,

weighted precision sebesar 2,45%, *weighted recall* sebesar 2,67%, dan *weighted F1-score* sebesar 2,55%.

Hal ini menunjukkan bahwa walaupun SVM optimal memiliki nilai performansi yang lebih tinggi untuk saat ini, ada kemungkinan performansi model berubah jika ditambahkan lebih banyak data. Oleh karena itu, selain melalui evaluasi secara kuantitatif, dilakukan evaluasi secara kualitatif untuk memperkuat keputusan dalam penggunaan model yang akan diteruskan dalam pembangunan prototipe sistem informasi. Hasil evaluasi kualitatif tersebut ditunjukkan pada Tabel 4.

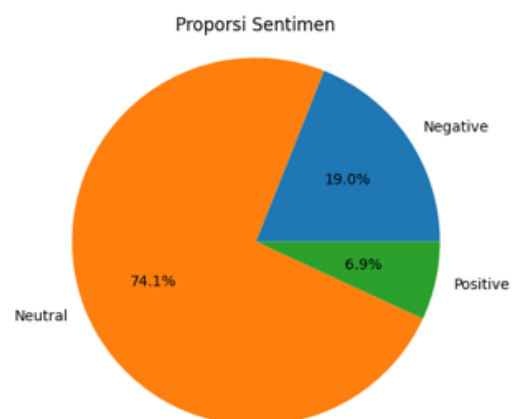
Tabel 4. Hasil evaluasi kualitatif

Aspek	Penilaian Keunggulan		Alasan
	SVM Optimal	Bi-LSTM	
Dependensi dan jumlah data	✓		Tweeepy hanya memperbolehkan pengguna untuk menarik data sampai seminggu ke belakang sehingga jika ingin dilakukan pengumpulan data untuk perbaikan model, maka proses akan berlangsung cukup lama. Karena sekarang jumlah <i>dataset</i> masih relatif kecil, maka model SVM lebih baik dalam menggenerasi akurasi yang wajar.
Dependensi perangkat keras (<i>hardware</i>)	✓		Agar pengguna dapat mengoperasikan sistem pada lebih banyak mesin, maka dipilih model SVM yang dapat bekerja pada mesin kelas <i>low-end</i> .
<i>Feature engineering</i>		✓	Tidak perlu menyimpan dan memuat <i>file</i> terkait hasil <i>feature extraction</i> pada prototipe.
Waktu eksekusi data latih	✓		Sudah dibuktikan sendiri membutuhkan waktu lebih sedikit untuk melatih data.
Waktu <i>testing</i>	✓		Sudah dibuktikan sendiri membutuhkan waktu lebih sedikit untuk menguji data.
Kemudahan interpretasi	✓		Parameter relatif lebih mudah diinterpretasikan.
Kompleksitas		✓	Diharapkan mampu menangkap pola yang lebih kompleks.

Berdasarkan hasil evaluasi secara kualitatif, model SVM optimal dipilih untuk digunakan dalam perancangan prototipe sistem yang dapat digunakan untuk menerapkan model *text mining* dalam melakukan ekstraksi dan analisis sentimen dari Twitter MyPertamina.

Pada penelitian ini, digunakan tiga jenis visualisasi data untuk prototipe yang dibangun. Visualisasi pertama menggunakan *pie chart* untuk menunjukkan proporsi persebaran sentimen dari keseluruhan data yang diekstraksi pada percobaan tersebut. Gambar 3 menunjukkan contoh tampilan *pie chart* dari prototipe yang dibangun. Ukuran *pie* dari setiap bagian merepresentasikan proporsi dari setiap sentimen. Agar mendapatkan data yang lebih akurat, ditampilkan juga label persentase dari setiap sentimen.

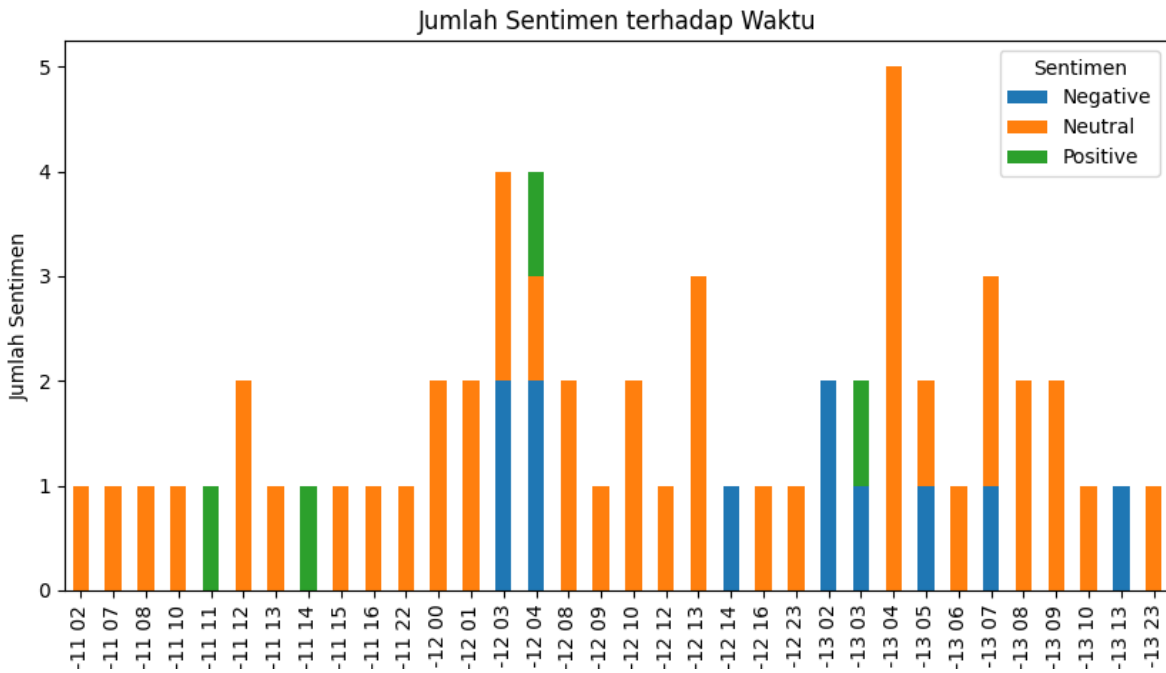
Visualisasi kedua menggunakan *stacked column chart*. *Stacked column chart* pada prototipe juga berfungsi untuk menunjukkan jumlah persebaran sentimen. Perbedaannya dengan *pie chart* adalah pada *stacked column chart*, persebaran sentimen dinilai setiap jam. Oleh karena itu, dapat terlihat jelas kapan pengguna Twitter banyak menuliskan ulasan bersentimen positif, negatif, dan netral di media sosial mereka. Gambar 4 menampilkan *stacked column chart* dari interval waktu 1 Februari 2023 00:00 hingga 3 Februari 00:00. Diperoleh bahwa dalam interval waktu tersebut, ulasan paling banyak muncul yaitu pada 1 Februari 2023 pukul 07:00 – 07:59 yang didominasi oleh sentimen positif dan netral. Pengguna dapat menghubungkan dengan kondisi di dunia nyata yang sedang terjadi terkait MyPertamina. Hal ini juga akan lebih berguna untuk mendeteksi kejadian yang memicu sentimen negatif.



Gambar 3. Contoh tampilan *pie chart* pada prototipe

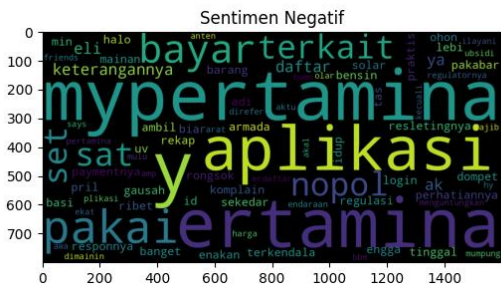
Visualisasi ketiga menggunakan *wordcloud* untuk menampilkan data teks dalam bentuk visual sehingga mudah dibaca dan dipahami dalam waktu singkat. *Wordcloud* dapat memberikan *insight* terkait kata-kata yang

paling sering muncul dalam *dataset* teks yang dianalisis. Pada *wordcloud*, ukuran dari setiap kata merepresentasikan frekuensi kemunculan kata. Kata terbesar merupakan topik yang paling sering disebut.

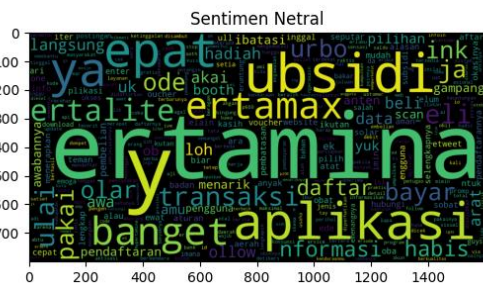


Gambar 4. Contoh tampilan *stacked column chart* pada prototipe

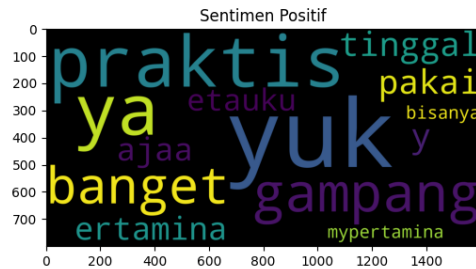
Wordcloud ini dibuat sebanyak tiga kali untuk masing-masing sentimen. Gambar 5, Gambar 6, dan Gambar 7 berturut-turut merepresentasikan kata-kata yang memiliki sentimen negatif, netral, dan positif yang sering muncul pada *dataset tweet* yang diekstraksi dan dianalisis.



Gambar 5. Contoh tampilan *wordcloud* sentimen negatif pada prototipe



Gambar 6. Contoh tampilan *wordcloud* sentimen netral pada prototipe



Gambar 7. Contoh tampilan *wordcloud* sentimen positif pada prototipe

Ulasan pengguna memiliki pengaruh yang signifikan terhadap MyPertamina, Ulasan positif dapat meningkatkan kepercayaan pengguna dan mendorong penggunaan lebih lanjut, sementara ulasan negatif dapat memiliki dampak yang merugikan seperti menurunnya citra perusahaan di mata konsumen. Ulasan negatif dapat menjadi indikator rendahnya kepuasan pelanggan, yang perlu ditanggapi dengan cepat oleh perusahaan.

Hasil penelitian ini diharapkan dapat memberikan manfaat bagi pihak MyPertamina. Karena data yang digunakan merupakan data dari konsumen yang diperoleh secara langsung, penggunaan data tersebut dalam pengembangan bisnis MyPertamina dapat membantu perusahaan untuk melakukan *data-driven design*. Penerapan *data-driven design* ini

akan membantu perusahaan dalam menghasilkan solusi yang lebih efektif dan efisien. Efektif karena melalui prototipe yang dibangun, perusahaan dapat memetakan keluhan maupun pujian dari pelanggan sehingga dapat membantu proses perbaikan kualitas layanan dan produk MyPertamina yang sesuai dengan masalah dan kebutuhan pasar. Efisien karena proses pemetaan itu dapat dilakukan dengan cepat melalui visualisasi dalam bentuk grafik maupun *wordcloud*. Dengan demikian, waktu dalam memfilter *tweet* yang kurang relevan dapat dipersingkat. Selain itu, penggunaan prototipe *dashboard* juga tidak memerlukan biaya karena menggunakan perangkat *open source* dan dibangun tanpa menggunakan penyedia jasa tertentu.

Kesimpulan

Pada penelitian ini, telah dilakukan analisis sentimen terhadap MyPertamina dengan sumber data Twitter menggunakan model *machine learning* dan model *deep learning*. Pada penelitian ini juga dilakukan perancangan prototipe yang digunakan untuk menerapkan model terpilih dalam melakukan ekstraksi dan analisis sentimen dari Twitter MyPertamina.

Pada penelitian ini, data ulasan pengguna yang digunakan untuk melakukan analisis sentimen diperoleh dari media sosial Twitter. Namun, untuk menghasilkan pemahaman yang lebih komprehensif mengenai sentimen pengguna, dapat dilakukan penelitian lebih lanjut dengan membandingkan data yang berasal dari Twitter dengan data dari media sosial lain atau data ulasan pengguna MyPertamina yang ada di Google Play Store.

Model terbaik yang dipilih dalam perancangan prototipe adalah model *Support Vector Machine* (SVM) dengan *hyperparameter* 'C' = 1, 'gamma' = 'scale', dan 'kernel' = 'linear'. *Accuracy* yang diperoleh adalah sebesar 83,74%, *weighted precision* sebesar 83,96%, *weighted recall* sebesar 83,74% dan *weighted F1-score* sebesar 83,72%. Nilai-nilai evaluasi ini mengindikasikan masih terdapat ruang untuk perbaikan. Salah satunya yaitu dengan menambah data untuk pemodelan. Model lain yang dipertimbangkan pada penelitian ini mungkin saja dapat menghasilkan nilai evaluasi yang lebih tinggi sehingga dapat menggantikan model SVM eksisting.

Pada prototipe yang dirancang, terdapat tampilan visualisasi data berupa *pie chart*,

stacked column chart, dan *wordcloud*. Pada penelitian selanjutnya, dapat dipertimbangkan alternatif visualisasi lain untuk merepresentasikan topik atau fitur yang banyak disebutkan pada setiap sentimen dengan lebih jelas.

Daftar Pustaka

- Afidah, D. I., Dairoh, Handayani, S. F., Pratiwi, Wijayatun, R., & Sari, S. N. (2022). Sentimen Ulasan Destinasi Wisata Pulau Bali Menggunakan Bidirectional Long Short Term Memory. *Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, 21(3), 607-618.
- Alaskar, H., & Saba, T. (2021). *Machine Learning and Deep Learning: A Comparative Review*. Riyadh: Prince Sultan University.
- Alghifari, D. R., Edi, M., & Firmansyah, L. (2022). Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia. *Jurnal Manajemen Informatika (JAMIKA)*, 89-99.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaria, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *Journal of Big Data*.
- Barkved, K. (2022). *How To Know if Your Machine Learning Model Has Good Performance*, [Online], Diakses dari [obviously.ai](https://www.obviously.ai/post/machine-learning-model-performance): <https://www.obviously.ai/post/machine-learning-model-performance> [2022, 9 March].
- Bian, Y., Ye, R., Zhang, J., & Yan, X. (2022). Customer Preference Identification from Hotel Online Reviews: A Neural Network Based Fine-Grained Sentiment Analysis. *Computers & Industrial Engineering*, 172.
- Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A Comprehensive Survey on Sentiment Analysis: Approaches, Challenges and Trends. *Knowledge-Based Systems* 226.
- Bisio, F., Oneto, L., & Cambria, E. (2017). Sentic Computing for Social Network Analysis. *Sentiment Analysis in Social Networks* (hal. 71-90). Cambridge: Elsevier.
- Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in*

- Analyzing Unstructured Data*. New York, NY: Cambridge University Press.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques* (3rd ed.). Waltham, MA: Morgan Kaufmann.
- Habyba, A. N., Rahmawati, N., & Triwulandari. (2021). Analisis Sentimen Mahasiswa untuk Perbaikan Desain Afektif Ruang Kelas Jurusan Teknik Industri, Universitas Trisakti. *Jurnal Rekayasa Sistem Industri*, 10(1), 27-34.
- Key, B., Kohl, A., Elflein, J., Puri-Mirza, A., Sapun, P., & Cherowbrier, J. *Leading countries based on number of Twitter users as of January 2022*, [Online], Diakses dari Statista: <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/> [2022, 22 November].
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Engineering Journal*, 5(4), 1093-1113.
- Maulana, R., Voutama, A., & Ridwan, T. (2023). Analisis Sentimen Ulasan Aplikasi Mypertamina pada Google Play Store Menggunakan Algoritma NBC. *Jurnal Teknologi Terpadu*, 9(1), 42-48.
- Mustasaruddin, Budianita, E., Fikry, M., & Yanto, F. (2023). Klasifikasi Sentiment Review Aplikasi MyPertamina Menggunakan Word Embedding FastText dan SVM (Support Vector Machine). *Jurnal Sistem Komputer dan Informatika (JSON)*, 526-534.
- Nikmah, T., Ammar, M., Allatif, Y., Husna, R., Kurniasari, P., & Bahri, A. (2022). Comparison of LSTM, SVM, and Naive Bayes for Classifying Sexual Harassment Tweets. *Journal of Soft Computing Exploration*, 3(2), 131 - 137.
- Schröer, C., Kruse, F., & Marx Gómez, J. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science*, 181, 526-534.
- Sholahudin, A. (2022). *Analisis Sentimen terhadap Kebijakan beli Peralite menggunakan Aplikasi MyPertamina pada Twitter menggunakan Metode Naive Bayes Classifier (NBC)*. Tangerang: Universitas Muhammadiyah Tangerang.
- Xu, G., Meng, Y., Qiu, X., Yu, Z., & Wu, X. (2019). Sentiment Analysis of Comment Texts Based on BiLSTM. *IEEE Access* 7, 51522-51532.