

# Analisis Sentimen terhadap Aplikasi PLN Mobile

Author:

Mochammad Reyhan Mauluddi 6032222003

Okky Wicaksono Adi 6032222108

Suryo Utomo 6032222062

*Abstract. The PLN Mobile application, developed by PT PLN (Persero) to facilitate customer access to electricity services, faces low customer satisfaction rates. This research aims to analyze customer sentiment towards the PLN Mobile application using TF-IDF based sentiment analysis and test the performance of three classification algorithms: Logistic Regression, Naive Bayes, and Support Vector Machine (SVM). Descriptive analysis reveals a dominance of negative sentiment, with words like "complaint" and "power outage" dominating the word cloud, indicating issues faced by users regarding electricity services. Evaluation results show that Logistic Regression exhibits the best performance in classifying sentiment, achieving higher accuracy, precision, recall, and F1-score compared to the other two models. The study concludes that users of the PLN Mobile application have a low level of satisfaction, with numerous complaints expressed regarding electricity services. These findings can be utilized by PT PLN to improve the quality of the PLN Mobile application and service responsiveness, thereby enhancing customer satisfaction and building a positive image for the application.*

## Latar Belakang

Revolusi digital telah membawa perubahan signifikan dalam berbagai sektor, termasuk sektor ketenagalistrikan. Seiring dengan perkembangan teknologi informasi dan komunikasi, PT PLN (Persero) meluncurkan aplikasi PLN Mobile sebagai upaya untuk meningkatkan kualitas layanan dan mempermudah akses pelanggan terhadap berbagai layanan kelistrikan. Aplikasi ini menawarkan beragam fitur yang praktis dan mudah diakses, seperti pengecekan tagihan, pembayaran tagihan, pengajuan pasang baru, perubahan daya, pelaporan gangguan, hingga pemantauan konsumsi listrik.

Meskipun aplikasi ini memiliki potensi besar untuk meningkatkan kepuasan pelanggan dan efisiensi layanan, tantangan dalam mencapai tujuan tersebut tidak dapat diabaikan. Salah satu tantangan utama adalah memastikan bahwa aplikasi PLN Mobile benar-benar menjawab kebutuhan dan harapan pelanggan. Untuk memahami persepsi pelanggan terhadap aplikasi tersebut, penelitian ini akan menggunakan teknik analisis sentimen.

Analisis sentimen akan memberikan gambaran objektif mengenai kepuasan pelanggan terhadap aplikasi PLN Mobile. Data sentimen yang dikumpulkan dari platform digital seperti ulasan pengguna di Google Play Store, akan diproses menggunakan teknik pemrosesan bahasa alami (NLP) untuk mengidentifikasi sentimen positif, negatif, dan netral. Sehingga, penelitian ini bertujuan untuk menganalisis kualitas aplikasi PLN Mobile dengan menggunakan teknik analisis sentimen. Analisis sentimen merupakan teknik yang memanfaatkan pemrosesan bahasa alami (NLP) untuk mengidentifikasi dan mengekstrak opini dan sentimen yang terkandung dalam teks, baik itu berupa ulasan pelanggan, komentar di media sosial, maupun forum diskusi online (Sinaga et al., 2021).

Urgensi analisis sentimen terhadap aplikasi PLN Mobile terletak pada kebutuhan untuk memahami persepsi pelanggan terhadap aplikasi tersebut. Sentimen positif menunjukkan tingkat kepuasan yang tinggi dan mendorong pelanggan untuk menggunakan aplikasi secara berkelanjutan. Sebaliknya, sentimen negatif mengindikasikan adanya masalah dan ketidakpuasan yang perlu segera ditangani (Gupta et al., 2020). Melalui analisis sentimen, PLN dapat memperoleh informasi yang berharga mengenai aspek-aspek aplikasi yang perlu ditingkatkan, seperti kelancaran proses transaksi, responsivitas layanan pelanggan, dan kemudahan navigasi.

Dengan memahami sentimen pelanggan terhadap aplikasi PLN Mobile, PT PLN dapat mengambil langkah-langkah strategis untuk meningkatkan kualitas layanan. Langkah-langkah tersebut dapat berupa perbaikan pada fitur aplikasi, peningkatan responsivitas layanan pelanggan, dan penyelesaian masalah yang dihadapi pengguna. Hasil analisis sentimen ini diharapkan dapat membantu PLN dalam meningkatkan kepuasan pelanggan dan membangun citra positif di mata publik.

Artikel ini memberikan kontribusi penting dalam memahami persepsi pelanggan terhadap aplikasi PLN Mobile. Hasil penelitian ini diharapkan dapat menjadi acuan bagi PT PLN dalam mengoptimalkan pengembangan aplikasi PLN Mobile, sehingga dapat memberikan layanan yang lebih baik dan meningkatkan kepuasan pelanggan.

### **Kajian Pustaka**

Analisis sentimen, sebuah teknik yang semakin populer dalam pemrosesan bahasa alami (NLP), berfokus pada identifikasi dan ekstraksi opini dan sentimen dari teks (Kavitha et al., 2022). Salah satu metode yang umum digunakan dalam analisis sentimen adalah *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF adalah metode yang membantu dalam menentukan pentingnya suatu kata atau frasa dalam sebuah dokumen, dengan mempertimbangkan frekuensi kemunculan kata tersebut dalam dokumen dan koleksi dokumen (Bafna et al., 2016). TF-IDF menggabungkan dua metrik utama: *Term Frequency* (TF) yang menunjukkan seberapa sering sebuah kata muncul dalam dokumen tertentu, dan *Inverse Document Frequency* (IDF) yang menunjukkan seberapa jarang sebuah kata muncul dalam koleksi dokumen. Nilai TF-IDF dihitung dengan mengalikan nilai TF dan IDF, sehingga kata yang sering muncul dalam sebuah dokumen dan jarang muncul dalam koleksi dokumen akan memiliki nilai TF-IDF yang tinggi, menunjukkan kata tersebut penting dan unik untuk dokumen tersebut.

Dalam konteks analisis sentimen, TF-IDF berperan penting dalam mengidentifikasi kata-kata atau frasa yang memiliki nilai sentimen tinggi. Kata-kata dengan nilai TF-IDF tinggi cenderung menjadi kata kunci yang mencerminkan sentimen positif atau negatif dalam dokumen. Metode TF-IDF menawarkan beberapa keuntungan dalam analisis sentimen, seperti kemampuannya untuk mempertimbangkan frekuensi kata dalam dokumen, mempertimbangkan kelangkaan kata dalam koleksi dokumen, dan kemudahan implementasinya. Namun, TF-IDF juga memiliki kelemahan, yaitu tidak mempertimbangkan konteks kata dalam kalimat atau paragraf, serta tidak mempertimbangkan makna sebenarnya dari kata tersebut.

Untuk mengatasi kelemahan ini, TF-IDF dapat dikombinasikan dengan classifier seperti Support Vector Machine (SVM), Logistic Regression, dan Naive Bayes. Classifier ini membantu dalam mengklasifikasikan sentimen berdasarkan fitur-fitur yang diekstraksi menggunakan TF-IDF. SVM menggunakan pemisah hiperplane untuk mengklasifikasikan data (Li et al., 2005), sedangkan Logistic Regression menggunakan fungsi sigmoid untuk memprediksi probabilitas kelas (Kavitha et al., 2022), dan Naive Bayes mengasumsikan independensi antar fitur untuk mengklasifikasikan data (Kim et al., 2002).

### **Metodologi**

Analisis sentimen terhadap aplikasi PLN Mobile dilakukan melalui rangkaian tahapan yang terstruktur untuk mendapatkan pemahaman yang komprehensif tentang persepsi pelanggan. Tahap pertama meliputi pengumpulan data sentimen dari berbagai platform digital seperti ulasan pengguna di Google Play Store. Data yang diperoleh kemudian dibersihkan dan diproses untuk mempersiapkannya untuk analisis sentimen. Tahap preprocessing ini mencakup langkah-langkah seperti *lowercasing*, tokenisasi, *stop word removal*, *stemming/lemmatization*, normalisasi, dan penghapusan karakter khusus.

Setelah data bersih dan siap diproses, tahap berikutnya adalah ekstraksi fitur menggunakan teknik TF-IDF. TF-IDF membantu dalam mengidentifikasi kata-kata atau frasa yang memiliki nilai sentimen tinggi, yang selanjutnya digunakan sebagai fitur untuk klasifikasi sentimen. Model klasifikasi, seperti SVM, Logistic Regression, dan Naive Bayes, dilatih dengan menggunakan fitur-fitur yang telah diekstraksi. Model terbaik yang dipilih berdasarkan performanya pada data pelatihan kemudian digunakan untuk mengklasifikasikan sentimen dalam data uji. Hasil klasifikasi kemudian dievaluasi menggunakan metrik seperti akurasi, presisi, recall, dan F1-score untuk menilai kemampuan model dalam mengklasifikasikan sentimen dengan tepat.

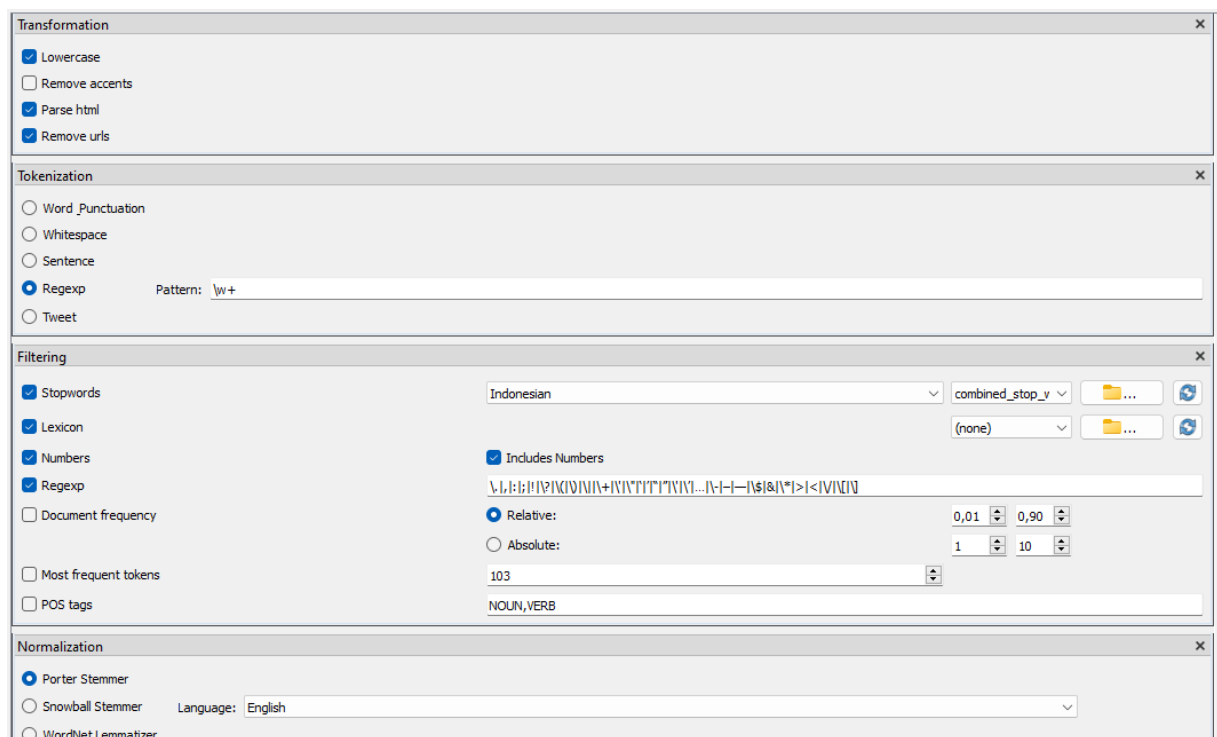
Seluruh rangkaian metodologi ini dilakukan dengan menggunakan software Orange. Orange adalah perangkat lunak yang bersifat open-source dan mudah digunakan untuk analisis data dan pembelajaran mesin. Orange menawarkan antarmuka grafis yang intuitif, yang memungkinkan pengguna untuk membangun alur kerja analisis data dengan mudah dengan menggunakan berbagai komponen visual. Komponen ini memungkinkan pengguna untuk melakukan preprocessing data, ekstraksi fitur, klasifikasi, dan evaluasi model. Orange juga menyediakan berbagai algoritma pembelajaran mesin, termasuk SVM, Logistic Regression, dan Naive Bayes, yang dapat digunakan untuk membangun model klasifikasi sentimen yang akurat.

## Pengumpulan dan Pembersihan Data

Data sentimen untuk penelitian ini dikumpulkan dari ulasan pengguna di Google Play Store dari bulan Januari – Mei 2024 sejumlah 3338 melalui proses scraping. Scraping memungkinkan pengumpulan data secara otomatis dengan menggunakan tools khusus seperti Python libraries (Beautiful Soup, Selenium). Data yang dikumpulkan kemudian dibersihkan melalui proses preprocessing untuk mempersiapkannya untuk analisis sentimen. Proses preprocessing ini meliputi beberapa langkah penting, yaitu:

- **Lowercasing:** Semua teks dalam ulasan pengguna diubah menjadi huruf kecil untuk memastikan konsistensi dan menghindari duplikasi data. Misalnya, "Bagus" dan "bagus" dianggap sebagai kata yang sama.
- **Tokenisasi:** Teks dipecah menjadi kata-kata atau frasa individual (token), yang memungkinkan sistem untuk memproses dan menganalisis setiap kata secara terpisah. Misalnya, kalimat "Aplikasi PLN Mobile sangat bagus" akan diubah menjadi token-token: "Aplikasi", "PLN", "Mobile", "sangat", "bagus".
- **Stop Word Removal:** Kata-kata yang tidak memiliki makna penting dalam analisis sentimen, seperti kata hubung (preposisi, konjungsi), dihapus. Contohnya, kata "dan", "atau", "di", "ke", "dari", dan "pada" akan dihilangkan. Hal ini membantu fokus pada kata-kata yang memiliki nilai sentimen lebih tinggi.
- **Stemming/Lemmatization:** Kata-kata diubah ke bentuk dasarnya untuk mengurangi keragaman bentuk kata yang memiliki makna sama. Stemming melibatkan pemotongan akhiran kata, sedangkan lemmatization menggunakan kamus untuk mengubah kata ke bentuk dasarnya. Misalnya, kata-kata "menjalankan", "jalankan", dan "menjalankan" akan diubah menjadi "jalan" atau "jalankan".
- **Normalisasi:** Kata-kata diubah menjadi bentuk yang seragam, seperti mengubah singkatan menjadi kata penuh. Misalnya, singkatan "dpt" diubah menjadi "dapat", dan "sya" diubah menjadi "saya".
- **Penghapusan Karakter Khusus:** Karakter khusus yang tidak relevan, seperti tanda baca yang tidak perlu, dihilangkan dari data. Misalnya, tanda seru (!), tanda tanya (?), dan tanda kutip (") dihapus.

Proses pembersihan data yang sistematis ini memastikan data yang digunakan dalam analisis sentimen bersih, terstruktur, dan siap diproses untuk menghasilkan hasil yang akurat dan reliable. Gambar 1 menunjukkan proses pembersihan data pada penelitian ini yang dilakukan melalui software Orange.

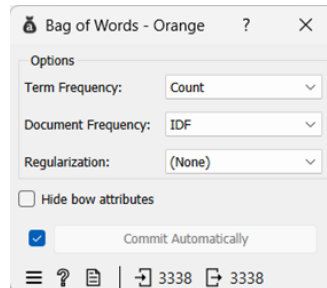


Gambar 1. Proses Pembersihan Data pada Orange

## Ekstraksi Fitur dengan TF-IDF dan Klasifikasi Sentimen

Setelah data sentimen terkumpul dan dibersihkan, langkah selanjutnya adalah ekstraksi fitur menggunakan teknik *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF membantu dalam mengidentifikasi kata-kata

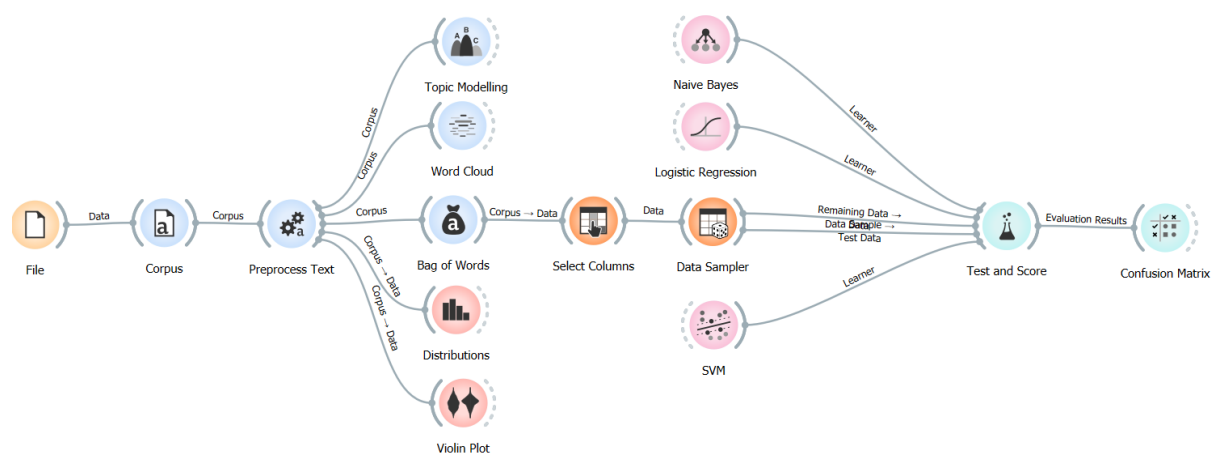
atau frasa yang memiliki nilai sentimen tinggi dalam ulasan pengguna. Metode ini bekerja dengan menghitung frekuensi kemunculan kata dalam sebuah ulasan (*Term Frequency*) dan seberapa jarang kata tersebut muncul dalam keseluruhan koleksi ulasan (*Inverse Document Frequency*). Kata-kata yang sering muncul dalam ulasan dan jarang muncul dalam keseluruhan koleksi ulasan, cenderung menjadi kata-kata yang penting dan spesifik untuk menyatakan sentimen tertentu. Nilai TF-IDF yang tinggi menunjukkan bahwa kata tersebut memiliki bobot penting dalam menentukan sentimen positif atau negatif.



Gambar 2. TF- IDF pada Orange

Fitur-fitur yang telah diekstraksi dengan TF-IDF kemudian digunakan untuk melatih model klasifikasi sentimen. Sebelum proses pelatihan, data dibagi menjadi dua bagian dengan aturan 80/20, yaitu 80% data digunakan untuk pelatihan model dan 20% sisanya digunakan untuk evaluasi model. Pembagian data ini bertujuan untuk memastikan bahwa model yang dilatih dapat menggeneralisasi dengan baik pada data yang belum pernah dilihat sebelumnya.

Dalam penelitian ini, tiga model klasifikasi yang umum digunakan akan dievaluasi, yaitu Logistic Regression, Support Vector Machine (SVM), dan Naive Bayes. Logistic Regression menggunakan fungsi sigmoid untuk memprediksi probabilitas kelas, SVM menggunakan pemisah hiperplane untuk mengklasifikasikan data, dan Naive Bayes menggunakan asumsi independensi antar fitur untuk mengklasifikasikan data. Pilihan model ini didasarkan pada keunggulan masing-masing dalam menangani data teks dan klasifikasi sentimen. Logistic Regression cocok untuk masalah klasifikasi biner dan memberikan interpretasi yang mudah, sementara SVM dikenal dengan kemampuannya untuk menggeneralisasi dan menangani data non-linear. Naive Bayes, meskipun sederhana, efektif untuk data teks dengan jumlah fitur yang besar dan merupakan pilihan yang baik untuk aplikasi dengan keterbatasan komputasi. Dengan mengevaluasi ketiga model ini, penelitian ini bertujuan untuk mengidentifikasi model terbaik yang dapat mengklasifikasikan sentimen ulasan pengguna terhadap aplikasi PLN Mobile dengan akurasi yang tinggi.



Gambar 3. Desain Model Sentimen Analisis

### Evaluasi model

Setelah model klasifikasi dilatih dengan menggunakan data pelatihan, tahap selanjutnya adalah mengevaluasi performa model menggunakan data uji. Evaluasi ini dilakukan untuk menilai seberapa baik model tersebut dapat mengklasifikasikan sentimen yang belum pernah dilihat sebelumnya. Beberapa metrik evaluasi umum digunakan untuk menilai performa model klasifikasi sentimen, yaitu:

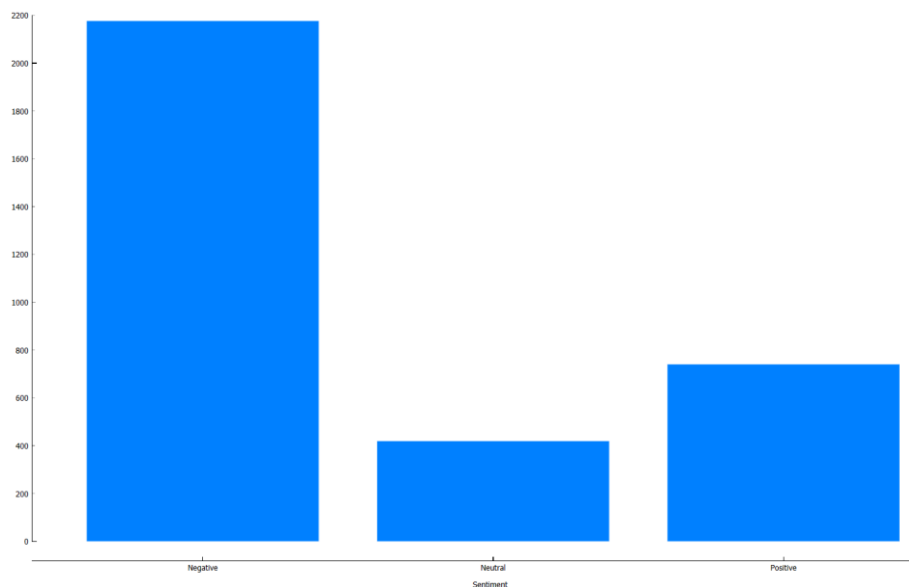
- Akurasi: Proporsi data yang diklasifikasikan dengan benar, dihitung dengan membagi jumlah data yang diklasifikasikan dengan benar dengan total jumlah data. Akurasi memberikan gambaran umum tentang ketepatan model dalam mengklasifikasikan sentimen.
- Presisi: Proporsi data positif yang diklasifikasikan dengan benar, dihitung dengan membagi jumlah data positif yang diklasifikasikan dengan benar dengan total jumlah data yang diklasifikasikan sebagai positif. Presisi memberikan gambaran tentang kemampuan model untuk menghindari kesalahan positif (false positive).
- Recall: Proporsi data positif yang ditemukan oleh model, dihitung dengan membagi jumlah data positif yang diklasifikasikan dengan benar dengan total jumlah data positif. Recall memberikan gambaran tentang kemampuan model untuk menemukan semua data positif.
- F1-score: Harmonic mean dari presisi dan recall, yang memberikan gambaran komprehensif tentang kemampuan model untuk mengklasifikasikan sentimen dengan benar.

Dalam penelitian ini, model klasifikasi Logistic Regression, SVM, dan Naive Bayes akan dievaluasi menggunakan metrik-metrik tersebut. Model dengan nilai akurasi, presisi, recall, dan F1-score yang tinggi menunjukkan kemampuan yang lebih baik dalam mengklasifikasikan sentimen. Hasil evaluasi ini akan digunakan untuk membandingkan performa ketiga model dan memilih model terbaik yang akan digunakan untuk analisis sentimen terhadap aplikasi PLN Mobile.

## Hasil dan Pembahasan

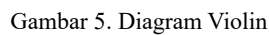
### *Statistik Deskriptif*

Data sentimen yang dikumpulkan dari ulasan pengguna di Google Play Store dan Apple Store berjumlah 3338. Data ini dibagi menjadi dua bagian dengan aturan 80/20, yaitu 80% (2671 data) untuk pelatihan model dan 20% (667 data) untuk evaluasi model. Dataset ini terdiri dari satu fitur, yaitu ulasan pengguna berupa teks, dan target variable dengan tiga kelas yang dihasilkan: negatif (2177 data, 65.22%), netral (420 data, 12.58%), dan positif (741 data, 22.20%). Analisis deskriptif menunjukkan bahwa sentimen negatif mendominasi dengan proporsi 65.22%, diikuti oleh sentimen positif (22.20%) dan netral (12.58%). Dominasi sentimen negatif mengindikasikan bahwa sebagian besar pengguna memiliki pengalaman negatif dengan aplikasi PLN Mobile. Distribusi sentiment dapat dilihat pada Gambar 4.



Gambar 4. Histogram Target Variable

Diagram violin, pada Gambar 5, yang menggambarkan distribusi data berdasarkan nilai median, menunjukkan bahwa mayoritas ulasan negatif (65.22%) memiliki rating bintang antara 1 dan 2, dengan beberapa ulasan mencapai rating 3. Sebaliknya, ulasan positif (22.20%) cenderung memiliki rating bintang yang lebih tinggi, meskipun distribusi data lebih merata, mengindikasikan variabilitas rating yang lebih tinggi. Ulasan netral (12.58%) memiliki distribusi yang lebih sempit dan cenderung terpusat pada rating bintang 3.

[illegible]

Gambar 6. World Cloud

Topic	Topic keywords
1	aplikasi, token, listrik, sdh, data, gagal, pln, udah, terbaru, jam
2	a, listrik, pln, aplikasi, pengaduan, mati, beli, token, petugas, jam
3	susah, bagus, email, aplikasi, buka, pelayanan, buruk, aplikasinya, verifikasi, error
4	aplikasi, listrik, pengaduan, pln, token, pelanggan, kwh, pembayaran, sisa, beli
5	aplikasi, pln, tagihan, bayar, tgl, lemot, mobile, meter, i, nggk
6	mati, lampu, aplikasi, pln, bayar, listrik, beli, hp, root, gangguan
7	aplikasi, pln, listrik, udah, pengaduan, padam, pemadaman, update, error, masuk
8	pln, padam, listrik, mati, pelayanan, lampu, cepat, mei, jam, terbaik
9	pln, mobile, aplikasi, meteran, token, listrik, laporan, tdk, mudah, kwh
10	aplikasi, pln, listrik, masuk, rumah, lampu, lambat, mati, petugas, jam

Gambar 7. Topic Modelling

*Topic Modelling*, adalah teknik di bidang pembelajaran mesin dan pemrosesan bahasa alami (NLP) yang digunakan untuk mengidentifikasi topik-topik yang terdapat dalam sekumpulan dokumen. Teknik ini secara otomatis menganalisis teks untuk menemukan kelompok kata (disebut "topik") yang sering muncul bersama dan untuk memahami struktur tersembunyi dalam data teks. Dua algoritma yang paling umum digunakan untuk topic modeling adalah Latent Dirichlet Allocation (LDA) dan Non-Negative Matrix Factorization (NMF). Berikut adalah penjelasan pada masing-masing topik dari Gambar 7.

Topik 1:

Kata kunci: aplikasi, token, listrik, sdh, data, gagal, pln, udah, terbaru, jam.

Penjelasan: Topik ini mungkin berkaitan dengan masalah yang sering dihadapi pengguna aplikasi PLN terkait dengan token listrik, seperti kegagalan data atau proses yang sudah dilakukan namun tidak berhasil.

Topik 2:

Kata kunci: aplikasi, pln, aplikasi, pengaduan, mati, beli, token, petugas, jam

Penjelasan: Topik ini fokus pada keluhan atau pengaduan pengguna terkait aplikasi PLN, terutama mengenai listrik yang mati, pembelian token yang gagal, dan kebutuhan untuk menghubungi petugas PLN.

Topik 3:

Kata kunci: susah, bagus, email, aplikasi, buka, pelayanan, buruk, aplikasinya, verifikasi, error

Penjelasan: Topik ini mencerminkan kesulitan pengguna dalam menggunakan aplikasi, seperti verifikasi yang error atau pelayanan yang buruk.

Topik 4:

Kata kunci: aplikasi, listrik, pengaduan, pln, token, pelanggan, kwh, pembayaran, sisa, beli

Penjelasan: Topik ini mencakup pengaduan pelanggan terkait penggunaan listrik, pembelian token, dan pembayaran kwh.

Topik 6:

Kata kunci: mati, lampu, aplikasi, pln, bayar, listrik, beli, hp, root, gangguan

Penjelasan: Topik ini berhubungan dengan gangguan listrik (lampu mati), masalah pembayaran, serta keluhan yang disampaikan melalui aplikasi PLN.

Topik 7:

Kata kunci: aplikasi, pln, listrik, udah, pengaduan, padam, pemadaman, update, error, masuk

Penjelasan: Topik ini mencakup pengaduan tentang listrik yang padam, masalah masuk ke aplikasi, serta error dan update aplikasi.

Topik 8:

Kata kunci: pln, padam, listrik, mati, pelayanan, lampu, cepat, mei, jam, terbaik

Penjelasan: Topik ini terkait dengan keluhan tentang listrik padam, pelayanan PLN, serta penanganan gangguan listrik yang diharapkan cepat.

Topik 9:

Kata kunci: pln, mobile, aplikasi, meteran, token, listrik, laporan, tdk, mudah, kwh

Penjelasan: Topik ini berfokus pada penggunaan aplikasi mobile PLN, masalah dengan meteran dan token listrik, serta laporan yang tidak mudah dilakukan oleh pengguna.

Topik 10:

Kata kunci: aplikasi, pln, listrik, masuk, rumah, lampu, lambat, mati, petugas, jam

Penjelasan: Topik ini menggambarkan keluhan pengguna tentang kesulitan masuk ke aplikasi, masalah listrik di rumah, serta layanan petugas yang lambat dalam menanggapi keluhan tentang listrik mati.

### *Perbandingan Performa Algoritma Klasifikasi*

Penelitian ini menggunakan tiga algoritma klasifikasi sentimen: Logistic Regression, Naive Bayes, dan Support Vector Machine (SVM) untuk mengklasifikasikan sentimen pengguna terhadap aplikasi PLN Mobile. Logistic



Regression menggunakan fungsi sigmoid untuk memprediksi probabilitas kelas, model ini cocok untuk masalah klasifikasi biner dan memberikan interpretasi yang mudah. Naive Bayes menggunakan asumsi independensi antar fitur untuk mengklasifikasikan data, model ini sederhana dan efektif untuk data teks dengan jumlah fitur yang besar, serta cocok untuk aplikasi dengan keterbatasan komputasi. Support Vector Machine (SVM) menggunakan pemisah hiperplane untuk mengklasifikasikan data, model SVM dikenal dengan kemampuannya untuk menggeneralisasi dan menangani data non-linear.

Sebelum proses pelatihan, data yang dikumpulkan dibagi menjadi dua bagian dengan aturan 80/20, yaitu 80% untuk pelatihan model dan 20% untuk evaluasi model. Pembagian data ini bertujuan untuk memastikan bahwa model yang dilatih dapat menggeneralisasi dengan baik pada data yang belum pernah dilihat sebelumnya. Model dilatih dengan menggunakan data pelatihan dan kemudian dievaluasi menggunakan data uji untuk mengukur performanya. Hasil evaluasi yang disajikan dalam tabel sebelumnya didasarkan pada data uji sebesar 20% yang belum pernah digunakan dalam proses pelatihan. Hal ini memungkinkan peneliti untuk menilai kemampuan model dalam mengklasifikasikan data baru dan mengukur generalisasinya.

Model	AUC	CA	F1	Prec	Recall	MCC
Logistic Regression	0.981	0.918	0.905	0.913	0.918	0.833
Naive Bayes	0.947	0.619	0.652	0.875	0.619	0.554
SVM	0.963	0.879	0.831	0.799	0.879	0.772

Gambar 8. Performa Metode dalam Klasifikasi

Hasil evaluasi pada Gambar 8 menunjukkan bahwa Logistic Regression memiliki performa terbaik dibandingkan dengan Naive Bayes dan SVM. Hal ini terlihat dari nilai AUC, CA, F1, Prec, Recall, dan MCC yang lebih tinggi. AUC, yang menunjukkan kemampuan model dalam membedakan antara kelas positif dan negatif, mencapai nilai 0.981 pada Logistic Regression, mengindikasikan kemampuan yang sangat baik dalam membedakan kelas. Akurasi (CA) Logistic Regression mencapai 0.918, menunjukkan bahwa model ini dapat mengklasifikasikan data dengan benar. Nilai F1-score, yang menunjukkan keseimbangan antara presisi dan recall, mencapai 0.905, menunjukkan bahwa model memiliki keseimbangan yang baik antara presisi dan recall. Nilai presisi yang lebih tinggi (0.913) menunjukkan kemampuan model dalam menghindari kesalahan positif, sedangkan recall yang lebih tinggi (0.918) menunjukkan bahwa model dapat menemukan sebagian besar data positif. Matthews Correlation Coefficient (MCC) yang mencapai 0.833 menunjukkan bahwa model memiliki korelasi yang kuat dengan hasil aktual. Secara keseluruhan, perbandingan performa ketiga algoritma klasifikasi menunjukkan bahwa Logistic Regression memiliki performa terbaik dalam mengklasifikasikan sentimen ulasan pengguna terhadap aplikasi PLN Mobile.

Selain itu, confusion matrix pada Gambar 9 memberikan gambaran rinci tentang hasil klasifikasi model, menunjukkan berapa banyak data yang diklasifikasikan dengan benar dan salah. Pada penelitian ini, confusion matrix dianalisis untuk ketiga model klasifikasi: Logistic Regression, SVM, dan Naive Bayes.

		Predicted			Σ
		Negative	Neutral	Positive	
Actual	Negative	431.2	12.3	0.5	444
	Neutral	24.7	27.3	28.0	80
	Positive	0.7	16.0	126.3	143
Σ		457	56	155	667
Logistics Regression					

		Predicted			Σ
		Negative	Neutral	Positive	
Actual	Negative	144.5	158.5	141.0	444
	Neutral	0.9	77.9	1.3	80
	Positive	0.0	21.9	121.0	143
Σ		145	258	263	667
SVM					

		Predicted			Σ
		Negative	Neutral	Positive	
Actual	Negative	180.2	211.6	52.2	444
	Neutral	0.0	77.7	2.2	80
	Positive	0.0	12.6	130.4	143
Σ		180	302	185	667
Naïve Bayes					

Gambar 9. Confusion Matrix Tiga Metode Klasifikasi

Logistic Regression menunjukkan performa yang baik dalam mengklasifikasikan sentimen negatif dan positif, dengan kesalahan klasifikasi yang relatif kecil. Model ini berhasil mengklasifikasikan 431.2 data negatif dengan benar, 27.3 data netral dengan benar, dan 126.3 data positif dengan benar. Namun, model ini cenderung mengalami kesulitan dalam mengklasifikasikan sentimen netral dengan tepat, seperti terlihat dari kesalahan klasifikasi 12.3 data negatif sebagai netral, 0.5 data negatif sebagai positif, 24.7 data netral sebagai negatif, 28 data netral sebagai positif, 0.7 data positif sebagai negatif, dan 16 data positif sebagai netral.

SVM menunjukkan performa yang kurang optimal dibandingkan dengan Logistic Regression. Model ini berhasil mengklasifikasikan 144.5 data negatif dengan benar, 77.9 data netral dengan benar, dan 121 data positif dengan



benar. Namun, model ini cenderung salah mengklasifikasikan data negatif sebagai netral dan positif, seperti terlihat dari kesalahan klasifikasi 158.5 data negatif sebagai netral, 141 data negatif sebagai positif, 0.9 data netral sebagai negatif, 1.3 data netral sebagai positif, 21.9 data positif sebagai netral.

Naive Bayes menunjukkan performa yang lebih rendah dibandingkan dengan Logistic Regression dan SVM. Model ini berhasil mengklasifikasikan 180.2 data negatif dengan benar, 77.7 data netral dengan benar, dan 130.4 data positif dengan benar. Namun, model ini cenderung salah mengklasifikasikan data negatif sebagai netral dan positif, seperti terlihat dari kesalahan klasifikasi 211.6 data negatif sebagai netral, 52.2 data negatif sebagai positif, 2.2 data netral sebagai positif, 12.6 data positif sebagai netral.

Secara keseluruhan, confusion matrix menunjukkan bahwa Logistic Regression memiliki performa yang lebih baik dalam mengklasifikasikan sentimen, terutama untuk data negatif dan positif, dibandingkan dengan SVM dan Naive Bayes. Hasil ini mendukung kesimpulan sebelumnya bahwa Logistic Regression adalah model yang paling efektif untuk mengklasifikasikan sentimen ulasan pengguna terhadap aplikasi PLN Mobile.

#### *Interpretasi hasil Analisis Sentimen Aplikasi PLN Mobile*

Analisis sentimen terhadap aplikasi PLN Mobile mengungkapkan berbagai permasalahan yang dihadapi oleh pengguna dan memberikan gambaran tentang persepsi mereka terhadap layanan yang diberikan. Analisis deskriptif menunjukkan dominasi sentimen negatif (65.22%), yang mengindikasikan bahwa mayoritas pengguna mengalami ketidakpuasan dengan aplikasi tersebut. Diagram violin memperlihatkan bahwa ulasan negatif cenderung memiliki rating bintang yang lebih rendah (1-2) dibandingkan dengan ulasan positif yang cenderung memiliki rating lebih tinggi.

Dominasi sentimen negatif ini diperkuat oleh hasil analisis word cloud yang menunjukkan kemunculan kata-kata seperti "aplikasi", "pln", "listrik", "pengaduan", dan "mati". Dominasi kata "pengaduan" dan "mati" menunjukkan bahwa pengguna sering kali mengalami masalah dengan layanan kelistrikan dan mengajukan keluhan melalui aplikasi.

Penelitian ini kemudian membandingkan performa tiga algoritma klasifikasi: Logistic Regression, Naive Bayes, dan SVM. Hasil evaluasi menunjukkan bahwa Logistic Regression memiliki performa yang paling baik dalam mengklasifikasikan sentimen, dengan akurasi, presisi, recall, dan F1-score yang lebih tinggi dibandingkan dengan dua model lainnya. Hal ini mengindikasikan bahwa Logistic Regression lebih mampu mengidentifikasi sentimen secara akurat dan meminimalkan kesalahan klasifikasi.

Secara keseluruhan, hasil analisis sentimen ini mengungkap beberapa isu penting yang dihadapi oleh PT PLN dalam menyediakan layanan digital melalui aplikasi PLN Mobile:

- **Tingkat Kepuasan yang Rendah:** Dominasi sentimen negatif menunjukkan bahwa pengguna aplikasi PLN Mobile tidak merasa puas dengan layanan yang diberikan.
- **Masalah Layanan Kelistrikan:** Dominasi kata "pengaduan" dan "mati" dalam word cloud mengindikasikan bahwa pengguna seringkali mengalami masalah dengan layanan kelistrikan dan memerlukan penyelesaian melalui aplikasi.
- **Pentingnya Responsivitas Aplikasi:** Hasil analisis menunjukkan bahwa pengguna mengharapkan aplikasi PLN Mobile dapat memberikan solusi yang efektif dan responsif terhadap masalah yang mereka hadapi.

Hasil penelitian ini dapat digunakan sebagai panduan bagi PT PLN untuk meningkatkan kualitas aplikasi PLN Mobile. Peningkatan fitur aplikasi, peningkatan responsivitas layanan pelanggan, dan penyelesaian masalah yang dihadapi pengguna, merupakan beberapa langkah penting yang dapat diambil oleh PT PLN untuk meningkatkan kepuasan pelanggan dan membangun citra positif bagi aplikasi PLN Mobile.

#### **Kesimpulan**

Penelitian ini bertujuan untuk menganalisis sentimen pelanggan terhadap aplikasi PLN Mobile dengan menggunakan teknik analisis sentimen berbasis TF-IDF. Analisis deskriptif menunjukkan bahwa sebagian besar ulasan pengguna terhadap aplikasi PLN Mobile bernada negatif (65.22%), diikuti oleh sentimen positif (22.20%) dan netral (12.58%). Diagram violin memperlihatkan bahwa ulasan negatif cenderung memiliki rating bintang yang lebih rendah (1-2), sedangkan ulasan positif memiliki rating yang lebih tinggi. Analisis word cloud menunjukkan bahwa kata-kata "aplikasi", "pln", "listrik", "pengaduan", dan "mati" mendominasi, menunjukkan bahwa pengguna sering kali mengalami masalah dan mengajukan pengaduan terkait dengan layanan kelistrikan.

Penelitian ini mengevaluasi tiga model klasifikasi sentimen: Logistic Regression, Naive Bayes, dan Support Vector Machine (SVM). Hasil evaluasi menunjukkan bahwa Logistic Regression memiliki performa terbaik dengan nilai akurasi, presisi, recall, dan F1-score yang lebih tinggi dibandingkan dengan Naive Bayes dan SVM. Model ini berhasil mengklasifikasikan sebagian besar data dengan tepat, dengan kesalahan klasifikasi yang minimal, terutama untuk sentimen negatif dan positif.

Secara keseluruhan, penelitian ini menunjukkan bahwa pengguna aplikasi PLN Mobile memiliki pengalaman negatif yang signifikan, dengan banyak keluhan yang diutarakan dalam ulasan. Analisis sentimen menggunakan model Logistic Regression menunjukkan kemampuan yang baik dalam mengidentifikasi sentimen pengguna dan mengungkap aspek-aspek yang perlu diperbaiki dalam aplikasi PLN Mobile. Hasil penelitian ini dapat memberikan informasi yang berharga bagi PT PLN untuk meningkatkan kualitas layanan dan responsivitas aplikasi PLN Mobile, serta meningkatkan kepuasan pelanggan.

#### Daftar Pustaka

- Bafna, P., Pramod, D., & Vaidya, A. (2016). Document clustering: TF-IDF approach. *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, 61–66. <https://ieeexplore.ieee.org/abstract/document/7754750/>
- Gupta, P., Mehlawat, M. K., Khaitan, A., & Pedrycz, W. (2020). Sentiment analysis for driver selection in fuzzy capacitated vehicle routing problem with simultaneous pick-up and drop in shared transportation. *IEEE Transactions on Fuzzy Systems*, 29(5), 1198–1211. <https://ieeexplore.ieee.org/abstract/document/8977391/>
- Kavitha, M., Naib, B. B., Mallikarjuna, B., Kavitha, R., & Srinivasan, R. (2022). Sentiment Analysis using NLP and Machine Learning Techniques on Social Media Data. *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 112–115. <https://ieeexplore.ieee.org/abstract/document/9823708/>
- Kim, S.-B., Rim, H.-C., Yook, D., & Lim, H.-S. (2002). Effective Methods for Improving Naive Bayes Text Classifiers. In M. Ishizuka & A. Sattar (Eds.), *PRICAI 2002: Trends in Artificial Intelligence* (Vol. 2417, pp. 414–423). Springer Berlin Heidelberg. [https://doi.org/10.1007/3-540-45683-X\\_45](https://doi.org/10.1007/3-540-45683-X_45)
- Li, Y., Bontcheva, K., & Cunningham, H. (2005). SVM Based Learning System for Information Extraction. In J. Winkler, M. Niranjan, & N. Lawrence (Eds.), *Deterministic and Statistical Methods in Machine Learning* (Vol. 3635, pp. 319–339). Springer Berlin Heidelberg. [https://doi.org/10.1007/11559887\\_19](https://doi.org/10.1007/11559887_19)
- Sinaga, R. B., Fajri, H. R. A., Mubarak, H., Pangestu, A. D., & Prasvita, D. S. (2021). Analisis Sentimen Pengguna Twitter terhadap Konflik antara Palestina dan Israel Menggunakan Metode Naïve Bayesian Classification dan Support Vector Machine. *Prosiding Seminar Nasional Mahasiswa Bidang Ilmu Komputer dan Aplikasinya*, 2(2), Article 2. <https://conference.upnvj.ac.id/index.php/senamika/article/view/1632>