

ANALISIS SENTIMEN MASYARAKAT TERHADAP PELAYANAN JASA EKSPEDISI JNE DAN J&T EXPRESS MENGGUNAKAN METODE LEXICON-BASED

Sebastianus Adi Santoso Mola[✉], Dinda Permata Mbatu, Dony Martinus Sihotang

Fakultas Sains dan Teknik, Universitas Nusa Cendana, Kota Kupang, Indonesia

Email: adimola@staf.undana.ac.id

DOI: <https://doi.org/10.46880/jmika.Vol9No1.pp56-65>

ABSTRACT

JNE and J&T Express are two of the largest and most popular courier companies in Indonesia, leading to various public opinions regarding the quality of their services. This research employs a lexicon-based method using the InSet dictionary, a simple scientific approach where the system calculates the weight of words and classifies them as positive, negative, or neutral sentiments. The analysis process begins with data collection of reviews using scraping techniques, followed by text processing including cleaning, case folding, normalization, tokenization, stemming, and stopword removal. Out of 3,565 reviews for JNE and 3,967 reviews for J&T, the sentiment analysis indicates that the majority of the public holds negative opinions towards the services of both courier companies. The analysis accuracy reaches 82% for JNE data, with a precision value of 95% for negative sentiment, 54% for positive sentiment, and 7% for neutral sentiment. The sensitivity values are 83% for negative sentiment, 82% for positive sentiment, and 15% for neutral sentiment. Data for J&T shows an accuracy of 78%, with a precision value of 97% for negative sentiment, 28% for positive sentiment, and 4% for neutral sentiment. Sensitivity values are 80% for negative sentiment, 82% for positive sentiment, and 4% for neutral sentiment.

Keyword: JNE, J&T, Sentiment Analysis, Lexicon-Based, InSet.

ABSTRAK

JNE dan J&T Express adalah dua perusahaan pengiriman terbesar dan terpopuler di Indonesia, mengakibatkan banyak opini masyarakat mengenai kualitas layanan mereka. Metode lexicon-based dengan kamus InSet digunakan dalam penelitian ini, sebuah metode ilmiah sederhana di mana sistem menghitung bobot kata dan mengklasifikasikannya sebagai sentimen positif, negatif, dan netral. Proses analisis dimulai dengan pengumpulan data ulasan menggunakan teknik scraping, diikuti oleh text processing yang meliputi cleaning, case folding, normalisasi, tokenizing, stemming, dan stopword removal. Dari 3565 ulasan JNE dan 3967 ulasan J&T, hasil sentimen menunjukkan mayoritas masyarakat beropini negatif terhadap pelayanan kedua jasa ekspedisi tersebut. Akurasi analisis mencapai 82% untuk data JNE dengan nilai presisi sentimen negatif 95%, sentimen positif 54%, sentimen netral 7% sedangkan sensitivitas sentimen negatif 83%, sentimen positif 82%, sentimen netral 15%. Data J&T memiliki akurasi sebesar 78% dengan nilai presisi sentimen negatif 97%, sentimen positif 28%, sentimen netral 4% sedangkan sensitivitas sentimen negatif 80%, sentimen positif 82%, sentimen netral 4%.

Kata Kunci: JNE, J&T, Analisis Sentimen, Lexicon-Based, InSet.

PENDAHULUAN

Kebutuhan akan penggunaan jasa ekspedisi barang meningkat dengan cepat dalam memenuhi berbagai kebutuhan seluruh masyarakat Indonesia. Kehadiran berbagai layanan ekspedisi tidak hanya memberikan kemudahan bagi masyarakat, tetapi juga bagi pengusaha atau penjual dimana dalam bisnis online, mereka memanfaatkan jasa ekspedisi untuk mengirimkan barang ke pembeli yang berada di berbagai daerah di Indonesia (Nitami & Februriyanti, 2022). Hal yang paling utama bagi sebuah perusahaan jasa ekspedisi barang adalah kemampuan perusahaan

tersebut dalam membangun hubungan yang harmonis dengan penggunanya (Akbar et al., 2022). Meningkatkan kepuasan pengguna dengan memberikan pelayanan yang ramah dan berkualitas adalah salah satu cara untuk membantu perusahaan mencapai target laba yang optimal (Affandi et al., 2022). Jika layanan yang diberikan kurang memuaskan bagi pengguna atau masyarakat, hal itu dapat membuat mereka mempertimbangkan kembali untuk menggunakan jasa ekspedisi tersebut di masa mendatang. Hal ini juga menjadi bahan pertimbangan oleh masyarakat dalam menggunakan jasa ekspedisi.

Terdapat beberapa perusahaan jasa pengiriman terbesar dan terpopuler di Indonesia, yakni **JNE dan J&T Express**. PT Tiki Jalur Nugraha Ekakurir, atau dikenal sebagai JNE, merupakan salah satu perusahaan yang bergerak dalam layanan distribusi barang di Indonesia. JNE menyediakan pengiriman paket dan dokumen dengan lebih dari 8.000 titik layanan di seluruh Indonesia. Pada tahun 2014, JNE memperkenalkan aplikasi MY JNE, yang bertujuan membantu pelanggan untuk mengecek tarif, melacak paket, menemukan lokasi konter terdekat, serta memfasilitasi transaksi jual-beli antara penjual dan pembeli individual (JNE, 2024). Sementara itu, PT Global Jet Express, atau J&T, adalah perusahaan jasa pengiriman yang memiliki 4.000 titik operasi tersebar di berbagai wilayah Indonesia. Kehadiran titik operasi ini bertujuan mendukung distribusi barang pelanggan. Pada tahun 2015, J&T meluncurkan aplikasi J&T Express untuk memberikan kemudahan kepada pengguna dalam memesan, melacak paket, menghitung ongkos kirim, serta menemukan drop point (J&T, 2024). Berdasarkan hasil Google Trends 2024, JNE dan J&T menempati peringkat pertama dan kedua sebagai jasa pengiriman yang sering dicari di Indonesia dalam dua belas bulan terakhir, dibandingkan dengan jasa pengiriman lain seperti Anteraja, Sicepat, dan Pos Indonesia. Hal ini menunjukkan bahwa JNE dan J&T menjadi dua perusahaan jasa pengiriman yang cukup populer di mesin pencari Google di wilayah Indonesia.

Popularitas JNE dan J&T Express sebagai penyedia jasa pengiriman memiliki dampak signifikan terhadap opini masyarakat mengenai kualitas layanan mereka. Sejumlah besar ulasan di *google playstore* mencerminkan tingkat kepuasan atau kekecewaan masyarakat terhadap layanan maupun aplikasi masing-masing penyedia jasa pengiriman tersebut. Opini masyarakat ini memiliki potensi besar untuk mempengaruhi reputasi JNE dan J&T Express (Sitorus, 2021). Mengingat dampak signifikan yang dapat ditimbulkan oleh berbagai opini, penting untuk segera mengambil langkah guna menganalisis opini masyarakat terhadap layanan yang diberikan oleh kedua perusahaan ini (Rahmadan & Marpaung, 2024). Langkah ini sangat krusial untuk mengidentifikasi masalah dan mencari solusi yang efektif guna meningkatkan kualitas layanan dan memperbaiki pengalaman masyarakat. Salah satu cara yang dapat digunakan dalam menganalisis opini masyarakat adalah analisis sentimen.

Analisis sentimen merupakan studi komputasional dari opini, sentimen dan emosi yang diekspresikan dalam bentuk teks, yang termasuk dalam

domain penelitian *text mining* dan mulai berkembang sejak tahun 2013 (Alita & Isnain, 2020). Kecenderungan penelitian tentang analisis sentimen umumnya berfokus pada opini yang mengungkapkan atau mencerminkan sentimen positif, negatif atau netral (Kurniawan, 2017).

Untuk melakukan analisis sentimen jasa pengiriman **JNE dan J&T Express**, diperlukan sejumlah data ulasan dari masyarakat. Informasi ini diambil dari *google play store*, di mana ulasan pengguna menjadi salah satu fitur ulasan yang dapat diakses. Fitur ulasan tersebut kerap digunakan sebagai alat untuk memperoleh informasi mengenai suatu produk atau layanan sebelum digunakan (Rezki et al., 2020).

Dari masalah yang telah disampaikan sebelumnya, penulis memilih untuk menggunakan *text mining* dan menerapkan metode *lexicon based* dalam pembuatan model analisis sentimen. Metode *lexicon based* merupakan pendekatan yang sederhana tanpa harus melakukan pelatihan data set (Singh et al., 2018). Pendekatan ini relatif mudah diterapkan dan praktis, sehingga menjadi pilihan yang baik untuk menganalisis sentimen dari ulasan atau komentar (Fatihin, 2022).

TINJAUAN PUSTAKA

Dari sejumlah penelitian yang telah dilakukan, terdapat temuan yang konsisten mengenai penggunaan metode *lexicon based* dalam analisis sentimen. Penelitian oleh Mujilawati (2024) menunjukkan bahwa metode ini mampu mengklasifikasikan komentar dalam kelas positif, negatif, dan netral berdasarkan nilai *polarity*. Begitu pula dengan penelitian oleh Arief & Imanuel (2019), Ismail & Hakim (2023), serta Sumitro dkk (2021), yang menghasilkan temuan serupa mengenai dominasi sentimen positif dalam konteks yang berbeda. Hasil-hasil ini mendukung kesimpulan bahwa metode *lexicon based* dapat digunakan untuk menganalisis sentimen dalam berbagai konteks, seperti penilaian terhadap wisata, vaksin, atau topik viral.

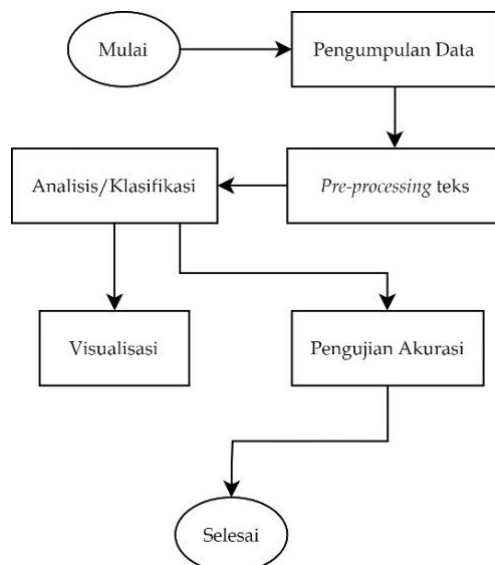
Penelitian yang dilakukan oleh Syaifuddin & Muslimin (2022) dan Musfiroh dkk (2021) juga menunjukkan hasil serupa dalam konteks kebijakan dan perkuliahan daring, dengan mayoritas tweet mengandung sentimen negatif. Temuan ini menunjukkan bahwa isu-isu tertentu, seperti pendidikan online atau kebijakan tertentu, dapat menimbulkan reaksi negatif dari masyarakat.

Di sisi lain, penelitian oleh Undap dkk (2021) dan Hernikawati (2021) menyoroti variasi dalam opini masyarakat terhadap topik yang berbeda, seperti

pembajakan artikel penelitian dan vaksin *Sinovac*, dengan dominasi opini netral. Begitu pula dengan penelitian oleh Fadhillah dkk (2019), yang menunjukkan variasi dalam respons masyarakat terhadap partai politik peserta pemilu 2019. Dengan demikian, hasil-hasil penelitian ini memberikan kontribusi yang berharga dalam pemahaman tentang bagaimana sentimen masyarakat dapat dianalisis dan dipahami melalui pendekatan *lexicon Based*.

METODE PENELITIAN

Tahap pertama dalam penelitian ini adalah pengumpulan dan pengambilan data, diikuti oleh tahap *preprocessing* teks, analisis teks atau klasifikasi sentimen menggunakan metode berbasis leksikon (*lexicon-based*), visualisasi hasil klasifikasi, serta pengujian akurasi dengan *confusion matrix*. Tahapan penelitian dapat dilihat pada Gambar 1.



Gambar 1 Diagram Alur Penelitian

Berdasarkan gambar diatas, alur atau tahap penelitian adalah sebagai berikut.

Tahap 1: Dataset dalam penelitian ini diperoleh dari situs *google play store* menggunakan teknik *web scraping* dengan program *python*. Data yang terkumpul kemudian akan dianalisis lebih lanjut untuk memahami sentimen masyarakat terhadap masing-masing layanan.

Tahap 2: Tahap *pre-processing* merupakan langkah pertama dalam *text mining* di mana data dibersihkan dari kata-kata yang dianggap tidak relevan, sehingga proses selanjutnya dapat menghasilkan hasil yang lebih optimal (Basryah et al., 2021). Text processing dalam penelitian ini meliputi *cleaning*, *case folding*,

normalisasi, *tokenizing*, *stemming*, dan *stopword removal*.

Tahap 3: Setelah dilakukannya *preprocessing*, selanjutnya adalah melakukan klasifikasi menggunakan metode berbasis leksikon (*lexicon-based*). Metode *lexicon-based* merupakan pendekatan ilmiah yang sering dipakai dalam penelitian analisis sentimen. Pendekatan ini melibatkan penggunaan kamus kata yang telah diberi bobot pada setiap katanya sebagai sumber leksikal atau bahasa dan termasuk dalam kategori *machine learning* yang bersifat *unsupervised*. Tingkat keberhasilan tergantung pada kualitas kamus kata yang digunakan. Menggunakan metode ini relatif sederhana dimana sistem akan menghitung bobot kata-kata yang ditemukan dan mengklasifikasikannya ke dalam sentimen tertentu (Fadhillah et al., 2019). Metode *lexicon based* melibatkan penghitungan orientasi dokumen dari orientasi semantik kata atau frasa dalam dokumen (Turney, 2001). Dengan demikian dalam penelitian ini dibutuhkan *lexicon* (kamus) yang dimana peneliti menggunakan kamus kata leksikon berbahasa Indonesia yaitu *InSet* yang dibangun oleh Koto & Rahmaningtyas, kamus ini dianotasi dengan orientasi semantik, atau polaritas kata digunakan untuk klasifikasi sentimen bahasa Indonesia. Dalam kamus *InSet* sudah memiliki bobot atau skor masing-masing pada tiap kata mulai dari -5 untuk sangat negatif sampai dengan +5 untuk sangat positif. (Taboada et al., 2011).

Dengan demikian, pendekatan ini menekankan pentingnya menganalisis setiap komponen kata atau frasa untuk memahami sentimen keseluruhan suatu teks. Sentimen pada data teks ditentukan berdasarkan kata-kata yang ada dalam kamus *lexicon*, yang dikategorikan sebagai negatif, positif, atau netral. Skor dihitung dari jumlah kata yang teridentifikasi pada setiap teks atau kalimat. (Ismail & Hakim, 2023). Maka, untuk klasifikasi sentimen dibutuhkan *Polarity score* yang dilakukan dengan cara sebagai berikut (Fadhillah et al., 2019) :

a. Menentukan skor sentimen setiap kata :

Pada tahap ini, setiap kata diberi skor sentimen berdasarkan *InSet* leksikon, skor ini kemudian digunakan untuk menentukan nilai *polarity* (Rajput et al., 2016).

b. Menghitung nilai *Polarity score* :

Perhitungan *Polarity Score* dilakukan dengan menjumlahkan setiap skor sentimen setiap kata dalam satu kalimat menggunakan rumus pada

Persamaan 1 untuk menentukan orientasi akhir dari sentimen (Yerzi & Sibaroni, 2021).

Keterangan :

$\sum S_{score}$ = Jumlah nilai skor

$\sum Nw$ = Jumlah kata

c. Menentukan orientasi sentimen :

Setelah nilai polaritas ditentukan, langkah selanjutnya adalah memberi label atau mengkategorikan kelas sentimen dari setiap kalimat atau data. Sentimen diklasifikasikan menurut persamaan berikut (Nurkasanah & Hayaty, 2022) :

$$Sentiment = \begin{cases} positif & \text{if } x > 0 \\ netral & \text{if } x = 0 \\ negatif & \text{if } x < 0 \end{cases} \quad (2)$$

Keterangan :

x = *Polarity score*

Tahap 4: Hasil klasifikasi yang diperoleh akan disajikan dalam bentuk visualisasi data dengan menggunakan *pie chart* histogram dan *word cloud*. Untuk memudahkan pemahaman hasil proses klasifikasi, data disajikan dalam bentuk *pie chart*. Histogram menunjukkan persentase untuk setiap kelas polaritas, sedangkan wordcloud menampilkan kata-kata yang paling sering muncul pada masing-masing sentimen (Amaliah & Dwi Nuryana, 2022; Musfiroh et al., 2021).

Tahap 5: Pada tahap pengujian akurasi, *Confusion Matrix* digambarkan dalam Tabel 1 terdapat jumlah data uji yang diklasifikasi dengan benar dan salah oleh model klasifikasi (Normawati & Prayogi, 2021).

Tabel 1. *Confusion matrix*

	<i>Predicted Class</i>			
	<i>Class</i>	Positif	Negatif	Netral
<i>Actual Class</i>	Positif	TPP	PFN	PFNet
	Negatif	NFP	TNN	NegFNet
	Netral	NetFP	NetFN	TNetNet

Keterangan:

1. TPP (*True Possitive Possitive*) : Kelas aktual dan prediksi positif.
2. TNN (*True Negative Negative*) : Kelas aktual dan prediksi negatif.
3. TNetNet (*True Netral Netral*) : Kelas aktual dan prediksi netral.
4. PFN (*Positive False Negatif*) : Aktual positif, prediksi negatif.
5. NFP (*Negatif False Positiv*) : Aktual negatif, prediksi positif.

6. PFNet (*Positive False Netral*) : Aktual positif, prediksi netral.

$$PolarityScore = \frac{\sum S_{score}}{\sum Nw} \quad (1)$$

7. NetFP (*Netral False Positive*) : Aktual netral, prediksi positif.
8. NetFN (*Netral False Negatif*) : Aktual netral, prediksi negatif.
9. NegFNet (*Negatif False Netral*) : Aktual negatif, prediksi netral.

Pengujian dalam penelitian ini dilakukan dengan cara menampilkan *classification report* yaitu menghitung *accuracy* (akurasi), *precision* (Presisi) dan *recall* (sensitivitas):

a. Akurasi

Akurasi menunjukkan perhitungan persentase jumlah data yang diklasifikasikan dengan benar, dengan cara data prediksi yang terklasifikasi dengan benar, dibagi dengan jumlah keseluruhan data prediksi. Dimana rumus dari perhitungan akurasi dapat dilihat pada Persamaan 3.

$$Akurasi = \frac{TPP + TNegNeg + TNetNet}{Total} \times 100 \quad (3)$$

b. Presisi

Presisi menunjukkan perhitungan tingkat ketepatan hasil klasifikasi kelas aktual yang diprediksi dengan benar oleh model, dengan cara jumlah data aktual yang diklasifikasi dengan benar (*True Possitive Possitive*), dibagi dengan jumlah data yang diklasifikasikan dengan benar dan salah (*True Possitive Possitive, Negatif False Positiv, Netral False Positive*). Dimana rumus dari perhitungan Presisi dapat dilihat pada Persamaan 4.

$$Presisi = \frac{TPP}{(TPP + NegFP + NetFP)} \times 100 \quad (4)$$

c. Sensitivitas

Sensitivitas adalah perhitungan tingkat keberhasilan model klasifikasi dalam mengklasifikasi suatu kelas, dengan cara jumlah klasifikasi entitas oleh model yang bernilai benar (*True Possitive Possitive*), dibagi dengan jumlah entitas (*True Possitive Possitive, Positive False Negatif, Positive False Netral*). Dimana rumus dari perhitungan sensitivitas dapat dilihat pada Persamaan 5.

$$Sensitivitas = \frac{TPP}{(TPP + PFNeg + PFNet)} \times 100 \quad (5)$$

HASIL DAN PEMBAHASAN

Pada tahap pengambilan data menggunakan teknik *web scraping*, menghasilkan data JNE sebanyak 3565 data ulasan dan J&T Express sebanyak 3967 data

ulasan. Setelah data didapatkan maka akan dilakukan *preprocessing*. Ulasan yang diperoleh dari situs *google play store* seringkali mengandung karakter atau kata yang tidak formal dan tidak terstruktur. Oleh karena itu, dalam tahap ini, karakter dan kata-kata tersebut dihilangkan atau diubah ke dalam bentuk yang lebih terstruktur untuk memudahkan proses klasifikasi.

Cleaning

Dalam penelitian ini, tahap pembersihan (*Cleaning*) digunakan untuk menghilangkan tanda baca, emotikon, karakter khusus, angka, dan tautan dari teks ulasan. Contoh hasil dari tahap *cleaning* dapat dilihat pada Tabel 2.

Tabel 2. Contoh Hasil *Cleaning*

<i>text</i>	<i>cleaning</i>
Expedisi kebanyakan tidur ini. Paket g jalan2 udh seminggu lebih..	Expedisi kebanyakan tidur ini Paket g jalan udh seminggu lebih

Case folding

Dalam penelitian ini *Case folding* digunakan untuk mengubah bentuk teks menjadi *lower case*. Contoh hasilnya dapat dilihat pada Tabel 3.

Tabel 3. Contoh Hasil *Case Folding*

<i>cleaning</i>	<i>case_folding</i>
Expedisi kebanyakan tidur ini Paket g jalan udh seminggu lebih	expedisi kebanyakan tidur ini paket g jalan udh seminggu lebih

Normalisasi

Tahap ini melakukan pencarian kata berslang/kata gaul dalam kamus *slangwords* dan diubah menjadi kata baku. Contoh hasilnya dapat dilihat pada Tabel 4.

Tabel 4. Contoh Hasil Normalisasi

<i>case_folding</i>	<i>normalisasi</i>
expedisi kebanyakan tidur ini paket g jalan udh seminggu lebih	expedisi kebanyakan tidur ini paket tidak jalan sudah seminggu lebih

Tokenizing

Dalam proses ini memecah suatu kalimat menjadi potongan-potongan kata berdasarkan spasi. Contoh hasilnya terlihat pada Tabel 5.

Tabel 5. Contoh Hasil *Tokenizing*

<i>normalisasi</i>	<i>tokenize</i>
expedisi kebanyakan tidur ini paket tidak jalan sudah seminggu lebih	['expedisi', 'kebanyakan', 'tidur', 'ini', 'paket', 'tidak', 'jalan', 'sudah', 'seminggu', 'lebih']

Stemming

Stemming merupakan tahap mengganti tiap kata menjadi bentuk awal (dasar), dimana hasil *stemming* dapat dilihat pada Tabel 6.

Tabel 6. Contoh Hasil *Stemming*

<i>tokenize</i>	<i>stemming</i>
['expedisi', 'kebanyakan', 'tidur', 'ini', 'paket', 'tidak', 'jalan', 'sudah', 'seminggu', 'lebih']	expedisi banyak tidur ini paket tidak jalan sudah minggu lebih

Stopword Removal

Stopword Removal adalah tahapan menghilangkan kata-kata yang umumnya muncul dalam teks namun kurang memberikan kontribusi penting, seperti ada, untuk, di, ke, dari, dan sebagainya. Contoh hasil dari tahap *cleaning* dapat dilihat pada Tabel 7.

Tabel 7. Hasil *Stopword Removal*

<i>stemming</i>	<i>stopwordremoval</i>
expedisi banyak tidur ini paket tidak jalan sudah minggu lebih	expedisi banyak tidur paket tidak jalan minggu lebih

Hasil Klasifikasi Menggunakan *Lexicon-Base*

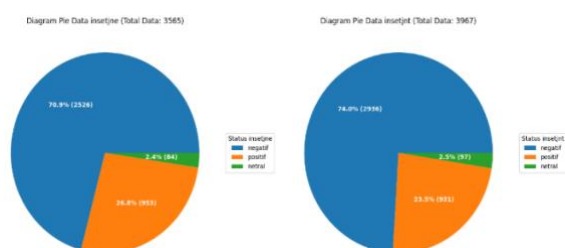
Setelah *preprocessing*, kemudian akan dilakukan klasifikasi, hasil klasifikasi ulasan dibagi menjadi sentimen positif, sentimen negatif dan sentimen netral. Semua data JNE dan J&T yang telah di *preprocessing* akan diklasifikasikan menggunakan algoritma metode *lexicon based* dengan memanfaatkan *InSet Lexicon* (Koto & Rahmaningtyas, 2017a). Sistem akan memeriksa setiap kata dalam data dengan *InSet Lexicon*. Kata dalam data akan memiliki bobot, jika kata tersebut terdapat dalam *InSet Lexicon*. Bobot tersebut akan diakumulasi untuk menghasilkan nilai *polarity* tiap ulasan. *Polarity score* ini yang akan menentukan sentimen tiap kalimat. Apabila nilai *polarity* lebih dari 0 atau bernilai positif maka kalimat tersebut berorientasi sentimen positif, sebaliknya jika nilai *polarity* kurang dari 0 atau bernilai negatif maka kalimat tersebut berorientasi sentimen negatif. Kalimat akan berorientasi sentimen netral jika nilai *polarity*

adalah 0. Contoh hasil *polarity score* dari sejumlah ulasan dapat dilihat pada tabel 8.

Tabel 8. Contoh Hasil Klasifikasi

Ulasan <i>preprocessing</i>	<i>Polarity score</i>	Sentimen
kecewa banget jne kirim paket isi bobol terima cuma dus	-1	Negatif
pertama kirim barang karawang tangerang nyata hari langsung mantap	1,444	Positif
aplikasi kurang mudah kurang tunjuk informasi tingkat kembang	-0,3	Negatif
sangat bantu sekali kirim barang jadi tahu harga cek resi	0,5	Positif
kasih bintang meski kadang lambat kirim nama manusia lebih kurang semangat terus buat seluruh jasa kirim indoensia	0,059	Positif

Untuk mempermudah dalam melihat hasil klasifikasi, seluruh data ulasan yang telah diklasifikasikan akan disajikan dalam bentuk diagram lingkaran (*pie chart*) yang dapat dilihat pada Gambar 2 dibawah ini.

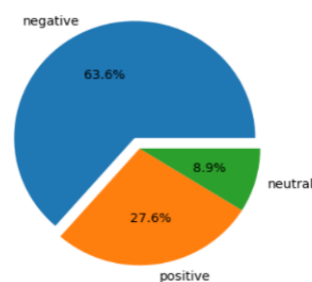


Gambar 2. Pie Chart Hasil Klasifikasi Sentimen JNE dan J&T

Berdasarkan pie chart hasil klasifikasi sentimen JNE dan J&T, dapat dilihat persentase opini masyarakat untuk klasifikasi sentimen negatif sebesar 70,9% untuk JNE dan 74,0% untuk J&T, atau sebanyak 2936 ulasan J&T menunjukkan opini, keluhan, atau pandangan negatif terhadap pelayanan maupun aplikasi kedua jasa ekspedisi tersebut. Sementara itu, persentase opini masyarakat untuk klasifikasi sentimen positif sebesar 26,8% untuk JNE dan 23,5% untuk J&T. Hal ini menunjukkan bahwa hanya 26,8% dari keseluruhan data ulasan JNE dan hanya 23,5% dari keseluruhan data ulasan J&T menampilkan opini positif, baik berupa

rasa senang, terbantu, merasakan manfaat, atau mendukung pelayanan dan aplikasi kedua jasa ekspedisi tersebut. Sisanya adalah 2,4% ulasan JNE dan 2,5% ulasan J&T, yang termasuk dalam sentimen netral.

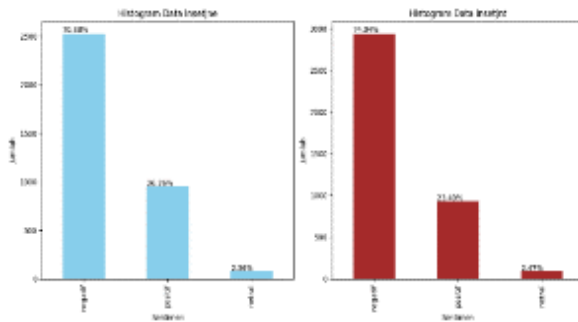
Terdapat juga hasil temuan serupa yang dilakukan oleh Desi Musfiroh bersama teman-temannya pada tahun 2021 dengan judul “Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon” (Musfiroh et al., 2021). Dapat dilihat diagram persentase hasil klasifikasi sentimen yang dilakukan oleh mereka pada Gambar 3 dibawah ini.



Gambar 3. Diagram Persentase Hasil Klasifikasi Sentimen

Dari diagram yang ditampilkan pada Gambar 3, dari data hasil *preprocessing* yang berjumlah 5811. Sebagian besar sentimen yang terdeteksi adalah negatif, yaitu sebesar 63,6% (3696 tweet), yang mencerminkan keluhan, atau pandangan kurang mendukung terhadap pelaksanaan perkuliahan daring di Indonesia. Sentimen positif tercatat sebanyak 27,6% (1604 tweet), berisi tanggapan seperti rasa senang, persetujuan, penerimaan, atau dukungan terhadap perkuliahan daring. Sisanya, 8,9% (517 tweet), menunjukkan sentimen netral. Dapat dilihat bahwa hasil klasifikasi sentimen positif, sentimen negatif, dan sentimen netral memiliki hasil yang sama yaitu lebih dominan adalah hasil klasifikasi sentimen negatif.

Berbeda dengan penelitian yang dilakukan oleh (Musfiroh et al., 2021), Untuk melihat frekuensi setiap kata yang memiliki persentase kemunculan tertinggi, data hasil klasifikasi akan disajikan dalam bentuk histogram, sedangkan dalam penelitian tersebut tidak menggunakan visualisasi histogram. Dimana hasil histogram dalam penelitian ini dapat dilihat pada Gambar 4.

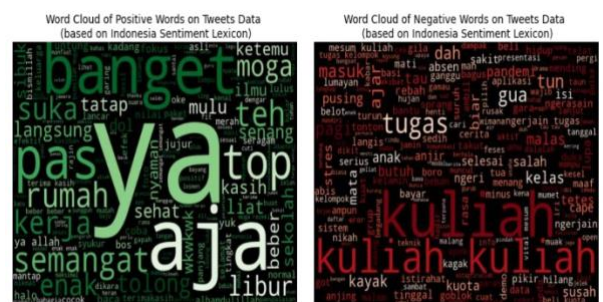


Kata-kata yang muncul pada keseluruhan data ulasan JNE dan J&T akan disajikan dalam bentuk *Word Cloud*. Visualisasi *Word Cloud* JNE dan J&T dapat dilihat pada Gambar 5.



paling mendominasi. Sesuai dengan topik yang dibahas mengenai jasa ekspedisi atau jasa pengiriman, kata **'kirim'** sering mendominasi dalam ulasan karena kata tersebut langsung berkaitan dengan pengalaman masyarakat terhadap layanan pengiriman. Selain itu, kata **'tidak'** sering mendominasi dalam ulasan karena digunakan untuk mengekspresikan perasaan masyarakat terhadap layanan pengiriman.

Sesuai dengan topik yang diangkat, pada hasil penelitian yang dilakukan oleh (Musfiroh et al., 2021) yaitu mengenai ‘Perkuliah Daring’, *wordcloud* sentimen positif dan negatif dapat dilihat pada Gambar 6 dibawah ini, namun untuk *wordcloud* sentimen netral tidak ditunjukkan.



Gambar 6. *WordCloud* Sentimen Positif dan Negatif

Hasil Pengujian Akurasi

Dalam Penelitian ini, pengujian akurasi menggunakan data klasifikasi manual yang dilakukan oleh pakar Bahasa Indonesia. Data klasifikasi manual tersebut akan dibandingkan dengan hasil klasifikasi yang sudah dihasilkan oleh model.

Total dataset awal yang dimasukkan ke dalam sistem adalah sebanyak 3565 ulasan pada JNE dan 3967 ulasan pada J&T. Setelah dilakukannya pembersihan teks dan klasifikasi kemudian dilakukan pengujian, sehingga hasil pengujian memperoleh *confusion matrix* berikut (Tabel 9):

Tabel 9. Hasil pengujian *Confusion Matrix*

JNE		Kelas Prediksi		
		Negatif	Netral	Positif
Kelas Aktual	Negatif	2406	70	424
	Netral	14	6	19
	Positif	107	8	511

J&T		Kelas Prediksi		
		Negatif	Netral	Positif
Kelas Aktual	Negatif	2838	86	612
	Netral	50	4	54
	Positif	49	8	265

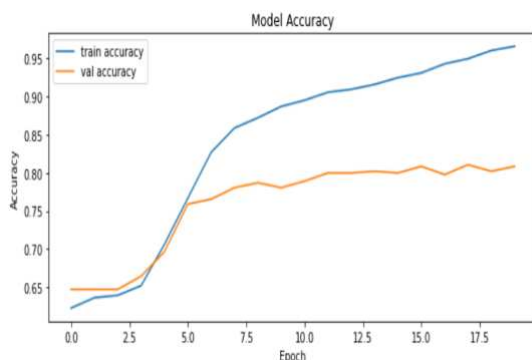
Dari Tabel 9, diperoleh diperoleh hasil akurasi, presisi, dan sensitivitas yang dapat dilihat pada Tabel 10 dibawah ini.

Tabel 10. Hasil Pengujian

JNE	Presisi	Sensitivitas	Akurasi
Negatif	95%	83%	82%
Netral	7%	15%	
Positif	54%	82%	
J&T	Presisi	Sensitivitas	Akurasi
Negatif	97%	80%	78%
Netral	4%	4%	
Positif	28%	82%	

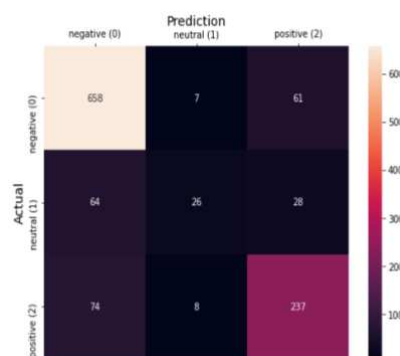
Berdasarkan Tabel 10, akurasi analisis mencapai 82% untuk data JNE dengan nilai presisi untuk sentimen negatif 95%, sentimen positif 54%, sentimen netral 7% sedangkan sensitivitas untuk sentimen negatif 83%, sentimen positif 82%, sentimen netral 15%. Untuk data J&T memiliki akurasi sebesar 78% dengan nilai presisi untuk sentimen negatif 97%, sentimen positif 28%, sentimen netral 4% sedangkan sensitivitas untuk sentimen negatif 80%, sentimen positif 82%, sentimen netral 4%. Secara keseluruhan model dapat mengklasifikasi sebagian data ulasan dengan benar namun hasil prediksi model dalam memprediksi kelas netral dan positif memiliki nilai persentase yang rendah. Kemampuan model dalam memprediksi kelas positif sangat baik walaupun kemampuan model dalam memprediksi untuk kelas netral sangat rendah, hal ini dikarenakan adanya ketidakseimbangan dalam dataset.

(Musfiroh et al., 2021) melakukan penelitian dengan data klasifikasi dibagi menjadi data latih dan data uji, dengan perbandingan 8:2. Dari total data, 80% (4648 tweet) digunakan sebagai data latih, untuk sisanya, yaitu 20% (1163 tweet) digunakan sebagai data uji. Hasil pengujian model klasifikasi menggunakan teknik *cross-validation* dapat dilihat pada Gambar 7.



Gambar 7. Hasil Pengujian Akurasi Model dengan *Cross Validation*

Gambar 7 diatas menunjukkan perbandingan akurasi antara proses training dan testing selama 20 epoch. Grafik memperlihatkan penggunaan epoch yang tinggi dapat meningkatkan akurasi jika jumlah data training mencukupi. Namun, terjadi overfitting, di mana akurasi pada proses training tinggi tetapi rendah pada proses testing. Hasil pengujian 1163 data uji ditampilkan dalam *confusion matrix* pada Gambar 8.



Gambar 8. Hasil *Confusion Matrix* dari Model Klasifikasi

Dengan merujuk pada *confusion matrix* tersebut, didapatkan nilai *accuracy*, *precision*, *recall* dari model klasifikasi yang dapat dilihat pada tabel berikut ini.

Tabel 11. Hasil Pengujian

	Presisi	Sensitivitas	Akurasi
Negatif	72.7%	74.2%	79.2%
Netral	63.4%	23.7%	
Positif	82.6%	90.6%	

Dari keseluruhan hasil pengujian akurasi, terlihat bahwa akurasi pada topik ‘Perkuliahan Daring’ tidak lebih tinggi dibandingkan dengan topik ‘Pelayanan Jasa Ekspedisi JNE’. Namun, nilai akurasinya lebih baik dibandingkan dengan topik ‘Pelayanan Jasa Ekspedisi J&T’.

KESIMPULAN

Terlihat bahwa model memiliki akurasi tinggi sebesar 82% untuk data JNE dan data J&T sebesar 78%. Secara keseluruhan model dapat mengklasifikasi sebagian data ulasan dengan benar namun hasil prediksi model dalam memprediksi kelas netral dan positif memiliki nilai persentase yang rendah. Sementara itu kemampuan model dalam memprediksi kelas positif sangat baik walaupun kemampuan model dalam memprediksi kelas netral sangat rendah.

Visualisasi *pie chart* dan histogram menunjukkan bahwa sebagian besar ulasan pada kedua jasa ekspedisi cenderung negatif. Hal ini tercermin dengan dalam *pie chart* mengilustrasikan proporsi sentimen negatif yang lebih besar dibandingkan sentimen positif dan netral, serta tingginya frekuensi data dengan sentimen negatif pada gambar histogram. Berdasarkan visualisasi *wordcloud*, kata ‘*kirim*’ dan ‘*tidak*’ paling sering muncul dalam ulasan JNE dan J&T, mencerminkan perasaan dan pengalaman masyarakat terkait pengiriman. **Sentimen positif** menunjukkan kepuasan dengan layanan yang berjalan baik dan lancar. **Sentimen negatif** mencerminkan kekecewaan masyarakat terhadap keterlambatan dan masalah pengiriman. **Sentimen netral** bersifat deskriptif, menggambarkan emosi yang tidak kuat.

DAFTAR PUSTAKA

- Affandi, I. R., Hasan, F. N., Pratiwi, N., & Halim, Z. (2022). Analisis Sentimen Opini Masyarakat Terkait Pelayanan Jasa Ekspedisi Anteraja Dengan Metode Naive Bayes. *Jurnal Linguistik Komputasional*, 5(2), Article 2. <https://doi.org/10.26418/jlk.v5i2.107>
- Akbar, M. N., Darmatasia, D., & Ardana, Y. (2022). Analisis Sentimen Terhadap Jasa Ekspedisi Pos Indonesia Pada Sosial Media Twitter Menggunakan Naïve Bayes Classifier. *Journal Software, Hardware and Information Technology*, 2(2), Article 2. <https://doi.org/10.24252/shift.v2i2.34>
- Alita, D., & Isnain, A. R. (2020). Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *jurnal komputasi*, 8(2). <https://doi.org/10.23960/komputasi.v8i2.2615>
- Amaliah, F., & Dwi Nuryana, I. K. (2022). Perbandingan Akurasi Metode Lexicon Based Dan Naive Bayes Classifier Pada Analisis Sentimen Pendapat Masyarakat Terhadap Aplikasi Investasi Pada Media Twitter. *Journal of Informatics and Computer Science (JINACS)*, 3(03), 384–393. <https://doi.org/10.26740/jinacs.v3n03.p384-393>
- Basryah, E. S., Erfina, A., & Warman, C. (2021). Analisis sentimen aplikasi dompet digital di era 4.0 pada masa pandemi covid-19 di play store menggunakan algoritma naive bayes classifier.
- Fadhillah, I. F. N., Herdiani, A., & Astuti, W. (2019). Analisis Sentimen Berbasis Leksikon InSet Terhadap Partai Politik Peserta Pemilu 2019 Pada Media Sosial Twitter. *eProceedings of Engineering*, 6(3), Article 3. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/11216>
- Fatihin, A. (2022). Analisis sentimen terhadap ulasan aplikasi mobile menggunakan metode support vector machine (SVM) dan pendekatan lexicon based (Bachelor's thesis, Fakultas Sains dan Teknologi Universitas Islam Negeri Syarif Hidayatullah Jakarta).
- Ismail, A. R., & Hakim, R. B. F. (2023). Implementasi Lexicon Based Untuk Analisis Sentimen Dalam Mengetahui Trend Wisata Pantai Di DI Yogyakarta Berdasarkan Data Twitter. 1.
- Koto, F., & Rahmaningtyas, G. (2017). InSet Lexicon: Evaluation of a Word List for Indonesian Sentiment Analysis in Microblogs. <https://doi.org/10.1109/IALP.2017.8300625>
- Kurniawan, T. (2017). Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Media Mainstream Menggunakan Naïve Bayes Classifier Dan Support Vector Machine. *Institut Teknologi Sepuluh Nopember*, 2774-1702.
- Musfiroh, D., Khaira, U., Utomo, P. E. P., & Suratno, T. (2021). Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon: Sentiment Analysis of Online Lectures in Indonesia from Twitter Dataset Using InSet Lexicon. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), 24–33. <https://doi.org/10.57152/malcom.v1i1.20>
- Nitami, M. T., & Februariyanti, H. (2022). Analisis Sentimen Ulasan Ekspedisi J&T Express Menggunakan Algoritma Naïve Bayes. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 5(1), 20–29. <https://doi.org/10.36595/misi.v5i1.396>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. 5.
- Nurkasanah, A., & Hayaty, M. (2022). Feature Extraction using Lexicon on the Emotion Recognition Dataset of Indonesian Text. *Ultimatics : Jurnal Teknik Informatika*, 14(1), 20–27. <https://doi.org/10.31937/ti.v14i1.2540>
- Ramadhan, A. S., & Marpaung, N. (2024). Pengaruh Word Of Mouth, Dissatisfaction, Variety Seeking Terhadap Brand Switching Jasa Kurir J&T Express. *Transekonomika: Akuntansi, Bisnis Dan Keuangan*, 4(2), 132-146.
- Rajput, Q., Haider, S., & Ghani, S. (2016). Lexicon-Based Sentiment Analysis of Teachers' Evaluation. *Applied Computational Intelligence and Soft Computing*, 2016, 1–12. <https://doi.org/10.1155/2016/2385429>
- Rezki, M., Kholifah, D. N., Faisal, M., Priyono, P., & Suryadithia, R. (2020). Analisis Review Pengguna Google Meet dan Zoom Cloud Meeting Menggunakan Algoritma Naïve Bayes. *Jurnal Infortech*, 2(2), Article 2. <https://doi.org/10.31294/infortech.v2i2.9286>
- Singh, V., Singh, G., Rastogi, P., & Deswal, D. (2018). Sentiment Analysis Using Lexicon

- Based Approach. *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 13–18.
<https://doi.org/10.1109/PDGC.2018.8745971>
- Sitorus, P. R. (2021). *Analisis Sentimen Data Ulasan Aplikasi Indriver pada Situs Google Play Menggunakan Metode Naïve Bayes Classifier dan Support Vector Machine* [Thesis, Universitas Sumatera Utara].
<https://repositori.usu.ac.id/handle/123456789/46908>
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, 37(2), 267–307.
https://doi.org/10.1162/COLL_a_00049
- Turney, P. D. (2001). Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*, 417.
<https://doi.org/10.3115/1073083.1073153>
- Yerzi, F. S., & Sibaroni, Y. (2021). Analisis Sentimen Terhadap Kebijakan Pemerintah Dalam Menangani Covid-19 Dengan Pendekatan Lexicon Based. *eProceedings of Engineering*, 8(5).