

## Rekomendasi Peningkatan Layanan Okejek dengan *Root Cause Analysis* berdasarkan Hasil Analisis Sentimen Ulasan Pengguna

Andhika Mifta Alauddin<sup>1</sup>, Nanang Yudi Setiawan<sup>2</sup>, Mochamad Chandra Saputra<sup>3</sup>

Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Brawijaya  
Email: <sup>1</sup>dikatokin@ub.ac.id, <sup>2</sup>nanang@ub.ac.id, <sup>3</sup>andra@ub.ac.id

### Abstrak

PT. Okejek Kreasi Indonesia merupakan perusahaan yang bergerak di bidang transportasi daring dengan produk unggulannya yaitu Okejek. Namun demikian, Okejek belum mampu bersaing dengan raksasa di bidang yang sama. Untuk mengetahui penyebab utama dari permasalahan tersebut, salah satu cara yang dapat dilakukan yaitu dengan menggunakan analisis sentimen. Oleh karena itu, sebuah analisis sentimen dengan menggunakan 1890 data yang bersumber dari ulasan pengguna pada Google Play Store kemudian dilakukan. Dengan menggunakan algoritma Recurrent Neural Network, didapatkan nilai akurasi dan nilai F1 score sebesar 96%. Setelah analisis sentimen berhasil dilakukan, maka dilanjutkan dengan Root Cause Analysis yang berdasarkan kepada 3 aspek yaitu aspek pelayanan, aspek harga, dan aspek staff. Dari ketiga aspek tersebut ditemukan permasalahan utama yang terletak pada masalah di aplikasi, perilaku mitra pengemudi dan sistem pembayaran yang ada. Oleh karenanya, dilakukan penyusunan rekomendasi bersama dengan stakeholder yang menghasilkan solusi berupa peningkatan aplikasi, memperketat SOP dan juga memperbaiki layanan pembayaran

**Kata kunci:** *analisis sentimen, recurrent neural network, root cause analysis, aspek, rekomendasi*

### Abstract

*PT. Okejek Kreasi Indonesia is one of the company that engaged in the field of online transportation with its flagship product, Okejek. However, Okejek still unable to compete with the giants in the field of online transportation. To find out the main causes of these problems, one way that can be done is by using sentiment analysis. Therefore, a sentiment analysis using 1890 data from user reviews on Google Play Store is performed. By using Recurrent Neural Network algorithm, an accuracy value and F1 score of 96% is obtained. After the sentiment analysis was successfully carried out, it is followed by Root Cause Analysis that based on 3 aspects which are service, price, and staff aspects. From these 3 aspects, the main problem were found to be a problem in the application, driver-partner's behavior, and the existing payment system. Therefore, a joint recommendation were made with stakeholders which resulted in the form of solutions that are improving applications, tightening SOPs, and also improving payment services.*

**Keywords:** *sentiment analysis, recurrent neural network, root cause analysis, aspects, recommendations*

### 1. PENDAHULUAN

Sejak hadirnya Gojek pada tahun 2016, wajah transportasi Indonesia berubah drastis. Jika pada awalnya masyarakat harus menaiki transportasi umum atau menggunakan transportasi pribadi seperti sepeda motor dan mobil, maka saat ini kemudahan dalam melakukan mobilisasi didukung oleh kemunculan layanan transportasi daring berupa ojek sepeda motor dan mobil. Aziah et al. (2018) menerangkan bahwa kehadiran layanan ojek

daring ini menawarkan berbagai keuntungan bagi penggunanya mulai dari segi estimasi biaya yang jelas, pelacakan kendaraan melalui sistem GPS, dan masih banyak keunggulan lain. Menurut Alvara Institute (2020), layanan ojek daring pada umumnya menjadi salah satu jenis aplikasi yang paling banyak diunduh oleh masyarakat dengan rerata diatas 40%. Selain itu, fakta lain yang diungkapkan oleh riset Alvara yaitu sebanyak 78% generasi milenial dan Gen-Z menghabiskan 78% pengeluaran digital mereka untuk keperluan transportasi daring.

Fakta-fakta tersebut tentunya dapat menjadi bukti bahwa bisnis transportasi daring memiliki potensi yang tinggi. Hal tersebut pada akhirnya menjadi pemicu lahirnya penyedia jasa baru di bidang transportasi daring. Para pemain baru ini meliputi NUjek, Maxim, Anterin dan Okejek.

Okejek merupakan sebuah penyedia jasa layanan transportasi daring yang berada di bawah naungan PT. Okejek Kreasi Indonesia yang berdiri pada tahun 2015. Meskipun demikian, kehadiran okejek masih jarang ditemui di jalan raya. Menurut riset yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (2020), Okejek masih belum masuk kedalam jajaran layanan jasa transportasi daring yang dipilih oleh masyarakat. Oleh karena itu, perlu ditelusuri penyebab mengapa Okejek belum menjadi pilihan masyarakat dengan menggunakan analisis sentimen untuk meningkatkan kepuasan pelanggan.

Penelitian yang dilakukan oleh Mas Pintoko & Muslim (2018) dan Garbian Nugroho et al. (2016) menunjukkan bahwa analisis sentimen dapat digunakan untuk mengetahui tingkat kepuasan pelanggan terhadap jasa transportasi daring. Keduanya menggunakan metode Naïve Bayes dalam proses analisis sentimen dan mendapatkan hasil masing-masing 86% dan 80% tingkat akurasi keberhasilan. Sementara itu, Merinda Lestandy et al. (2021) menggunakan metode RNN untuk melakukan analisis sentimen terhadap tweet vaksin COVID-19 dan mendapatkan hasil 97% tingkat akurasi keberhasilan. Dengan metode yang sama, Kurniasari & Setyanto (2020) mendapatkan tingkat akurasi sebesar 92%. Dengan demikian, digunakan metode RNN sebagai algoritma dalam melakukan analisis sentimen.

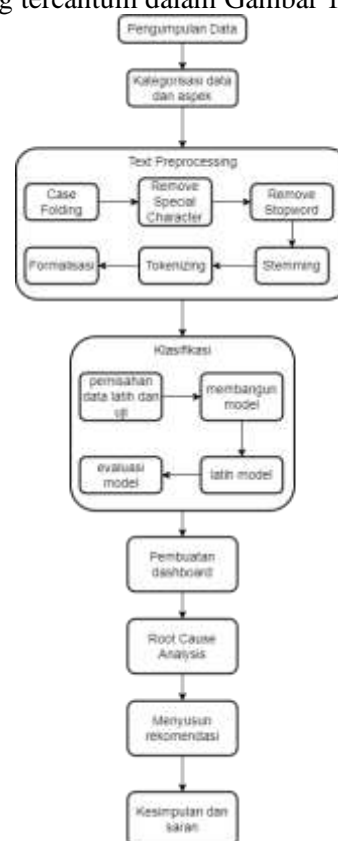
Selanjutnya, untuk dapat memberikan rekomendasi perbaikan terhadap layanan Okejek, diperlukan analisis lebih mendalam terhadap klasifikasi sentimen tersebut. Salah satu cara yang dapat digunakan dalam melakukan analisis yaitu dengan menggunakan Root Cause Analysis. Pada penelitian yang dilakukan oleh Arunachalam & Sarkar (2013), analisis sentimen berbasis aspek dapat digabungkan dengan Root Cause Analysis untuk menentukan faktor utama dari sentimen negatif pengguna. Selain itu, penelitian yang dilakukan Park & Kim (2021) juga menunjukkan bahwa Root Cause Analysis dapat digunakan untuk memperoleh hubungan antar aspek sentimen dari

analisis sentimen yang telah dilakukan.

Berdasarkan kajian terhadap beberapa penelitian diatas, maka dapat dibuktikan bahwa analisis sentiment dapat digunakan untuk dapat meningkatkan kualitas layanan objek serta sebagai bahan pertimbangan bagi pemangku kepentingan dalam pengambilan keputusan. Penelitian ini akan melakukan analisis sentimen terhadap ulasan pengguna Aplikasi Okejek dengan data yang didapatkan dari Google Play Store. Data diambil menggunakan metode web scraping dengan bantuan Selenium Webdriver. Selanjutnya akan dilakukan preprocessing terhadap data yang telah didapatkan. Selanjutnya, akan dilakukan analisis sentimen terhadap aspek-aspek yang telah ditentukan dengan menggunakan metode Recurrent Neural Network (RNN). Hasil dari analisis tersebut kemudian akan dimanfaatkan untuk menyusun rekomendasi perbaikan bagi layanan Okejek dengan pendekatan Root Cause Analysis.

## 2. METODOLOGI PENELITIAN

Alur penelitian analisis sentimen terhadap layanan Okejek dibagi menjadi beberapa bagian seperti yang tercantum dalam Gambar 1.



Gambar 1. Metodologi

### 2.1. Pengumpulan Data

Dalam penelitian ini, semua data bersumber

dari ulasan pengguna aplikasi Okejek pada platform Google Play Store. Data yang diekstrak merupakan hasil ulasan murni dari Google Play Store tanpa mengekstrak identitas maupun bintang dari pengguna. Data diekstrak dengan menggunakan bantuan web scraper Selenium berbasis Python. Dalam melakukan web scraping tersebut, berhasil didapatkan data ulasan pengguna sebanyak 2241 ulasan dari 3 tahun terakhir.

**2.2. Kategorisasi Data dan Aspek**

Pada tahap ini, dilakukan pelabelan terhadap data yang telah diekstrak. Terdapat 2 jenis label yang akan digunakan yaitu labelisasi polaritas negatif positif serta labelisasi aspek. Dalam melakukan pelabelan tersebut, penulis dibantu oleh beberapa ahli. Menurut Goodman (1987 disitasi dalam Keeney et al., 2001), seorang ahli dapat didefinisikan sebagai seseorang yang spesialis dalam bidang yang spesifik. Oleh karena itu, penulis memilih ahli yang memiliki keahlian di bidang Sastra Indonesia. Profil ahli yang membantu penulis dalam melakukan pelabelan dapat dilihat pada Tabel 1.

Tabel 1. Profil Ahli

Nama	Usia	Profesi
Rahma Dewi Nur Hanif Ahfanti	22 th	Mahasiswi S1 Sastra Indonesia
Cinta Putri Prameswari	20 th	Mahasiswi S1 Sastra Indonesia

Setelah dilakukan observasi pada proses pelabelan, didapatkan 3 aspek besar yaitu pelayanan, performa staff/mitra, dan harga atau pembayaran. Hal ini didukung oleh penelitian Ching & de Dios Bulos (2019) yang menyatakan bahwa aspek tersebut merupakan aspek-aspek yang memiliki pengaruh terhadap sentimen dari pengguna layanan. Penelitian serupa oleh Joseph & Varghese (2019) juga membuktikan bahwa aspek-aspek tersebut dapat mewakili sentimen pengguna pada layanan AirBnB. Ketiga aspek tersebut kemudian dibagi kembali menjadi dua tipe polaritas yaitu ulasan positif dan ulasan negatif. Adapun contoh pemberian aspek dapat dilihat pada tabel 2.

Tabel 2. Pengkategorian aspek

Ulasan	Aspek	Label
Sangat membantu aktifitas	Pelayanan	Positif
Ulasan	Aspek	Label

Susah banget untuk dapet driver nya (daerah papua)	Staff	Negatif
Murahnya beneran tanpa embel-embel	Harga	Positif

**2.3. Text Preprocessing**

Setelah dataset berhasil disusun dan diberi label baik label aspek maupun polaritas, langkah selanjutnya yaitu mengolah data. Tahap pertama dalam pengolahan data adalah *text preprocessing*. Menurut Malini et al. (2014), *Preprocessing* merupakan langkah yang penting dan kritis di dalam *text mining*. Dalam area *text mining*, *preprocessing* digunakan untuk mengekstrak informasi yang menarik dan tidak sederhana. Text preprocessing ini terdiri dari beberapa tahap antara lain case folding, penghapusan karakter spesial, penghapusan *stopwords*, *stemming*, tokenisasi dan formalisasi. Hasil dari *text preprocessing* ini dapat dilihat pada gambar 2.

Ulasan	Case Folding	Special Character	Stopwords	Stemming	Tokenisasi	Formalisasi
Ulasan bagus sih, cuma sering eror aja. Kasian driver nya jd gadapef orderan	udahn bagus sih, cuma sering eror aja. kasiann driver nya jd gadapef orderan	udah bagus sih. cuttla sering eror aja. kasiann driver nya jd gadapef orderan	udah bagus sih. cuma sering eror aja. kasiann driver nya jd gadapef orderan	udah bagus sih. cuma sering eror aja. kasiann driver nya jd gadapef order	{udah, 'bagus', 'sih', 'cuma', 'sering', 'eror', 'aja', 'kasi', 'kasinn', 'driver', 'nya', 'jd', 'gadapef', 'order'}	sudah bagus. cuma eror. kasiinn. pengemudi. order.

Gambar 2. Text Preprocessing

**2.4. Proses Klasifikasi**

Setelah data berhasil dibersihkan melalui preprocessing, langkah selanjutnya adalah pembuatan model klasifikasi. Terdapat beberapa tahap yang dilakukan dalam perancangan model ini. Tahapan tersebut meliputi pembagian data menjadi data latih, data validasi dan data uji. Selanjutnya dilakukan pembangunan model dengan menggunakan Recurrent Neural Network (RNN). Setelah model berhasil dibangun, model kemudian akan dilatih menggunakan data latih. Selain itu, hasil latihan tersebut akan divalidasi dengan menggunakan data validasi sebelum masuk kedalam evaluasi model sebagai tahap terakhir.

Tahap pertama yang akan dilakukan dalam perancangan ini adalah pemisahan dataset menjadi paling tidak 2 bagian yaitu data latih dan data uji. Namun pada sebagian kasus, dataset akan dipecah menjadi 3 bagian yaitu data latih, data validasi dan data uji. Pada penelitian ini, dataset akan dibagi menjadi 3 bagian dengan 2 macam proporsi yaitu 70%-10%-20% dan 80%-5%-15% yang masing-masing mewakili data

latih, data validasi, dan data uji.

Setelah dataset dibagi, maka langkah selanjutnya adalah merancang struktur model. Pada penelitian ini, model yang akan digunakan yaitu model dengan arsitektur Recurrent Neural Network (RNN) dengan implementasi Long-Short Term Memory (LSTM). Konfigurasi model RNN yang akan dibangun dapat dilihat pada gambar 3.



Gambar 3. Arsitektur model RNN

Pada layer LSTM, ulasan akan diolah dengan memperhatikan hubungan antar kata dengan memperhatikan riwayat perhitungan sebelumnya. Optimasi dengan fungsi ReLU dan tanh juga dilakukan untuk menghasilkan prediksi yang akurat. Pada layer LSTM, akan dilakukan penggabungan dengan layer Bidirectional yang memungkinkan layer untuk bekerja maju atau mundur dalam dimensi waktu. Akan ada beberapa konfigurasi pada layer embedding dan lstm yang masing-masing menggunakan 32 dan 64 unit embedding serta 50 dan 100 unit lstm.

Setelah model berhasil dibangun, maka langkah selanjutnya adalah melatih model. Model akan dilatih menggunakan data train dengan konfigurasi Batch Size sebesar 64 data per batch dan 100 kali perulangan latihan atau Epoch. Model sebelumnya akan di compile terlebih dahulu dengan pengaturan fungsi loss Binary Crossentropy dan fungsi optimizer Adam. Pelatihan model menggunakan supervised learning dimana ulasan yang akan dimasukkan telah memiliki label khusus sehingga pada setiap perulangan latihan, model akan mempelajari ulasan yang masuk untuk kemudian disesuaikan dengan label yang telah ditentukan. Setelah model dilatih menggunakan data latih, maka model akan dievaluasi menggunakan data validasi.

Setelah selesai dilatih, maka model akan diuji tingkat presisi dan akurasi. Proses pengujian ini akan dilakukan dengan metode confusion matrix. Ulasan pada data uji akan dimasukkan kedalam model untuk diprediksi

polaritas dari ulasan tersebut. Nantinya hasil prediksi ini akan dibandingkan dengan label murni dari ulasan tersebut. Hasil perbandingan akan dihitung menggunakan confusion matrix sehingga dapat ditemukan tingkat akurasi, presisi dan nilai nilai lain dari model tersebut. Hasil confusion matrix juga akan divisualisasikan dalam bentuk grafik.

**2.5. Pembuatan Dashboard**

Setelah didapatkan model terbaik dan dilakukan prediksi terhadap dataset yang ada, dilakukan pembuatan hasil visualisasi terhadap hasil tersebut. Visualisasi yang akan dibuat berbentuk dashboard dengan bantuan Google Data Studio.

**2.6. Root Cause Analysis**

Setelah dashboard selesai dibuat, maka data tersebut akan kembali diolah untuk menemukan permasalahan utama dari rendahnya minat pengguna terhadap Okejek. Berdasarkan hasil sentimen negatif, akan dilakukan Root Cause Analysis untuk dapat mengetahui penyebab utama sentimen negatif dari semua aspek.

**2.7. Penyusunan Rekomendasi**

Setelah didapatkan akar utama permasalahan dari setiap aspek, maka langkah terakhir dalam proses ini adalah penyusunan rekomendasi. Penyusunan rekomendasi dilakukan dalam dua tahap yaitu penyusunan rekomendasi berdasarkan pada Root Cause Identification dan penyusunan rekomendasi bersama stakeholder.

**3. HASIL DAN ANALISIS**

**3.1. Hasil Akurasi Model**

Setelah dilakukan implementasi terhadap dataset beserta model yang akan digunakan, maka langkah selanjutnya yaitu memilih hasil terbaik dari konfigurasi-konfigurasi yang telah ditentukan sebelumnya. Pada konfigurasi pembagian dataset yang pertama yaitu 70%-10%-20%, didapatkan hasil akurasi confusion matrix yang dapat dilihat pada tabel 3.

Tabel 3. Hasil akurasi 70%-10%-20%

Embed ding Dim	LSTM Dim	Accu racy	Preci sion	Recall	F1 Score
32	50	96%	96%	96%	96%
32	100	93%	93%	93%	93%

64	50	92%	92%	92%	92%
64	100	93%	93%	93%	93%

Sedangkan pada konfigurasi pembagian dataset yang kedua yaitu 80%-5%-15%, didapatkan hasil akurasi confusion matrix yang dapat dilihat pada tabel 4.

Tabel 4. Hasil akurasi 80%-5%-15%

Embed ding Dim	LSTM Dim	Accu racy	Preci sion	Recall	F1 Score
32	50	94%	94%	94%	94%
32	100	93%	93%	93%	93%
64	50	95%	95%	95%	95%
64	100	92%	92%	92%	92%

Pada kedua proporsi tersebut, didapatkan hasil bahwa terdapat ketidakseimbangan proporsi data antara ulasan negatif dan ulasan positif. Pada proporsi pembagian 70%-10%-20%, didapatkan data uji sebanyak 396 data uji dengan proporsi 130 ulasan negatif dan 266 ulasan positif atau 33% ulasan negatif berbanding dengan 67% ulasan positif. Sedangkan pada proporsi pembagian 80%-5%-15% didapatkan data uji sebesar 297 data uji dengan proporsi 93 ulasan negatif dan 204 ulasan positif atau 32% ulasan negatif berbanding dengan 68% ulasan positif. Oleh karena itu, digunakan hasil pada proporsi 70%-10%-20% untuk mendapatkan hasil terbaik. Oleh karena itu, didapatkan model terbaik yaitu model dengan konfigurasi 32 dimensi pada layer embedding dan 32 dimensi pada layer LSTM dengan accuracy, precision, recall, dan F1 score sebesar 96%. Visualisasi uji confusion matrix dapat dilihat pada gambar 4.

	precision	recall	f1-score	support
Negatif	0.93	0.94	0.93	130
Positif	0.97	0.97	0.97	266
accuracy			0.96	396
macro avg	0.95	0.95	0.95	396
weighted avg	0.96	0.96	0.96	396

Gambar 4. Hasil Confusion Matrix

Namun demikian, penggunaan accuracy dirasa kurang tepat karena tidak dapat merepresentasikan performa model secara keseluruhan. Untuk itu, digunakan nilai F1 Score yang dapat merepresentasikan performa model karena nilai tersebut merupakan penggabungan dari nilai precision sebagai tingkat presisi dari hasil prediksi dan nilai recall sebagai kemampuan model dalam mengenali nilai yang benar (true positive rates). Selain *F1 Score*,

Menurut Al-Azani & El-Alfy (2017), penggunaan weighted average pada data yang tidak seimbang efektif untuk menentukan performa model karena perhitungan nilai-nilai precision, recall, dan F1 memperhatikan weight atau proporsi dari data uji tersebut.

### 3.2. Visualisasi Dashboard

Setelah menentukan model terbaik, maka langkah selanjutnya adalah memvisualisasikan hasil tersebut kedalam sebuah dashboard. Pada penelitian ini, pembuatan dashboard dilakukan dengan Google Data Studio.



Gambar 4. Dashboard Halaman Pertama

Gambar 4 merupakan visualisasi halaman pertama pada dashboard yang terdiri dari 3 bagian. Bagian pertama yang terdapat pada bagian kiri atas merupakan deskripsi data secara keseluruhan yang berupa jumlah ulasan yang berhasil diekstrak dan rentang tahun pengambilan ulasan. Bagian kedua berada pada bagian tengah hingga kanan atas yang berisi visualisasi pembagian klasifikasi data dalam bentuk diagram. Terdapat dua diagram yang memvisualisasikan pembagian data ulasan dari segi sentimen dan juga dari segi aspek. Sedangkan bagian terakhir yang berada pada bagian bawah merupakan tabel berisi seluruh ulasan yang berhasil diekstrak.



Gambar 5. Dashboard Halaman Kedua

Gambar 5 menunjukkan halaman kedua dashboard yang berfokus pada detail dari

sentimen ulasan yang telah diekstrak. Terdapat 3 bagian utama yang ada pada halaman kedua dashboard yaitu visualisasi wordcloud, polaritas sentimen, dan perangkungan ulasan. Pada bagian visualisasi wordcloud, terdapat satu diagram dari plugin Vega yang menampilkan hasil wordcloud dengan default pada semua aspek. Pada visualisasi dilengkapi dengan dropdown sehingga pengguna dapat memilih untuk menampilkan wordcloud dari aspek mana saja. Pada bagian polaritas sentimen, terdapat satu diagram yang berisi jumlah polaritas sentimen pada setiap aspek dan label. Sedangkan pada bagian perangkungan ulasan, terdapat dua tabel yang masing-masing merepresentasikan sentimen positif dan negatif. Bagian ini digunakan sebagai alat bantu untuk melakukan *root cause analysis*.

### 3.3. Root Cause Analysis

Setelah menentukan model terbaik serta memvisualisasikannya pada sebuah dashboard, maka langkah selanjutnya adalah melakukan proses Root Cause Analysis untuk menentukan masalah utama yang akan diperbaiki. Proses Root Cause Analysis pada penelitian ini menggunakan metode ekstraksi kata atau frasa yang berbentuk wordcloud. Menurut Arunachalam & Sarkar (2013), penggunaan hotword atau kata penting didalam sebuah ulasan dapat membantu mengurangi beban kerja dalam melakukan Root Cause Analysis.

Dalam membuat wordcloud, penulis menggunakan plugin Vega dalam Google Data Studio. Selain menggunakan hotword, perangkungan kata dalam ulasan per aspek seperti yang dijelaskan oleh Selvam et al. (2019) juga akan digunakan dalam penelitian ini. Perangkungan ulasan dilakukan dengan menggunakan metode Term Frequency yaitu dengan melihat jumlah frekuensi sebuah kata pada ulasan yang telah diekstrak. Hasil dari wordcloud dan perangkungan kemudian akan dikombinasikan dengan 4 tahapan yang dipaparkan oleh Rooney & vanden Heuvel (2004) yaitu mengumpulkan data, causal factor charting, root cause identification, dan recommendation generation.

Pada tahapan pertama yaitu pengumpulan data, data yang akan diproses merupakan data yang telah melalui tahapan text preprocessing serta memiliki sentimen negatif. Hal ini berdasar pada tingkat kemajemukan kata pada ulasan yang telah melalui text preprocessing lebih sedikit dari ulasan sebelum text preprocessing

sehingga pada saat dilakukan proses pembuatan wordcloud, kata-kata yang memiliki makna yang sama namun berbeda secara penulisan dapat diminimalisir untuk mendapatkan hasil yang optimal. Didapatkan total 1097 ulasan dengan rincian 746 ulasan pada aspek pelayanan, 236 ulasan pada aspek staff, dan 115 ulasan pada aspek harga.

Setelah data berhasil terkumpul, maka tahap selanjutnya yaitu melakukan Causal Factor Charting. Tahapan ini bertujuan untuk mencari hal-hal yang menjadi faktor utama dalam sebuah ulasan negatif. Proses Causal Factor Charting dilakukan dengan menggunakan dua metode yaitu menggunakan wordcloud dan perangkungan ulasan per aspek dengan menggunakan metode term frequency. Hasil dari wordcloud beserta perangkungan ini akan dijadikan dasar dalam melakukan root cause identification pada tahap selanjutnya.

Setelah berhasil melakukan ekstraksi pada kata dan frasa yang berkaitan dengan ulasan negatif, maka langkah selanjutnya yaitu Root Cause Identification. Pada proses ini, kata dengan frekuensi kemunculan terbesar pada wordcloud akan dicocokkan pada ranking ulasan apakah memang kata tersebut berada pada skor ulasan dengan ranking teratas atau tidak. Apabila kata tersebut ditemukan pada tabel ranking, maka akan dilakukan penelusuran terhadap ulasan kotor pada dataset untuk menentukan akar permasalahan yang dialami oleh pengguna.

Dari tahapan-tahapan tersebut didapatkan 3 *root cause* atau masalah utama yang masing-masing mewakili setiap aspek. Root cause ini dapat dilihat pada tabel 5.

Tabel 5. *Root Cause Identification*

<i>Root Cause</i>	<b>Temuan dalam Perangkungan</b>
Aplikasi	Aplikasi enggak
	aplikasi enggak jelas
Pengemudi ( <i>driver</i> )	pengemudi enggak mau antar
	pengemudi enggak
Saldo	tarif ongkos tidak sesuai aplikasi
	top up enggak masuk saldo baik aplikasi ribet

### 3.4. Penyusunan Rekomendasi

Setelah didapatkan akar utama permasalahan dari setiap aspek, maka langkah terakhir dalam proses ini adalah penyusunan rekomendasi. Penyusunan rekomendasi dilakukan dalam dua tahap yaitu penyusunan rekomendasi berdasarkan pada Root Cause Identification dan penyusunan rekomendasi bersama stakeholder. Pada tahap pertama, dirumuskan rekomendasi seperti yang tercantum pada tabel 6.

Tabel 6. Rekomendasi Awal

Aspek	Rekomendasi
Pelayanan	Melakukan perbaikan terhadap beberapa fitur seperti maps dan optimasi aplikasi
Staff	Menerapkan SOP yang ketat beserta sanksi yang tegas kepada driver
Harga	Memperbaiki sistem pembayaran pada aplikasi dan memberi jaminan bagi konsumen

Rekomendasi tersebut kemudian dikirimkan kepada stakeholder untuk dilakukan kajian dan diskusi mengenai rekomendasi apa yang telah tepat sasaran dan rekomendasi apa yang harus dihapus karena sudah tidak relevan. Setelah dilakukan proses tersebut, maka hasil rekomendasi akhir yang dapat diterapkan kepada pihak stakeholder Okejek dapat dilihat pada tabel 7.

Tabel 7. Rekomendasi Akhir

Aspek	Rekomendasi
Pelayanan	Optimasi pada layanan maps serta memberikan optimasi layanan pada setiap daerah
Staff	SOP serta sanksi lebih dipertegas dan diselaraskan di setiap daerah
Harga	Pemberian jaminan pada konsumen serta layanan aduan yang responsif

### 4. KESIMPULAN

Okejek sebagai salah satu penyedia jasa transportasi daring memang kurang dikenal dan dipercaya oleh masyarakat pada umumnya. Hal ini disebabkan oleh beberapa faktor yang telah

ditemukan dalam 3 aspek besar yaitu aspek pelayanan, aspek staff, dan aspek harga. Dengan melakukan ekstraksi data ulasan pengguna dari Google Play Store selama 3 tahun terakhir, ditemukan 1097 ulasan negatif dari total 2240 ulasan atau 50% dari total ulasan yang berhasil diekstrak.

Hal ini didukung dengan hasil analisis sentimen menggunakan model arsitektur Recurrent Neural Network yang berhasil meraih akurasi F1 score 96% secara keseluruhan dan 93% pada ulasan negatif. Kemudian, setelah dilakukan penelusuran menggunakan Root Cause Analysis, ditemukan tiga masalah utama yang ada dalam sistem pelayanan Okejek yaitu masalah pada aplikasi, driver mitra yang kurang profesional dan juga sistem pembayaran yang kurang maksimal.

Oleh karena itu, dilakukan diskusi bersama pihak dari Okejek selaku stakeholder untuk menyusun beberapa rekomendasi yang bisa diterapkan oleh Okejek. Rekomendasi ini meliputi Optimasi pada layanan maps serta memberikan optimasi layanan pada setiap daerah, sosialisasi SOP serta sanksi lebih dipertegas dan diselaraskan di setiap daerah, serta Pemberian jaminan saldo kembali pada konsumen sekaligus layanan aduan yang responsif.

### 5. DAFTAR PUSTAKA

Alvara Institute. (2020). *The Battle Of Our Generation INDONESIA GEN Z AND MILLENNIAL REPORT 2020: Januari 2020*.

Arunachalam, R., & Sarkar, S. (2013). The New Eye of Government: Citizen Sentiment Analysis in Social Media. *Proceedings of the IJCNLP 2013 Workshop on Natural Language Processing for Social Media (SocialNLP)*, 23–28. <http://www.jeffbullas.com/2012/11/28/the-latest-27->

Asosiasi Penyelenggara Jasa Internet Indonesia. (2020). *Grab dan Gojek, Layanan Transportasi Online Paling Populer di Masyarakat*. <https://databoks.katadata.co.id/datapublish/2020/11/11/grab-dan-gojek-layanan-transportasi-online-paling-populer-di-masyarakat>

Aziah, A., Popon, Adawia, R., & Sitasi, C. (2018). Analisis Perkembangan Industri

- Transportasi Online di Era Inovasi Disruptif (Studi Kasus PT Gojek Indonesia). *Cakrawala*, 18(2), 149–156. <https://doi.org/10.31294/jc.v18i2>
- Ching, M. R. D., & de Dios Bulos, R. (2019). Improving restaurants' business performance using yelp data sets through sentiment analysis. *ACM International Conference Proceeding Series*, 62–67. <https://doi.org/10.1145/3340017.3340018>
- Garbian Nugroho, D., Herry Chrisnanto, Y., Wahana Jurusan Informatika, A., & Matematika dan Ilmu Pengetahuan Alam Universitas Jenderal Achmad Yani Jalan Terusan Jenderal Sudirman, F. (2016). ANALISIS SENTIMEN PADA JASA OJEK ONLINE MENGGUNAKAN METODE NAÏVE BAYES. *Prosiding SNST Fakultas Teknik*, 1(1).
- Joseph, G., & Varghese, V. (2019). Analyzing Airbnb customer experience feedback using text mining. In *Big Data and Innovation in Tourism, Travel, and Hospitality: Managerial Approaches, Techniques, and Applications* (pp. 147–162). Springer Singapore. [https://doi.org/10.1007/978-981-13-6339-9\\_10](https://doi.org/10.1007/978-981-13-6339-9_10)
- Keeney, S., Hasson, F., & Mckenna, H. P. (2001). A critical review of the Delphi technique as a research methodology for nursing. *International Journal of Nursing Studies*, 38(2), 195–200. [www.elsevier.com/locate/ijnurstu](http://www.elsevier.com/locate/ijnurstu)
- Kurniasari, L., & Setyanto, A. (2020). Sentiment Analysis using Recurrent Neural Network. *Journal of Physics: Conference Series*, 1471(1). <https://doi.org/10.1088/1742-6596/1471/1/012018>
- Malini, W., Bashir, M., Rozaimée, A., Ch, D., Daud Associat Professor, A., Gurusamy, V., & Professor, A. (2014). Preprocessing Techniques for Text Mining RIDDHI REX Related papers A Word St emming Algorit hm for Hausa Language Applicat ions of St emming Algorit hms in Informat ion Ret rieval-A Review Rakesh Goswami Urdu language processing: a survey Preprocessing Techniques for Text Mining. *International Journal of Computer Science \& Communication Networks*, 5(1), 7–16.
- Mas Pintoko, B., & Muslim, K. (2018). Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naïve Bayes Classifier. *EProceedings of Engineering*, 5(3).
- Merinda Lestandy, Abdurrahim Abdurrahim, & Lailis Syafa'ah. (2021). Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(4), 802–808. <https://doi.org/10.29207/resti.v5i4.3308>
- Park, S. M., & Kim, Y. G. (2021). Root Cause Analysis Based on Relations Among Sentiment Words. *Cognitive Computation*, 13(4), 903–918. <https://doi.org/10.1007/s12559-021-09872-3>
- Rooney, J. J., & vanden Heuvel, L. N. (2004). Root Cause Analysis For Beginners. *Quality Progress*, 37(7), 45–56. [www.asq.org](http://www.asq.org)
- Selvam, B., Ravimaran, S., & Selvam, S. (2019). HYBRID APPROACH FOR ROOT CAUSE ANALYSIS OF ASPECTS USING ONLINE REVIEWS. *GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES*.