



# Implementasi Text Mining Dalam Penentuan Kinerja Layanan Antara Grab dan Gojek Berdasarkan Opini Masyarakat Menggunakan LDA

Tiwi Syafira<sup>\*</sup>, Ilka Zufria

Fakultas Sains dan Teknologi, Program Studi Ilmu Komputer, Universitas Islam Negeri Sumatera Utara, Medan  
Jl. William Iskandar Ps. V, Medan Estate, Kec. Percut Sei Tuan, Kabupaten Deli Serdang, Sumatera Utara, Indonesia

Email: <sup>1,\*</sup>tiwisyapira549@gmail.com, <sup>2</sup>ilkazufria@uinsu.ac.id

Email Penulis Korespondensi: tiwisyapira549@gmail.com

Submitted: 15/01/2024; Accepted: 29/01/2024; Published: 31/01/2024

**Abstrak**—Transportasi umum saat ini berbasis aplikasi dan internet, sehingga disebut transportasi online. Grab dan Gojek adalah dua penyedia jasa transportasi online yang ingin memberikan pelayanan terbaik. Pengguna memberikan tanggapan, pengalaman, kritik, dan saran melalui Twitter. Pada penelitian ini akan menggunakan algoritma LDA untuk memetakan topik yang sering dibicarakan para pengguna mengenai Gojek dan Grab. Latent Dirichlet Allocation (LDA) adalah algoritma pembelajaran tanpa pengawasan yang digunakan untuk mendeteksi topik dalam kumpulan dokumen teks. LDA mengasumsikan bahwa setiap dokumen terdiri dari campuran beberapa topik, dan bahwa setiap topik terdiri dari distribusi probabilitas atas kata-kata. LDA bekerja dengan memodelkan proses generatif dari dokumen. Proses ini dimulai dengan memilih topik untuk dokumen, lalu memilih kata-kata dari topik tersebut. Probabilitas memilih topik dan kata tertentu ditentukan oleh parameter yang dipelajari dari data. Data diperoleh melalui Twitter dengan kata kunci “#grabid” dan “#gojekindonesia”. Dari hasil penelitian didapatkan hasil pemetaan topik terbaik pada objek Gojek adalah 2 topik, yaitu promosi dan orderan. Sedangkan untuk objek Grab terbagi menjadi 4 topik, yaitu komunikasi, transaksi, orderan dan orderan.

**Kata Kunci:** Topik; Gojek; Grab; LDA; Layanan

**Abstract**—Public transportation is currently based on applications and the internet, so it is called online transportation. Grab and Gojek are two online transportation service providers who want to provide the best service. Users provide responses, experiences, criticism and suggestions via Twitter. This research will use the LDA algorithm to map topics that users frequently discuss regarding Gojek and Grab. Latent Dirichlet Allocation (LDA) is an unsupervised learning algorithm used to detect topics in a collection of text documents. LDA assumes that each document consists of a mixture of several topics, and that each topic consists of a probability distribution over words. LDA works by modeling the generative process of documents. This process begins by selecting a topic for the document, then selecting words from that topic. The probability of selecting certain topics and words is determined by parameters learned from the data. Data was obtained via Twitter with the keywords “#grabid” and “#gojekindonesia”. From the research results, it was found that the best topic mapping results for Gojek objects were 2 topics, namely promotions and orders. Meanwhile, Grab objects are divided into 4 topics, namely communication, transactions, orders and orders

**Keywords:** Topic; Gojek; Grab; LDA; Service

## 1. PENDAHULUAN

Transportasi online saat ini memiliki peranan sangat penting didalam kehidupan masyarakat sebagai sarana untuk berpindah dari satu tempat ke tempat lain tempat ke tempat lain. Seiring dengan perkembangan teknologi saat ini memberikan kontribusi terhadap kinerjasistem dalam semua aspek termasuk dalam aspek transportasi[1]. Transformasi transportasi umum ini membuka peluang bagi pelaku usaha untuk membuka usaha transportasi online. Adapun startup pertama kali yang hadir di Indonesia adalah Gojek dan diikuti startup GrabBike empat bulan setelahnya[2]. Ojek online saat ini menjadi tren transportasi publik terkini di kalangan masyarakat karena kemudahan dalam menggunakan jasa transportasi ini melalui perangkat aplikasi tanpa harus secara konvensional datang ke tempat konvensional untuk menggunakan layanan transportasi ini. Ada beberapa aplikasi transportasi online yang hadir di kota medan seperti Grab dan Gojek. Ojek online yang merupakan perlindungan keselamatan bagi penggunaan sepeda motor yang digunakan untuk kepentingan masyarakat dengan sebuah aplikasi[3]. Kedua transportasi online ini sama-sama identik dengan warna hijau. Meskipun layanan transportasi online mendapat respon positif namun terdapat masalah yang dihadapi yaitu banyaknya konsumen yang kecewa dan merasa tidak puas dengan pelayanan yang diberikan.

Hadirnya transportasi online merupakan fenomena yang cukup menyita perhatian, fenomena ojek online ini sangat digemari dikalangan masyarakat sebab karakteristik yaitu efiseien dalam hal waktu, cara pemesanan yang mudah, mampu melayani pada saat kita perlukan. Transportasi saat ini sangat menguntungkan, begitu cepatnya pertumbuhan para pelaku bisnis transportasi online atau sejenis seperti grab dan gojek. Grab dan gojek saat ini masih yang paling menonjol didalam persaingan pasar. Gojek bisa dikatakan yang pertama hadir di Indonesia dengan layanan yang semakin berkembang seiring dengan kebutuhan para konsumen[4].

Sistem transportasi online telah banyak digunakan oleh masyarakat selama beberapa tahun terakhir, di antaranya adalah Grab dan Gojek. Lalu muncul pertanyaan terkait perbandingan cara kerja grab dan gojek yang mana lebih bagus, pada dasarnya keduanya transportasi online ini, selain memiliki pengguna yang banyak, juga mempunyai keunggulan masing-masing. Popularitas keduanya bahkan bisa membuat mereka bertahan ditengah persaingan yang semakin ketat[5]. Ada beberapa perbedaan yang bisa dijadikan pertimbangan ketika harus



memilih antara kedua platform transportasi online, hal ini didasari pada selera kebutuhan masing-masing orang. Ada beberapa bahan perbandingan antara grab dan gojek mulai dari sistem aplikasi, layanan yang ditawarkan, tarif perjalanan, kemudahan pembayaran, diskon dan promo. Beberapa pengguna memberikan tanggapan tentang layanan yang diberikan, pengalaman, kritik maupun saran melalui sosial media salah satunya adalah Twitter. Apabila seketika kualitas layanan dinilai buruk, maka dapat menimbulkan perasaan dan tanggapan negatif dari pelanggannya. Bahkan ketika di biarkan tidak terselesaikan, dapat mengakibatkan penurunan kepercayaan, ketidakpuasan konsumen [6].

Seiring berjalannya waktu, grab dan gojek semakin diminati masyarakat apalagi dengan bantuan smartphone. Setiap pelanggan memiliki tingkat kepuasan berbeda terhadap layanan yang diberikan oleh grab dan gojek, sehingga selalu ada pro dan kontra berupa saran dan keluhan[7]. Pemrosesan terhadap saran dan keluhan kini dapat disampaikan secara real time melalui media sosial twitter, sehingga aplikasi tersebut dapat mengetahui tanggapan secara cepat tingkat kepuasan pelanggan terhadap pelayanan yang diberikan. Namun mengingat banyaknya jumlah pelanggan, tidak sedikit pula saran maupun keluhan yang ditujukan ke media sosial twitter yang diterima per hari[8]. Dilihat dari permasalahan yang ada, maka diperlukan sebuah solusi berupa analisis terhadap saran maupun keluhan yang diterima oleh perusahaan gojek maupun grab dari pelanggan yang mengirimkan melalui twitter sehingga dapat diketahui informasi sentimen pelanggan terhadap pelayanan yang telah diberikan secara cepat. Adapun salah satu cara terbaik untuk meningkatkan layanan adalah dengan melakukan analisis sentimen untuk mengetahui topik yang sering dibicarakan. Oleh karena itu data yang akan digunakan berupa sentimen di peroleh melalui proses crawling dengan menggunakan fasilitas Twitter API miring (italic)[9].

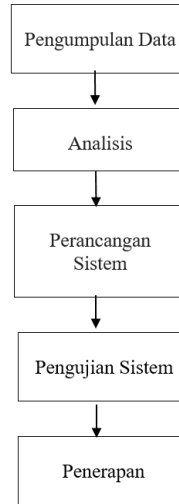
Analisis sentimen dan opinion mining adalah bidang studi yang menganalisis pendapat seseorang, sentimen seseorang, evaluasi seseorang, sikap seseorang dan emosi seseorang ke dalam bahasa tertulis. Ada beberapa penelitian terdahulu tentang analisis sentimen mengenai Gojek maupun Grab diantaranya. Hasil penelitian dari Immanuel O.D.P, et al. mengenai Implementasi Text Mining Untuk Analisis Opini Masyarakat Terhadap Kinerja Layanan Transportasi Online Dengan Analisis Faktor menghasilkan bahwa Aplikasi yang dibuat dapat mengumpulkan data terbaru secara real time. Algoritma Naïve Bayes Classifier pada aplikasi ini menghasilkan akurasi pada objek Grab 74,34%, pada objek Gojek 68,84% dan akurasi rata-rata dari kedua objek tersebut sebesar 71,59%[10]. Kemudian hasil penelitian lainnya dari Ilham Aminudin & Dyah Anggraini mengenai Analisis Peringkat Top Brand ojek online Menggunakan Jejaring Sosial Percakapan Twitter menghasilkan kesimpulan bahwa Setelah dilakukan perbandingan di dapat peringkat berdasarkan properti jaringan, yaitu grab unggul 4 properti sedangkan gojek hanya unggul 1 properti jaringan. Strategi keduanya sama ingin mempertahankan atau meningkatkan peringkatnya berdasarkan properti jaringan yaitu mereka harus membuat konten yang menarik agar dapat menarik perhatian pengguna sehingga otomatis akan menyebabkan banyak interaksi di dalamnya[11]. Ada juga hasil penelitian dari Alvian A.N., et al. mengenai Klasifikasi Tweets Masyarakat yang Membicarakan Layanan GoFood dan GoRide pada Gojek Dimedia Sosial Twitter Selama Masa Kenormalan Baru (New Normal) dengan Metode Naïve Bayes Classifier menyimpulkan bahwa Dari hasil penelitian metode Naïve Bayes Classifier sudah tepat dan sesuai digunakan untuk proses pengelompokan data tweets masyarakat yang membicarakan layanan GoFood dan GoRide pada Gojek dengan perolehan data dari proses crawling dari twitter[12]. Hasil penelitian dari Ni Luh Putu, Merawati, et al. mengenai Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation menyimpulkan bahwa Berdasarkan hasil analisis sentimen diperoleh 9.496 tweet dengan pembagian 8.996 tweet sehingga dapat disimpulkan bahwa lebih banyak wisatawan memberikan respon positif dari pada negatif terhadap pariwisata lombok[13]. Kemudian terdapat juga hasil penelitian dari M. Luvian mengenai Latent dirichlet allocation lda untuk mengetahui topik pembicaraan warganet twitter tentang omnibus law. Dari penelitian ini diperoleh jumlah topik sebanyak 5 dengan coherence score sebesar 0.5644. Kesimpulan dari setiap topiknya yaitu topik ke-1 mengenai penolakan UU Omnibus Law Cipta Kerja oleh sejumlah lapisan masyarakat dan kepala daerah. Topik ke-2 mengenai kekhawatiran masyarakat terhadap penyebaran virus Corona yang disebabkan oleh pelanggaran kebijakan PSBB yang dilakukan oleh demonstran dalam unjuk rasa omnibus law. Topik ke-3 mengenai dampak UU omnibus law cipta kerja terhadap kawasan hutan dan lingkungan. Topik ke-4 mengenai keistimewaan yang didapat warga asing karena omnibus law. Topik ke-5 kerusuhan yang terjadi pada saat demo tolak omnibus law di berbagai wilayah[14].

Dalam penelitian ini akan dilakukan penerapan Text Mining untuk menganalisis sumber data opini yang cukup banyak di media sosial Twitter. Twitter adalah media sosial gratis dan terpopuler serta menyediakan layanan jaringan yang memungkinkan pengguna untuk berbagi pendapat melalui pesan singkat atau sering dikenal dengan tweet[15]. Analisis sentimen adalah salah satu bagian ilmiah dalam mengklasifikasikan teks dan mengklasifikasikan teks berdasarkan pendapat dalam teks. Analisis sentimen adalah suatu proses dalam mendeteksi polaritas suatu teks secara kontekstual, ini menentukan suatu teks yang memiliki makna positif, negatif atau netral[16]. Analisis menggunakan LDA pada penelitian ini melalui beberapa tahapan diantaranya, crawling data, preprocessing, stemming, tokenization, stopword removal, classification, visualisasi dan evaluasi. Berdasarkan penelitian diatas, peneliti menggunakan LDA untuk mendeteksi topik – topik yang ada pada koleksi dokumen beserta proporsi kemunculan topik. Hasil dari penelitian yang dilakukan senantiasa bisa menjadi masukan terhadap pihak objek penelitian yaitu Gojek dan Grab untuk dapat meningkatkan layanannya masing-masing.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Penelitian berupa kumpulan fakta yang diperoleh dari suatu objek penelitian. Dalam hal ini objek penelitian adalah data tweet yang berkaitan dengan aplikasi transportasi online yang menggunakan hashtag #grabid dan #gojekindonesia. Dalam melakukan penelitian ini diperlukan perencanaan agar proses penelitian dapat berjalan sesuai dengan harapan. Gambar 1. dibawah ini merupakan diagram alur proses penelitian yang akan dilakukan.



Gambar 1. Diagram Alur Penelitian

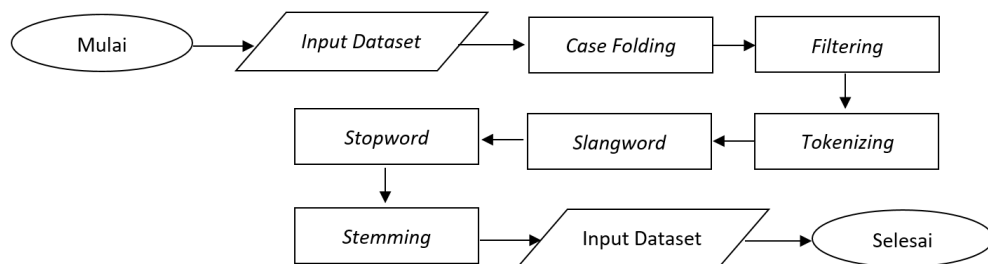
Pada gambar 1. Terdapat beberapa tahapan penelitian yang akan dilakukan, diantaranya adalah tahapan pengumpulan data, preprocessing data, vektorisasi, splitting data, pemodelan klasifikasi dan evaluasi.

### 2.2 Pengumpulan Data

Tahap pengumpulan data dilakukan dengan menggunakan crawling API Twitter dengan menggunakan library pada bahasa pemrograman Python. Data yang di crawl adalah tweet pengguna sejak Juli 2023 yang menyertakan hashtag #grabid dan #gojekindonesia. Dataset hasil crawling disimpan dalam format .csv untuk dilakukan proses selanjutnya.

### 2.3 Preprocessing Data

Setelah proses pengumpulan data dikarenakan dataset tersebut termasuk dalam unstructured data (data tidak terstruktur). Sebelum dilakukan analisis lebih lanjut, terlebih dahulu dataset melalui proses text preprocessing untuk menghilangkan serta mengatasi noisy data agar hasil perhitungan optimal. Adapun alur proses text preprocessing antara lain, case folding, filtering, tokenizing, dan stopword removal. Setelah data bersih maka dilakukan proses pelabelan secara manual. Berikut ini adalah diagram alur text preprocessing yang ditunjukkan pada Gambar 2.



Gambar 2. Preprocessing Data

### 2.4 Vektorisasi TF-IDF

Penelitian ini menerapkan algoritma pembobotan kata atau TF-IDF (Term Frequency-Inverse Document Frequency) yang berguna untuk melakukan pembobotan sebagai cara untuk memberi nilai pada kata (term) berdasarkan suatu dokumen/tweet. Tahapan pra-pemrosesan teks (text preprocessing) merupakan salah satu tahapan penting untuk mendapatkan model sentimen yang akurat karena teks komentar biasanya mengandung banya noise seperti slang word, tanda baca, kata tidak baku, singkatan, emoticon, dan sejenisnya sehingga

diperlukan metode prapemrosesan yang tepat[17]. Adapun rumus algoritma TF-IDF seperti pada persamaan (1) sebagai berikut:

$$w_{ij} = tf_{ij} \times \log\left(\frac{D+1}{df_{i+1}}\right) + 1 \tag{1}$$

Pada persamaan diatas  $w_{ij}$  merupakan bobot kata  $t$  terhadap dokumen yang ada,  $tf_{ij}$  adalah jumlah kemunculan kata  $t_i$ ,  $D$  adalah jumlah dokumen dan  $df_{ij}$  adalah jumlah kemunculan kata pada  $D_j$

### 2.5 Latent Dirichlet Allocation (LDA)

Pemodelan topik atau topic modelling adalah sebuah metode untuk mengekstrak dan merepresentasikan konteks yang digunakan sebagai sebuah arti kata dengan memanfaatkan komputasi statistik untuk sejumlah korpus yang besar dari teks. Tujuan pemodelan topik yaitu menemukan topik dan kata-kata yang terkandung dalam korpus tersebut. Dalam melakukan pemodelan topik, dapat diimplementasikan metode pengelompokan berdasarkan ukuran kedekatan (kemiripan) suatu data, salah satunya dengan menggunakan LDA. Prinsip LDA mengklasifikasikan dokumen sesuai topik terkait. Prinsip algoritma LDA bahwa setiap kata memiliki makna semantik. LDA adalah model probabilistik generatif dan algoritma unsupervised learning dimana dokumen direpresentasikan sebagai campuran acak topik-topik.[18]. Setelah melakukan tahap pra-processing, hasilnya digunakan untuk membuat pemodelan topik dengan LDA. Sebelum pemodelan topik, terlebih dahulu dilakukan pembentukan dictionary dan corpus. Dictionary merupakan format data yang mengandung himpunan kata unik dengan nomor indeks masing-masing corpus merupakan format data berbentuk bag-of-words reference yang akan digunakan untuk pembentukan model[19]. Proses pemodelan topik LDA dilakukan dengan menentukan jumlah topik yang menjadi kelompok klusterisasi kata dan penentuan jumlah passes dalam topic modeling yang merupakan jumlah iterasi proses pembelajaran dari training model. Validasi pemodelan topic dengan topic coherence bertujuan untuk memastikan model yang dibentuk dari proses topik modeling pada dokumen yang dihasilkan memiliki nilai probabilitas tertinggi. Hasil pemodelan topik LDA dengan jumlah topik yang sesuai berdasarkan nilai topic coherence tiap model[20].

## 3. HASIL DAN PEMBAHASAN

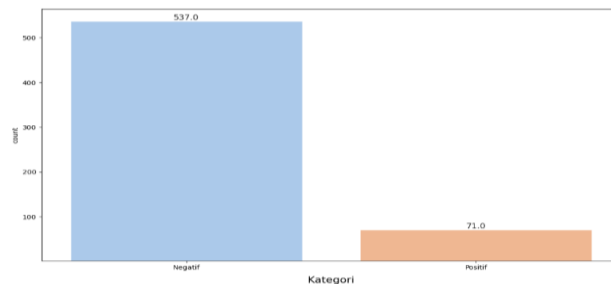
### 3.1 Persiapan Data

Pada penelitian ini penulis akan menganalisis data sentimen terhadap layanan aplikasi Gojek dan Grab pada tahun 2023. Tahap pertama pada penelitian adalah melakukan pengumpulan data. Proses pengumpulan data opini masyarakat mengenai layanan aplikasi Gojek dan Grab pada tahun 2023 dilakukan pada media sosial twitter dengan menggunakan library python sebagai scrapper data.

Proses scrapping data dilakukan dengan kata kunci nama kedua perusahaan besar tersebut, yaitu #gojekindonesia dan #grabid. Data yang diambil merupakan tweet pada bulan Juli tahun 2023. Hasil dari proses scrapping akan diubah ke dalam ekstensi .csv untuk kemudian diolah menggunakan bahasa pemrograman python. Jumlah data yang diperoleh dari hasil scrapping adalah 1208 data. Dimana 608 data berasal dari kata kunci aplikasi Gojek yaitu #gojekindonesia dan 600 data berasal dari kata kunci aplikasi Grab yaitu #grabid.

#### a. Persebaran data ulasan pada aplikasi Gojek

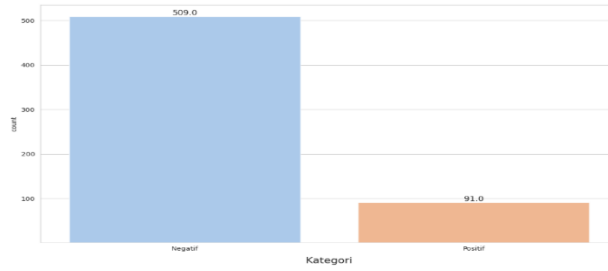
Pada aplikasi gojek diambil sebanyak 608 data. Keseluruhan data dikelompokkan secara manual kategori dari setiap sentimen sehingga diperoleh jumlah sentimen negatif pada data yang digunakan adalah sebanyak 537 dan sentimen positif berjumlah 71 sentimen. Gambar 3 merupakan grafik kategori sentimen dari aplikasi Gojek.



**Gambar 3.** Hasil splitting ulasan negatif dan positif pada Gojek

#### b. Persebaran data ulasan pada aplikasi Grab

Pada aplikasi Grab diambil sebanyak 600 data. Keseluruhan data dikelompokkan secara manual kategori dari setiap sentimen sehingga diperoleh jumlah sentimen negatif pada data yang digunakan adalah sebanyak 509 dan sentimen positif berjumlah 91 sentimen. Dibawah ini merupakan gambar 4 grafik kategori sentimen dari aplikasi Grab.



**Gambar 4.** Hasil splitting ulasan negatif dan positif pada Grab

Hasil pemetaan topik menggunakan algoritma Latent Dirichlet Allocation akan dikelompokkan ke dalam beberapa kelas. Berdasarkan referensi yang diperoleh, dirangkum 4 buah topik yang umum dibahas mengenai Gojek dan Grab. Adapun topik yang dimaksud tersaji pada tabel 1.

**Tabel 1.** Kategori topik

Topik	Kata Kunci
Promo	Voucher, kode, promo
Transaksi	Bayar, pakai, transfer
Orderan	Order, cancel, pickup, pesan
Komunikasi	Cek, dm, min, halo, balas

### 3.2 Text preprocessing

Setelah proses pengumpulan data Twitter selesai, terlebih dahulu dilakukan text preprocessing. Hal ini dilakukan dikarenakan dataset yang digunakan termasuk dalam data unstructured data (data tidak terstruktur). Adapun beberapa tahap text preprocessing yang dilakukan pada penelitian kali ini sebagai berikut :

- Case Folding adalah tahap perubahan semua huruf campuran baik itu uppercase, atau lowercase, menjadi lowercase (huruf kecil) semua.
- Filtering adalah tahap pembersihan data dari tanda baca, simbol atau elemen yang tidak dibutuhkan seperti URL.
- Tokenizing adalah tahap pemecahan teks menjadi kata-kata, seperti “saya pergi ke sekolah” menjadi “saya”, “pergi”, “ke”, “sekolah”.
- Stopword Removal adalah tahap penghapusan kata sambung seperti “ke” ,”di” ,”dan” , “dia”, ”kami”, ”aku”, ”saya”.

Gambar 2. merupakan contoh sentimen sebelum dilakukan text-preprocessing dan setelah pre-processing.

**Tabel 2.** Hasil Text Preprocessing

Sentimen	
Input	Lah mau promo gofood tp cm buat paylater doang yg gede, kembalikan yg buat gopay juga dong elah. @gojekindonesia
Output	['promo', 'gofood', 'paylater', 'doang', 'gede', 'kembali', 'gopay', 'elah']

### 3.3 Vektorisasi TF-IDF

Manfaat dari proses vektorisasi adalah untuk memberi nilai pada kata dalam suatu dokumen atau tweet agar dapat diolah dan dicari polanya ketika masuk dalam tahap klasifikasi. Proses vektorisasi dengan menggunakan algoritma perhitungan pembobotan kata atau TF-IDF yaitu dengan menganalisis sampel data tweet kemudian melakukan perhitungan berdasarkan rumus (1). Berikut ini merupakan hasil perhitungan TF-IDF terhadap 3 sampel data tweet yang dapat dilihat pada Tabel 3.

**Contoh :**

**Data Latih :**

D1 = ['pilih', 'ojol', 'baik']

D2 = ['driver', 'bijak', 'terus', 'suka', 'penumpang']

D3 = ['jalan', 'terus']

D4 = ['bodoh', 'tolak', 'orang', 'langgar', 'selalu', 'rekam']

D5 = ['rating', 'salah', 'rekam', 'pasang', 'ojol', 'salah', 'indonesia']

**Contoh hasil Vektorisasi:**

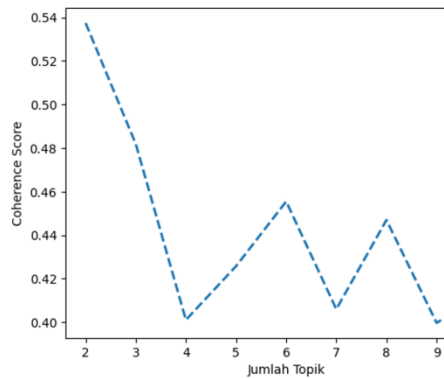
**Tabel 3.** Hasil Vektorisasi

Term	TF-IDF				
	D1	D2	D3	D4	D5
pilih	2.099	0	0	0	0

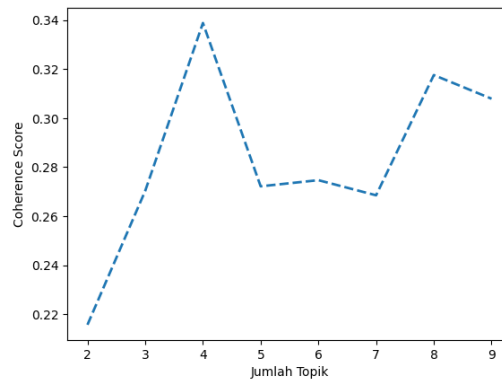
Term	TF-IDF				
	D1	D2	D3	D4	D5
ojol	1.693	0	0	0	1.693
baik	2.099	0	0	0	0
driver	0	2.099	0	0	0
bijak	0	2.099	0	0	0
terus	0	1.693	1.693	0	0
suka	0	2.099	0	0	0
penumpang	0	2.099	0	0	0
driver	0	0	2.099	0	0
bodoh	0	0	0	2.099	0
tolak	0	0	0	2.099	0
orang	0	0	0	2.099	0
langgar	0	0	0	2.099	0
selalu	0	0	0	2.099	0
rekam	0	0	0	2.099	0
rating	0	0	0	0	2.099
rekam	0	0	0	0	2.099
pasang	0	0	0	0	2.099
salah	0	0	0	0	4.198
indonesia	0	0	0	0	2.099

**3.4 Latent Dirichlet Allocation (LDA)**

Dalam menentukan hasil pemodelan dapat dilakukan dengan melihat pada visualisasi dari grafik coherencescore. Coherence score merupakan suatu ukuran yang akan digunakan untuk mengevaluasi Topic Modeling, dimana jika coherencescore topik yang tinggi maka model yang dihasilkan tersebut yang baik. Grafik dari coherence score yang dihasilkan terdapat naik dan turun, peneliti akan menunjukkan beberapa perbandingan dalam grafiknya yaitu dimana topik dimulai dari 2-9 yang dapat dilihat pada gambar dibawah ini.



**Gambar 5.** Grafik coherence untuk objek Gojek



**Gambar 6.** Grafik coherence untuk objek Grab

Berdasarkan Gambar 5. dan Gambar 6. diatas didapatkan informasi bahwa pada grafik coherence score diatas memiliki pola yang berulang dan semakin banyak limit topiknya maka semakin tinggi nilai coherence yang dihasilkan. Kemudian dari perbandingan grafik tersebut maka peneliti memutuskan untuk menggunakan topik sebanyak 2 untuk studi kasus Gojek dan topik sebanyak 5 untuk studi kasus Grab yang memiliki nilai coherence

terbesar yaitu sebesar 0.561480 dan 0.310411, dengan begitu jumlah topik tersebut yang akan menjadi acuan untuk membuat model selanjutnya. Bini merupakan tabel coherence score yang dihasilkan:

**Tabel 4.** Tabel coherence Gojek

JumlahTopik	NilaiCoherence
2	0.561480
3	0.469883
4	0.430672
5	0.441090
6	0.383735
7	0.409289
8	0.416514
9	0.420244

**Tabel 5.** Tabel coherence Grab

JumlahTopik	NilaiCoherence
2	0.215775
3	0.270331
4	0.338756
5	0.272194
6	0.274700
7	0.268537
8	0.317535
9	0.307898

### 3.5 Hasil PemetaanTopik

#### 3.5.1 Untuk studi kasus Gojek

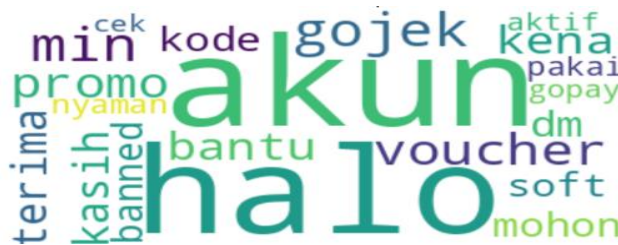
##### 1. Topik 1

Pada topik 1 untuk aplikasi gojek, didapatkan model LDA sebagai berikut.

$$\begin{aligned}
 & '0.043* "akun" + 0.040* "gojek" + 0.036* "voucher" + 0.032 \\
 & * "halo" + 0.032* "promo" + 0.028* "kasih" + 0.027* "terima" + \\
 & 0.027* "bantu" + 0.026* "kena" + 0.021* "kode"
 \end{aligned}$$

**Gambar 7.** Model LDA Gojek Topik 1

Berdasarkan model LDA yang diperoleh untuk topik ke-1 pada studi kasus Gojek pada gambar diatas, maka dapat disimpulkan bahwa pada model LDA topik ke-1 yang banyak muncul dan berkaitan satu sama lain yaitu mengenai **promosi**. Kemudian dari model tersebut juga dapat ditampilkan dengan bentuk wordcloud sebagai berikut.



**Gambar 8.** Wordcloud topik 1 Gojek

##### 2. Topik 2

Pada topik 2 untuk aplikasi gojek, didapatkan model LDA sebagai berikut

$$\begin{aligned}
 & '0.026* "min" + 0.019* "driver" + 0.018* "dm" + 0.014* "gopay" \\
 & + 0.013* "order" + 0.013* "pakai" + 0.011* "cek" + 0.009 \\
 & * "tolong" + 0.008* "kali" + 0.008* "halo"
 \end{aligned}$$

**Gambar 9.** Model LDA Gojek Topik 2

Berdasarkan model LDA yang diperoleh untuk topik ke-2 pada studi kasus Gojek pada gambar diatas, maka dapat disimpulkan bahwa pada model LDA topik ke-2 yang banyak muncul dan berkaitan satu sama lain yaitu

mengenai **orderan**. Kemudian dari model tersebut juga dapat ditampilkan dengan bentuk wordcloud sebagai berikut



Gambar 10. Wordcloud topik 2 Gojek

### 3.5.2 Untuk studi kasus Grab

#### 1. Topik 1

Pada topik 1 untuk aplikasi Grab, didapatkan model LDA sebagai berikut.

$$0.028 * "driver" + 0.018 * "min" + 0.016 * "pesan" + 0.015 * "tolong" + 0.014 * "grab" + 0.010 * "dm" + 0.010 * "grabfood" + 0.010 * "drivernya" + 0.010 * "order" + 0.009 * "cek"$$

Gambar 11 Model LDA Grab Topik 1

Berdasarkan model LDA yang diperoleh untuk topik ke-1 pada studi kasus Grab pada gambar diatas, maka dapat disimpulkan bahwa pada model LDA topik ke-1 yang banyak muncul dan berkaitan satu sama lain yaitu mengenai **komunikasi**. Kemudian dari model tersebut juga dapat ditampilkan dengan bentuk wordcloud sebagai berikut



Gambar 12. Wordcloud topik 1 Grab

#### 2. Topik 2

Pada topik 2 untuk aplikasi Grab, didapatkan model LDA sebagai berikut.

$$0.017 * "grab" + 0.012 * "pakai" + 0.010 * "driver" + 0.010 * "terima" + 0.010 * "bayar" + 0.009 * "aplikasi" + 0.006 * "potong" + 0.006 * "kali" + 0.006 * "kasih" + 0.006 * "min"$$

Gambar 13. Model LDA Grab Topik 2

Berdasarkan model LDA yang diperoleh untuk topik ke-2 pada studi kasus Grab pada gambar diatas, maka dapat disimpulkan bahwa pada model LDA topik ke-2 yang banyak muncul dan berkaitan satu sama lain yaitu mengenai **transaksi**. Kemudian dari model tersebut juga dapat ditampilkan dengan bentuk wordcloud sebagai berikut



Gambar 14. Wordcloud topik 2 Grab

### 3. Topik 3

Pada topik 3 untuk aplikasi Grab, didapatkan model LDA sebagai berikut

$$'0.030*\text{"driver"} + 0.023*\text{"order"} + 0.018*\text{"pakai"} + 0.018*\text{"grab"} + 0.012*\text{"enggak"} + 0.010*\text{"jam"} + 0.010*\text{"makan"} + 0.009*\text{"min"} + 0.009*\text{"grabfood"} + 0.009*\text{"pesan"}$$

Gambar 15. Model LDA Grab Topik 3

Berdasarkan model LDA yang diperoleh untuk topik ke-3 pada studi kasus Grab pada gambar diatas, maka dapat disimpulkan bahwa pada model LDA topik ke-3 yang banyak muncul dan berkaitan satu sama lain yaitu mengenai **orderan**. Kemudian dari model tersebut juga dapat ditampilkan dengan bentuk wordcloud sebagai berikut



Gambar 16. Wordcloud topik 3 Grab

### 4. Topik 4

Pada topik 4 untuk aplikasi Grab, didapatkan model LDA sebagai berikut.

$$'0.019*\text{"cancel"} + 0.016*\text{"driver"} + 0.010*\text{"grab"} + 0.010*\text{"halo"} + 0.009*\text{"order"} + 0.008*\text{"pakai"} + 0.007*\text{"tunggu"} + 0.007*\text{"motor"} + 0.006*\text{"pesan"} + 0.006*\text{"pickup"}$$

Gambar 17. Model LDA Grab Topik 4

Berdasarkan model LDA yang diperoleh untuk topik ke-4 pada studi kasus Grab pada gambar diatas, maka dapat disimpulkan bahwa pada model LDA topik ke-4 yang banyak muncul dan berkaitan satu sama lain yaitu mengenai **orderan**. Kemudian dari model tersebut juga dapat ditampilkan dengan bentuk wordcloud sebagai berikut



Gambar 18. Wordcloud topik 4 Grab

## 4. KESIMPULAN

Berdasarkan hasil analisis pemodelan topik dengan metode Latent Dirichlet Allocation (LDA) pada data ulasan pengguna aplikasi Gojek dan Grab pada tahun 2023, diperoleh bahwa aplikasi Grab memiliki cakupan layanan yang lebih luas dibandingkan aplikasi Gojek. Hal ini ditunjukkan dengan adanya empat topik utama pada aplikasi Grab, yaitu komunikasi, transaksi, orderan, dan orderan. Sedangkan aplikasi Gojek hanya memiliki dua topik utama, yaitu promosi dan orderan.

## UCAPAN TERIMAKASIH

Saya ingin mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dalam penelitian ini. Terimakasih kepada keluarga saya atas dukungan moral dan doa-doanya yang tak ternilai. Terima kasih kepada dosen pembimbing yang telah memberikan bimbingan dan arahan, nasihat, dan waktunya selama proses penelitian ini. Saya juga ingin mengucapkan terima kasih kepada teman-teman sejawat



yang telah berbagi pengetahuan dan pengalaman mereka, serta semua pihak yang telah memberikan akses dan bantuan dalam pengumpulan data. Penelitian ini tidak mungkin terlaksana tanpa dukungan dari semua pihak ini. Semoga hasil penelitian ini dapat bermanfaat bagi semua. Sekali lagi, terima kasih banyak atas segala dukungan yang telah diberikan.

## REFERENCES

- [1] E. Hernawan and Andy, “Analisis Faktor- Faktor Yang Mempengaruhi Pelanggan Gojek Dan Grab Online Di Jakarta,” *J. Ekon. Dan Bisnis - Vol. 17. No. 1*, vol. 17, no. 1, pp. 1–13, 2019.
- [2] W. San Taslim, “Analisis preferensi konsumen terhadap jasa angkutan online di Pontianak (Studi kasus: Gojek vs Grab),” *Obis*, vol. 3, no. 1, pp. 13–18, 2020.
- [3] A. Alfonsius, “Pelayanan Transportasi Online Di Era New Normal,” *J. Account. Manag. Innov.*, vol. 4, no. 2, pp. 91–100, 2020.
- [4] J. Wahani, J. D. D. M. Poluan, and J. Grace, “J . Wahani ., J . D . D . Massie ., J . G . Poluan DI KOTA MANADO ANALYSIS OF SERVICE QUALITY COMPARISON BETWEEN GO-JEK AND GRAB IN MANADO CITY Oleh : Jurnal EMBA Vol . 9 No . 3 Juli 2021 , Hal . 1762- 1774 J . Wahani ., J . D . D . Massie ., J . G . Polu,” *Anal. PERBANDINGAN KUALITAS PELAYANAN ANTARA GO-JEK DAN GRAB DI KOTA Manad.*, vol. 9, no. 3, pp. 1762–1774, 2021.
- [5] Y. Astutik and W. Suharso, “Evaluasi Kualitas Layanan Aplikasi Go-Jek Dan Grab Dengan Metode Ahp,” *Univ. Muhammadiyah Jember*, 2019.
- [6] I. Setyawan, “Kualitas Layanan Last-Mile Delivery : Studi Komparasi Dua Layanan Pesan-Antar Makanan,” *Kualitas Layanan Last-Mile Deliv. Stud. Komparasi Dua Layanan Pesan-Antar Makanan*, vol. 22, no. 3, pp. 2050–2060, 2022, doi: 10.33087/jjubj.v22i3.2692.
- [7] R. Septiani and N. Nurhadi, “Peran Mediasi Kepuasan Pelanggan Pada Pengaruh E-Service Quality, Persepsi Harga, Dan Promosi Penjualan Terhadap Loyalitas Pelanggan,” *J. Fokus Manaj. Bisnis*, vol. 10, no. 2, p. 249, 2020, doi: 10.12928/fokus.v10i2.2886.
- [8] Diana Khuntari, “Analisis Pengalaman Pengguna Aplikasi Gojek dan Grab dengan Pendekatan User Experience Questionnaire,” *J. Tek. Inform. dan Sist. Inf.*, vol. 8, no. April, pp. 275–286, 2022.
- [9] J. Waworundeng, G. Sandag, S. V. Ngeloh, and A. Lalong, “Analisis Tingkat Kepuasan Pelanggan terhadap Layanan Grab dan Gojek di Masa Pandemi Covid-19 Analysis of Customer Satisfaction Levels with Grab and Gojek Services during the Covid-19 Pandemic,” *Cogito Smart J.*, vol. 8, no. 1, pp. 111–121, 2022.
- [10] K. R. Prilianti, “Implementation of Text Mining for Online Transportation Service Analysis with Factor Analysis,” *Simantec*, vol. 8, no. 2, pp. 1–9, 2020.
- [11] I. Aminudin and D. Anggraini, “Analisis Peringkat Top Brand Ojek Online Menggunakan Jejaring Sosial Percakapan Twitter,” *J. Ilm. Inform. Komput.*, vol. 24, no. 2, pp. 88–104, 2019, doi: 10.35760/ik.2019.v24i2.2365.
- [12] A. A. Nabhan, B. Rahayudi, and D. E. Ratnawati, “Klasifikasi Tweets Masyarakat yang Membicarakan Layanan GoFood dan GoRide pada GoJek Dimedia Sosial Twitter Selama Masa Kenormalan Baru ( New Normal ) dengan Metode Naïve Bayes,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 7, pp. 3018–3025, 2021.
- [13] R. Sistem et al., “JURNAL RESTI Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok,” vol. 1, no. 10, pp. 123–131, 2021.
- [14] M. Luvian chisni chilmi, “Latent dirichlet allocation (LDA) untuk mengetahui topik pembicaraan warganet twitter tentang omnibus law,” *Univ. Islam Negeri Syarif Hidayatullah*, pp. 1–131, 2021.
- [15] S. Styawati, N. Hendrastuty, and A. R. Isnain, “Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine,” *J. Inform. J. Pengemb. IT*, vol. 6, no. 3, pp. 150–155, 2021, doi: 10.30591/jpit.v6i3.2870.
- [16] I. Noor Kabiru and P. Kencana Sari, “Analisa Konten Media Sosial E-Commerce Pada Instagram Menggunakan Metode Sentimen Analysis Dan Lda-Based Topic Modeling (Studi Kasus: Shopee Indonesia) Analysis of Content Social Media E-Commerce in Instagram Using Sentiment Analysis and Lda Based Topic M,” *e-Proceeding Manag.*, vol. 6, no. 1, p. 12, 2019.
- [17] B. B. Baskoro, I. Susanto, and S. Khomsah, “Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF (Studi Kasus: Ulasan Pelanggan Pada Situs TRIPADVISOR),” *INISTA (Journal Informatics Inf. Syst. Softw. Eng. Appl.)*, vol. 3, no. 2, pp. 21–29, 2021, doi: 10.20895/INISTA.V3.
- [18] N. Novarian, S. Khomsah, and A. B. Arifa, “Topic Modeling Tugas Akhir Mahasiswa Fakultas Informatika Institut Teknologi Telkom Purwokerto Menggunakan Metode Latent Dirichlet Allocation Nathanael,” *LEDGER J. Inform. Inf. Technol.*, vol. 2, no. 1, pp. 1–14, 2023.
- [19] Y. Matira, “Pemodelan Topik pada Judul Berita Online Detikcom Menggunakan Latent Dirichlet Allocation,” *Univ. Gadjah Mada*, vol. 4, no. 1, pp. 53–63, 2023, doi: 10.20956/ejsa.vi.24843.
- [20] S. Yoga, F. Nurul Isnaini, and O. Pamulatsih Dwi, “Pemodelan Topik Penelitian Bidang Keperawatan Indonesia pada Repository Jurnal Sinta Menggunakan Metode Topic Modelling LDA (Latent Dirichlet Allocation),” *Semin. Nas. Teknol. Inf. Ilmu Komput.*, vol. 1, pp. 90–102, 2020.