

## Analisis Sentimen Pengguna Transportasi Online Terhadap Layanan Grab Indonesia Menggunakan Multinomial Naive Bayes Classifier

Sartika Mandasari<sup>1</sup>, B. Herawan Hayadi<sup>2</sup>, Rudi Gunawan<sup>3</sup>

<sup>1,2</sup> Ilmu Komputer, Universitas Potensi Utama

<sup>3</sup>Sistem Informasi, STMIK Triguna Dharma

Email: [sartikamandasari12@gmail.com](mailto:sartikamandasari12@gmail.com), [b.herawan.hayadi@gmail.com](mailto:b.herawan.hayadi@gmail.com), [rudi\\_gunawan8899@yahoo.com](mailto:rudi_gunawan8899@yahoo.com)

Email Penulis Korespondensi: [sartikamandasari12@gmail.com](mailto:sartikamandasari12@gmail.com)

### Article History:

Received Jun 20<sup>th</sup>, 2022

Revised Jul 01<sup>th</sup>, 2022

Accepted Jul 16<sup>th</sup>, 2022

### Abstrak

Ojek *online* telah merambah dunia transportasi di Indonesia. Terobosan baru yang semakin diminati banyak orang ini, tidak hanya menyediakan transportasi angkutan penumpang saja namun juga melayani jasa kurir untuk pemesanan makanan, jasa pengiriman barang, dokumen, dan berbelanja. Banyaknya pelayanan yang diberikan perusahaan ojek *online*, semakin banyak pula opini yang dilontarkan masyarakat melalui *twitter*, mengenai kualitas dari setiap jenis layanan yang diberikan oleh perusahaan ojek *online*. Opini yang memiliki sentimen tersebut akan dianalisis sehingga dapat diketahui layanan mana yang mendapatkan sentimen positif, negatif, dan netral. Oleh sebab itu, diperlukan sebuah pendekatan yang dapat menganalisis sentimen masyarakat terhadap kualitas dari setiap layanan ojek online. Pada penelitian ini metode yang digunakan yaitu term frequency (tf) dan *multinomial naive bayes classifier*. Tahapan keseluruhan metode yang digunakan pada penelitian ini adalah *preprocessing* (*cleaning, case folding, tokenisasi, convert negation, stopword removal, stemming, dan normalisasi*), perhitungan frekuensi kemunculan kata (tf), dan klasifikasi sentimen. Hasil dari penelitian ini adalah mengklasifikasikan *tweet* ke dalam sentimen positif, negatif, netral dan mengetahui kualitas dari setiap jenis layanan ojek *online*. Dengan menggunakan algoritma dan metode ini, akurasi yang didapat adalah sebesar 86,57%. Pendekatan ini diharapkan akan sangat membantu pihak perusahaan ojek *online* untuk memperbaiki kualitas dari setiap jenis layanan yang ada.

**Kata Kunci :** Analisis Sentimen, *Multinomial Naive Bayes, Ojek Online, Term Frequency, Text Mining*

### Abstract

*Online motorcycle taxis have penetrated the world of transportation in Indonesia. This new breakthrough, which is increasingly in demand by many people, not only provides passenger transportation but also provides courier services for ordering food, delivering goods, documents, and shopping. The number of services provided by online motorcycle taxi companies, the more opinions are expressed by the public via Twitter, regarding the quality of each type of service provided by online motorcycle taxi companies. Opinions that have these sentiments will be analyzed so that it can be seen which services get positive, negative, and neutral sentiments. Therefore, we need an approach that can analyze public sentiment on the quality of each online motorcycle taxi service. In this research, the methods used are term frequency (tf) and multinomial naive bayes classifier. The overall stages of the method used in this research are preprocessing (cleaning, case folding, tokenization, convert negation, stopword removal, stemming, and normalization), calculation of word occurrence frequency (tf), and sentiment classification. The results of this study are to classify tweets into positive, negative, neutral sentiments and determine the quality of each type of online motorcycle taxi service. By using this algorithm and method, the accuracy obtained is 86.57%. This approach is expected to be very helpful for online motorcycle taxi companies to improve the quality of each type of service available.*

**Keyword :** Sentiment Analysis, *Multinomial Naive Bayes, Ojek Online, Term Frequency, Text Mining*

## 1. PENDAHULUAN

Analisis sentimen atau opinion mining merupakan proses memahami, mengekstrak, dan mengolah data tekstual secara otomatis untuk mendapatkan suatu informasi sentimen yang terkandung dalam suatu kalimat opini [1]. Secara umum, analisis sentimen bertujuan untuk menentukan sikap pembicara atau penulis terhadap suatu topik atau keseluruhan polaritas kontekstual pada suatu dokumen [2]. Sikap dapat berupa penilaian atau evaluasi, sisi emosional penulis pada saat menulis atau efek komunikasi emosional yang penulis inginkan terhadap pembacanya. Twitter sebagai salah satu jejaring sosial yang interaktif memungkinkan penggunaannya untuk mengkritisi suatu isu maupun sebuah fasilitas pelayanan secara real time. Masyarakat yang semula membutuhkan waktu lama untuk menyampaikan aspirasinya kini dapat melakukannya dengan mudah berkat kehadiran teknologi ini.

Salah satu transportasi online yang ramai dibicarakan oleh masyarakat adalah ojek online. Ojek online atau sering disebut Ojol merupakan transformasi dari ojek konvensional yang biasanya bertempat di pangkalan untuk menunggu pelanggan. Seiring berjalannya waktu, Ojol ini semakin diminati masyarakat apalagi dengan bantuan smartphone. Pelayanan yang disediakan tidak hanya transportasi angkutan penumpang saja, namun juga melayani jasa kurir untuk pemesanan makanan, jasa pembersih untuk bersih-bersih rumah dan kantor, pengiriman barang, dokumen, berbelanja, dan lain sebagainya.

Grab Indonesia sebagai salah satu perusahaan ojek online terkemuka di Indonesia memiliki jumlah pelanggan yang cukup besar dan mencakup hampir seluruh wilayah di Indonesia. Setiap pelanggan memiliki tingkat kepuasan berbeda terhadap layanan yang diberikan oleh Grab Indonesia, sehingga selalu ada pro dan kontra berupa saran dan keluhan. Pemrosesan terhadap saran dan keluhan kini dapat disampaikan secara real time melalui akun Twitter @GrabID, sehingga Grab Indonesia 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 akun @GrabID yang diterima per hari.

Dilihat dari permasalahan yang ada, maka diperlukan sebuah solusi berupa analisis terhadap saran maupun keluhan yang diterima oleh perusahaan ojek online Grab Indonesia dari pelanggan yang mengirimkan tweet ke akun user @GrabID pada media sosial Twitter sehingga dapat diketahui informasi sentimen pelanggan terhadap pelayanan yang telah diberikan secara cepat. Data dapat di peroleh melalui proses crawling dengan menggunakan fasilitas Twitter API.

Adapun permasalahan pengklasifikasian sebuah kalimat sentimen ke dalam kelas-kelas tertentu dapat diselesaikan dengan metode Multinomial Naive Bayes Classifier. Metode ini memiliki beberapa kelebihan antara lain, sederhana, cepat, dan berakurasi tinggi yang memiliki keunggulan dalam memproses teks. Metode MNBC untuk klasifikasi atau kategorisasi teks menggunakan atribut kata yang muncul dalam suatu dokumen sebagai dasar klasifikasinya.

## 2. METODOLOGI PENELITIAN

Metode yang diajukan pada penelitian ini yaitu untuk mengetahui tanggapan masyarakat melalui analisis sentimen mengenai kualitas dari beberapa layanan ojek online yaitu Grab Indonesia yang terdiri dari beberapa proses. Tahapan yang akan dilakukan adalah sebagai berikut: kumpulan opini memasuki proses text processing dimana opini akan dibersihkan dari simbol-simbol yang tidak penting dikarenakan simbol-simbol tersebut tidak berkaitan dan tidak memiliki arti penting dalam opini tersebut, contohnya @, RT, HTML, kata sambung, dan lain sebagainya. Setelah itu akan melewati proses stemming dengan menggunakan algoritma Nazief dan Andriani, selanjutnya akan melewati tahapan sorting dan grouping hingga proses klasifikasi dengan algoritma MNBC yang menentukan apakah kategori opini tersebut termasuk dalam kelompok negatif, positif, atau netral.

Algoritma Multinomial Naive Bayes Classifier ini menggabungkan teorema - teorema yang nantinya akan diterapkan pada permasalahan klasifikasi teks secara otomatis [3].

Dalam penelitian ini data yang digunakan diambil dari kumpulan tweet mention @GrabID berdasarkan 3 layanan yaitu GrabCar, GrabBike dan GrabFood melalui twitter search API. Banyaknya data yang terdapat dalam penelitian ini adalah 1050 data, yang pembagiannya 700 untuk data latih yang akan digunakan untuk membentuk model analisis sentimen, yang secara manual diberi label positif, negatif dan netral sesuai dengan sifat opini dari data yang telah didapatkan. Sedangkan 350 data dari hasil crawling akan digunakan sebagai data uji yang akan dilabeli sentimen secara otomatis oleh sistem.

Sebelum melakukan klasifikasi seluruh data training dan testing harus melewati tahap preprocessing [4]. Tahapan ini bertujuan untuk mempersiapkan dokumen teks yang tidak terstruktur menjadi data terstruktur yang siap digunakan untuk proses selanjutnya.

Setelah melalui proses perhitungan kemunculan kata pada tiap dokumen akan melalui tahap pengklasifikasian, pada proses ini akan digunakan algoritma *Multinomial Naive Bayes Classifier* [5].

### 3. HASIL DAN PEMBAHASAN

Berdasarkan analisis dan perancangan yang telah dibuat dalam menganalisis sentimen pada media sosial Twitter, sistem ini dibuat menggunakan bahasa pemrograman php dan javascript[6]. Dan untuk perancangan sistem membutuhkan beberapa fasilitas pendukung, berikut merupakan fasilitas pendukung yang dibutuhkan oleh sistem, baik perangkat keras maupun perangkat lunak.

#### 3.1 Pengujian Sistem

Tahapan uji coba berfungsi untuk menguji apakah sistem yang dibangun sudah berjalan dengan baik atau belum dengan membandingkan hasil yang diperoleh dari sistem dengan beberapa parameter. Untuk melakukan pemeriksaan terhadap sistem yang sudah dibangun, maka dilakukan pengujian sistem yang terdiri dari proses pelatihan dan pengujian model.

Adapun untuk menjelaskan proses pelatihan data training dan testing, diberikan tiga contoh data training dan satu contoh data testing dari total keseluruhan data training sebanyak 700 data tweet dan 350 data testing yang sebelumnya telah melewati proses preprocessing.

Tabel 1. Contoh Data Training

User (@)	Sebelum <i>Preprocessing</i>	Setelah <i>Preprocessing</i>	Sentimen Manual
14Maman	@GrabID Ayo para pelanggan pesan GrabFood nya sekarang hanya di mallpenvil ya.. barang kali gue yang antar	ayo langgan pesan grabfood nya mall penvil barang saya	Positif
_Ramdan9	@GrabID download aplikasi GrabFood merchant dimana di playstore kok gak ada ya?	download aplikasi grabfood merchant mana playstore tidak ada	Netral
amolohora	@GrabID min saya pesan GrabFood, karena lama tidak di pickup saya cancel, kok saldo saya kepotong ya. Dan tidak balik	pesan grabfood tidak pickup batal saldo potong tidak	Negatif

Tabel 2. Contoh Data *Testing*

User (@)	Sebelum <i>Preprocessing</i>	Setelah <i>Preprocessing</i>	Sentimen Manual
aigoosh	Min sekarang kok kode promo Grab motor, mobil, sama GrabFood di aplikasi udah gak pernah ada masuknotifikasi lagi ya? @GrabID	kode promo grab motor mobil grabfood aplikasiudah tidak ada masuk pemberitahuan	?

#### 3.2 Menghitung Term Frequency (TF) pada Data Latih dan Uji

Setelah data training melewati tahap preprocessing, Kemudian kumpulan term disusun ke dalam satu kolom dan dihitung kemunculan kata- nya (TF) pada tiap-tiap dokumen [7], yang dapat dilihat pada tabel dibawah ini.

Tabel 3. Daftar Term Beserta Frekuensi Kemunculannya

Terms	tfi (term frequency)		
	D1	D2	D3
ayo	1	0	0
langgan	1	0	0
pesan	1	0	0
grabfood	1	0	0
nya	1	0	0

mall	1	0	0
penvil	1	0	0
barang	1	0	0
saya	1	0	0
aplikasi	0	1	0
grabfood	0	1	0
merchant	0	1	0
mana	0	1	0
playstore	0	1	0
tidak	0	1	0
ada	0	1	0
pesan	0	0	1
grabfood	0	0	1
Tidak	0	0	1
pickup	0	0	1
batal	0	0	1
saldo	0	0	1
potong	0	0	1
tidak	0	0	1

### 3.3 Hasil Proses Sorting dan Grouping

Dimana sorting adalah untuk mengurutkan term sesuai dengan abjad dan kemudian term yang sama dikelompokkan menjadi satu (grouping) sehingga term- term tersebut dapat dihitung frekuensi kemunculannya pada tiap-tiap dokumen [8]. Proses sorting akan mengurutkan term-term yang ada pada Dokumen 1-3 (D1-D3) sesuai dengan abjad, semua term yang sama harus dimasukkan pada proses sorting, kemudian pada proses grouping term yang sama akan dikumpulkan menjadi satu untuk dihitung term frequency-nya pada tiap dokumen 1-3. Hasil proses sorting dan grouping dapat dilihat pada tabel dibawah ini.

Tabel 4. Hasil Proses *Sorting* dan *Grouping*

Terms	tfi (term frequency)		
	D1	D2	D3
ada	0	1	0
aplikasi	0	1	0
ayo	1	0	0
barang	1	0	0
batal	0	0	1
download	0	1	0
langgan	1	0	0
mall	1	0	0
mana	0	1	0
merchant	0	1	0
nya	1	0	0
penvil	1	0	0
pesan	1	0	1
pickup	0	0	1
playstore	0	1	0
potong	0	0	1
saldo	0	0	1
saya	1	0	0
tidak	0	1	2

### 3.4 Perhitungan Klasifikasi Data *Training* dengan Algoritma *Multinomial Naive Bayes Classifier*

Setelah melalui proses perhitungan kemunculan kata pada tiap dokumen, lalu akan melalui tahap pengklasifikasian.

Pada proses ini akan digunakan algoritma Multinomial Naive Bayes dengan menghitung prior probabilitas. Proses prior probabilitas adalah proses menghitung jumlah dokumen training dalam satu kelas (c) kemudian membaginya dengan jumlah seluruh dokumen training dalam seluruh kelas seperti pada persamaan [9], Kemudian dihitung conditional probabilitas dan laplace smoothing untuk menentukan kelas yang tepat pada dokumen latih. Proses matching pada data uji untuk mengetahui apakah term yang sama ada pada data training atau tidak [10]. Memangkatkan nilai conditional probabilitas pada data testing, bobot nilai term dikuadratkan jika ditemukan term yang sama pada satu dokumen yang sama. Setelah seluruh bobot didapatkan maka akan dilakukan perkalian dengan nilai prior probability yang sebelumnya telah dihitung.

Tabel 5. Hasil Perhitungan Prior Probabilitas untuk Tiap Kelas

Kelas	Prior probabilitas
Positif	85/350
Netral	134/350
Negatif	131/350

- a. Setelah melewati tahapan preprocessing dan menghitung kemunculan kata (tf) pada tiap dokumen seperti yang terdapat pada tabel 3.15, kemudian hitung Prior Probabilitas dengan menggunakan persamaan (3.1)

$$\text{Persamaan: } (c) = \frac{N_c}{N} \quad (3.1)$$

1.  $P(\text{negatif}) = \frac{8}{25}$
2.  $P(\text{positif}) = \frac{9}{25}$
3.  $P(\text{netral}) = \frac{8}{25}$

Keterangan :

- Nilai 8, 9, dan 8 adalah jumlah dokumen training dalam kelas negatif, positif, dan netral.
- Nilai 25 adalah jumlah keseluruhan dokumen training dalam seluruh kelas.

Nilai prior adalah untuk mengetahui seberapa banyak jumlah dokumen dalam satu kelas di data training. Nilai prior ini nantinya yang akan dikalikan dengan nilai probabilitas kemunculan kata yang ada pada data testing.

- b. Hitung Conditional Probabilitas dengan menggunakan persamaan (3.2) :

$$\text{Persamaan (2.3) : } P(tk|c) = \frac{T_{ct}}{\sum_{t \in V} T_{ct}}$$

1.  $P(\text{grabfood}|D1|\text{negatif}) = \frac{1}{8}$
2.  $P(\text{grabfood}|D2|\text{positif}) = \frac{1}{9}$
3.  $P(\text{grabfood}|D3|\text{netral}) = \frac{1}{8}$

Keterangan :

- D1, D2, dan D3 adalah contoh dokumen data 1, 2, dan 3 pada data training yang mewakili beberapa dokumen lainnya pada data latih.
  - Nilai 1 adalah term baik muncul sebanyak 1 kali dalam dokumen 1 pada kelas negatif.
  - Nilai 8, 9, dan 8 adalah jumlah seluruh term yang ada dalam kelas negatif, positif, dan netral.
- c. Laplace smoothing digunakan untuk menghilangkan nilai nol (0) dan merupakan tahap akhir dari proses training menggunakan algoritma Multinomial Naive Bayes. Hasil yang didapat dari proses ini akan menjadi model untuk melakukan klasifikasi [11].

Contoh perhitungan Conditional Probabilitas dan Laplace Smoothing pada dokumen 1 atau D1:

“ayo langgan pesan grabfood nya mall penvil barang saya” menggunakan persamaan (3.3) :

$$\text{Persamaan (2.5) : } P(tk|c) = \frac{T_{ct} + 1}{(\sum_{t \in V} T_{ct}) + B'}$$

1. Hitung bobot nilai setiap kata p12qwersentimen positif pada dokumen 1 (D1):
  - Kata “ayo|D1|positif” =  $\frac{1+1}{9+25} = \frac{2}{34} = 0,058823529$
2. Hitung bobot nilai setiap kata persentimen netral pada dokumen 1 (D1):

- Kata “ayo|D1|netral” =  $\frac{0+1}{8+25} = \frac{1}{33} = 0,03030303$

3. Hitung bobot nilai setiap kata persentimen negatif pada dokumen 1 (D1):

- Kata “ayo|D1|netral” =  $\frac{0+1}{8+25} = \frac{1}{33} = 0,03030303$

Hasil perhitungan dari probabilitas dan laplace smoothing dapat dilihat pada tabel dibawah ini:

Tabel 6. Hasil Perhitungan *Conditional P & Laplace S Per Term*

Terms	Nilai <i>Conditional Probabilitas dan Laplace Smoothing</i>		
	Positif	Netral	Negatif
ada	0,029411765	0,060606061	0,03030303
aplikasi	0,029411765	0,060606061	0,03030303
ayo	0,058823529	0,03030303	0,03030303
barang	0,058823529	0,03030303	0,03030303
batal	0,029411765	0,03030303	0,060606061
download	0,029411765	0,060606061	0,03030303
grabfood	0,058823529	0,060606061	0,060606061
langgan	0,058823529	0,03030303	0,03030303
mall	0,058823529	0,03030303	0,03030303
mana	0,029411765	0,060606061	0,03030303
merchant	0,029411765	0,060606061	0,03030303
nya	0,058823529	0,03030303	0,03030303
penvil	0,058823529	0,03030303	0,03030303
pesan	0,058823529	0,03030303	0,060606061
pickup	0,029411765	0,03030303	0,060606061
playstore	0,029411765	0,060606061	0,03030303
potong	0,029411765	0,03030303	0,060606061
saya	0,058823529	0,03030303	0,03030303
tidak	0,029411765	0,060606061	0,090909091

### 3.5 Perhitungan Klasifikasi Data Testing dengan Algoritma Multinomial Naive Bayes Classifier

a. Sama seperti data training, setelah melewati proses preprocessing data testing akan melalui tahap ascending dan grouping [12]. Hasil sorting dan grouping pada data uji dapat dilihat pada tabel dibawah ini

Tabel 7. Hasil Sorting Ascending dan Grouping Data Testing

Terms	Tfi (term frequency)
	D6
ada	1
aplikasi	1
grab	1
grabfood	1
kode	1
masuk	1
mobil	1
motor	1
pemberitahuan	1
promo	1
tidak	1

- b. Proses matching adalah proses untuk menemukan term yang sama-sama muncul pada model. Tabel dibawah ini merupakan hasil matching menggunakan data utuh. Yaitu hasil perhitungan menyeluruh terhadap data yang diproses. Untuk sebuah term yang kemunculannya lebih dari satu kali, pangkatkan nilai laplace smoothing-nya dengan term frequency berdasarkan kata yang sama. Kemudian kalikan nilainya untuk masing-masing kelas. Contoh, jika term ‘grabfood’ memiliki tf sebanyak 1 kali. Pangkatkan nilai CP&LS-nya untuk menyederhanakan perhitungan.  $P(\text{grabfood}|D1|\text{positif}) = 0,058823529^2 = 0,0034602075640138$

Tabel 8. Hasil Proses Matching Data Testing

Terms	Kelas		
	Positif	Netral	Negatif
ada	0	1	0
aplikasi	0	1	0
grabfood	1	1	1
tidak	0	1	2

- c. Setelah melewati proses matching ditemukan 4 term yang sama yang ditemukan pada data training yang terdapat pada tabel Hasil Perhitungan Conditional Probabilitas dan Laplace Smoothing Data Testing Setiap Kelas, kemudian hitung total nilai conditional probabilitas dan laplace smoothing pada setiap term yang ada di data testing dengan menggunakan persamaan.

1. Hitung bobot nilai setiap kata persentimen negatif pada dokumen 4 (D4):

- Kata “ada|D4|positif” =  $\frac{0+1}{9+25} = \frac{1}{34} = 0,029411765$
- Kata “aplikasi|D4|positif” =  $\frac{0+1}{9+25} = \frac{1}{34} = 0,029411765$
- Kata “grabfood|D4|positif” =  $\frac{1+1}{9+25} = \frac{2}{34} = 0,058823529$
- Kata “tidak|D4|positif” =  $\frac{0+1}{9+25} = \frac{1}{34} = 0,029411765$

2. Hitung bobot nilai setiap kata persentimen positif pada dokumen 4 (D4):

- Kata “ada|D4|netral” =  $\frac{1+1}{8+25} = \frac{2}{33} = 0,060606061$
- Kata “aplikasi|D4|netral” =  $\frac{1+1}{8+25} = \frac{2}{33} = 0,060606061$
- Kata “grabfood|D4|netral” =  $\frac{1+1}{8+25} = \frac{2}{33} = 0,060606061$
- Kata “tidak|D4|netral” =  $\frac{1+1}{8+25} = \frac{2}{33} = 0,060606061$

3. Hitung bobot nilai setiap kata persentimen netral pada dokumen 4 (D4):

- Kata “ada|D4|negatif” =  $\frac{1+1}{8+25} = \frac{1}{33} = 0,03030303$
- Kata “aplikasi|D4|negatif” =  $\frac{1+1}{8+25} = \frac{1}{33} = 0,03030303$
- Kata “grabfood|D4|negatif” =  $\frac{1+1}{8+25} = \frac{2}{33} = 0,060606061$
- Kata “tidak|D4|negatif” =  $\frac{2+1}{8+25} = \frac{3}{33} = 0,090909091$

Hasil perhitungan lengkap conditional probabilitas dan laplace smoothing pada data testing terdapat pada tabel dibawah ini:

Tabel 9. Hasil Perhitungan Conditional Probabilitas dan Laplace Smoothing Data Testing Setiap Kelas

Terms	Tf	Nilai Conditional Probabilitas dan Laplace Smoothing		
		Positif	Netral	Negatif
ada	1	0,029411765	0,060606061	0,03030303
aplikasi	1	0,029411765	0,060606061	0,03030303
grab	1	0,029411765	0,03030303	0,03030303
grabfood	1	0,058823529	0,060606061	0,060606061
kode	1	0,029411765	0,03030303	0,03030303
masuk	1	0,029411765	0,03030303	0,03030303
mobil	1	0,029411765	0,03030303	0,03030303
motor	1	0,029411765	0,03030303	0,03030303
pemberitahuan	1	0,029411765	0,03030303	0,03030303
promo	1	0,029411765	0,03030303	0,03030303
tidak	1	0,029411765	0,060606061	0,090909091
<b>TOTAL</b>		<b>0,382352941</b>	<b>0,484848485</b>	<b>0,454545455</b>

c. Selanjutnya untuk mendapatkan nilai probabilitas dari data testing terhadap seluruh kelas dengan cara mengalikan nilai prior probabilitas dengan total nilai conditional probabilitas dan laplace smoothing untuk masing-masing kelas dengan menggunakan persamaan [13]. Probabilitas masing-masing kelas terhadap data testing untuk D6 yaitu: Persamaan :  $P = N_c \times \frac{T_{cr} + 1}{N (t^f \in V T_{CtF}) + B'}$

1. Probabilitas D4 terhadap kelas positif :  
 $P(\text{positif} | D4) = 85 / 350 * 0,382352941 = 0,092857143$
2. Probabilitas D6 terhadap kelas netral :  
 $P(\text{netral} | D4) = 134 / 350 * 0,484848485 = 0,185627706$
3. Probabilitas D6 terhadap kelas netral :  
 $P(\text{negatif} | D4) = 131 / 350 * 0,454545455 = 0,17012987$

Dari perhitungan probabilitas diatas diketahui bahwa probabilitas D6 terhadap kelas netral memiliki nilai yang paling tinggi, yaitu **0,185627706**. Sehingga D4 masuk ke dalam kelas **netral**.

### 3.6 Hasil Pengujian Sistem

Menghitung akurasi data uji keseluruhan sistem, dengan menggunakan persamaan :

$$AKURASI = \frac{\text{Jumlah data benar}}{\text{jumlah data keseluruha}} \times 100\%$$

$$AKURASI = \frac{303 \times 100\%}{350} = 86,57\%$$

## 4. KESIMPULAN

Dari berbagai penjelasan yang telah diuraikan dalam penelitian ini, maka dapat disimpulkan berbagai hal sebagai berikut: Algoritma Multinomial naive bayes mampu mengetahui sentimen masyarakat terhadap tweet yang di lontarkan pada jasa layanan ojek online Grab Indonesia ke dalam kelas positif, netral, dan negatif. Hasil klasifikasi yang diambil dari tweet sentimen netral menunjukkan tingkat kualitas layanan Grab Indonesia dengan layanan terbaik pada GrabFood sebesar 42,29%, GrabBike sebesar 32,57%, dan GrabCar sebesar 25,14%. Pengujian analisis sentimen pada layanan Grab Indonesia dengan metode Multinomial Naive Bayes Classifier ini menghasilkan akurasi sebesar 86,57%.

## UCAPAN TERIMA KASIH

Alhamdulillah dan terima kasih saya ucapkan kepada bapak Dr. B. Herawan Hayadi, S. Kom., M. Kom sebagai penulis kedua serta dosen pembimbing saya yang telah banyak membimbing dan memberi motivasi kepada saya sehingga terselesainya artikel ini dan terima kasih kepada bapak Dr. Rudi Gunawan, SE., M. Si sebagai penulis ketiga dalam artikel

ini, yang telah memberikan motivasi dan semangat dalam pengerjaan artikel ini, kepada seluruh pihak terkait juga saya ucapkan terimakasih atas kerjasamanya sehingga terbitlah artikel ini.

## DAFTAR PUSTAKA

- [1] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 10, no. 1, pp. 34–40, 2022, doi: 10.23960/jitet.v10i1.2262.
- [2] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.
- [3] R. Amelia, N. S. Prastiwi, and M. E. Purbaya, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Masyarakat Indonesia Mengenai Drama Korea Pada Twitter," vol. 9, no. 2, pp. 338–343, 2022, doi: 10.30865/jurikom.v9i2.3895.
- [4] B. H. Hayadi and A. Basri, "1 , 2 1,2," vol. 2, no. 1, pp. 97–106, 2016.
- [5] A. Rozaq, Y. Yunitasari, K. Sussolaikah, E. Resty, N. Sari, and R. I. Syahputra, "Analisis Sentimen Terhadap Implementasi Program Merdeka Belajar Kampus Merdeka Menggunakan Naive Bayes , K-Nearest Neighbors Dan Decision Tree," vol. 6, no. April, pp. 746–750, 2022, doi: 10.30865/mib.v6i2.3554.
- [6] W. A. Prabowo and C. Wiguna, "Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM," *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 149, 2021, doi: 10.30865/mib.v5i1.2604.
- [7] P. Morgan, B. Show, and N. Media, "Analisis Sentimen Pengguna Youtube Terhadap Tayangan # Matanajwamenantiterawan Dengan Metode Naive Bayes Sentiment Analysis Of User Comments On Youtube Video # Matanajwamenantiterawan With Naive Bayes Classifier Method," vol. 5, no. 1, pp. 1–6, 2022, doi: 10.33387/jiko.
- [8] R. Puspita and A. Widodo, "Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS," *J. Inform. Univ. Pamulang*, vol. 5, no. 4, p. 646, 2021, doi: 10.32493/informatika.v5i4.7622.
- [9] F. Harahap, "Perbandingan Algoritma K Means dan K Medoids Untuk Clustering Kelas Siswa Tunagrahita," *TIN Terap. Inform. Nusant.*, vol. 2, no. 4, pp. 191–197, 2021.
- [10] W. Katrina, H. J. Damanik, F. Parhusip, D. Hartama, A. P. Windarto, and A. Wanto, "C.45 Classification Rules Model for Determining Students Level of Understanding of the Subject," *J. Phys. Conf. Ser.*, vol. 1255, no. 1, 2019, doi: 10.1088/1742-6596/1255/1/012005.
- [11] E. P. Cynthia and E. Ismanto, "Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.)*, vol. 3, no. July, p. 1, 2018, doi: 10.30645/jurasik.v3i0.60.
- [12] H. Al, R. Harpizon, R. Kurniawan, and I. Iskandar, "Analisis Sentimen Komentar Di YouTube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naive Bayes," vol. 5, no. 1, pp. 131–140, 2022.
- [13] A. Ikhwan *et al.*, "A novelty of data mining for promoting education based on FP-growth algorithm," *Int. J. Civ. Eng. Technol.*, vol. 9, no. 7, pp. 1660–1669, 2018.