

# **Indonesian Journal of Innovation Studies**

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

## **Table Of Content**

<b>Journal Cover</b> .....	2
<b>Author[s] Statement</b> .....	3
<b>Editorial Team</b> .....	4
<b>Article information</b> .....	5
Check this article update (crossmark) .....	5
Check this article impact .....	5
Cite this article .....	5
<b>Title page</b> .....	6
Article Title .....	6
Author information .....	6
Abstract .....	6
<b>Article content</b> .....	7

**ISSN (ONLINE) 2598-9936**



**INDONESIAN JOURNAL OF INNOVATION STUDIES**

PUBLISHED BY  
UNIVERSITAS MUHAMMADIYAH SIDOARJO

# **Indonesian Journal of Innovation Studies**

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

## **Originality Statement**

The author[s] declare that this article is their own work and to the best of their knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the published of any other published materials, except where due acknowledgement is made in the article. Any contribution made to the research by others, with whom author[s] have work, is explicitly acknowledged in the article.

## **Conflict of Interest Statement**

The author[s] declare that this article was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## **Copyright Statement**

Copyright © Author(s). This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

# **Indonesian Journal of Innovation Studies**

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

## **EDITORIAL TEAM**

### **Editor in Chief**

Dr. Hindarto, Universitas Muhammadiyah Sidoarjo, Indonesia

### **Managing Editor**

Mochammad Tanzil Multazam, Universitas Muhammadiyah Sidoarjo, Indonesia

### **Editors**

Fika Megawati, Universitas Muhammadiyah Sidoarjo, Indonesia

Mahardika Darmawan Kusuma Wardana, Universitas Muhammadiyah Sidoarjo, Indonesia

Wiwit Wahyu Wijayanti, Universitas Muhammadiyah Sidoarjo, Indonesia

Farkhod Abdurakhmonov, Silk Road International Tourism University, Uzbekistan

Bobur Sobirov, Samarkand Institute of Economics and Service, Uzbekistan

Evi Rinata, Universitas Muhammadiyah Sidoarjo, Indonesia

M Faisal Amir, Universitas Muhammadiyah Sidoarjo, Indonesia

Dr. Hana Catur Wahyuni, Universitas Muhammadiyah Sidoarjo, Indonesia

Complete list of editorial team ([link](#))

Complete list of indexing services for this journal ([link](#))

How to submit to this journal ([link](#))

## Article information

**Check this article update (crossmark)**



**Check this article impact (\*)**



**Save this article to Mendeley**



(\*) Time for indexing process is various, depends on indexing database platform

## **Comment Sentiment Analysis of JNE Using K-Nearest Neighbor (KNN) Method on Twitter**

*Analisis Sentimen Komentar terhadap JNE Menggunakan Metode K-Nearest Neighbor (KNN) pada Twitter*

**Ricky Renaldo Arisandi, cassaduabelas@gmail.com, (1)**

*Universitas Muhammadiyah Sidoarjo, Indonesia*

**Sumarno Sumarno, sumarno@umsida.ac.id, (0)**

*Universitas Muhammadiyah Sidoarjo, Indonesia*

**Hamzah Setiawan, hamzah@umsida.ac.id, (0)**

*Universitas Muhammadiyah Sidoarjo, Indonesia*

<sup>(1)</sup> Corresponding author

### **Abstract**

Social media has evolved into a prominent public space for virtual criticism, particularly on platforms like Twitter, facilitated by widespread smartphone usage. Netizens utilize Twitter as an effective communication channel due to its accessibility and vast reach. This study focuses on sentiment analysis of comments from the public on Twitter, aiming to expedite the acquisition of accurate information about the general sentiment towards JNE (a logistics company). The K-Nearest Neighbor (KNN) classifier is employed, employing the TF-IDF weighting method to classify Indonesian language comments and assess the achieved accuracy.

### **Highlights:**

- Study focused on sentiment analysis of Twitter comments concerning JNE services using the K-Nearest Neighbor (KNN) method with Indonesian language text.
- Employed the TF-IDF weighting to classify comments and achieved an impressive 90% accuracy in sentiment analysis.
- The obtained classification proves valuable in evaluating public perception of JNE's services based on feedback from the social media community on Twitter.

---

Published date: 2023-07-25 00:00:00

---

## Pendahuluan

PT. Tiki Jalur Nugraha Eka Kurir atau biasa disebut JNE Express merupakan salah satu perusahaan penyedia jasa pengiriman barang terbesar dengan jaringan terluas di Indonesia. Lingkup bisnis yang di jalani oleh PT. JNE EXPRESS adalah pengiriman (*courier*), penyimpanan dan pendistribusian (*logistic and distribution*), transportasi (*moving, trucking, air sea cargo*), (*custom clearance*), penjemputan bandara (*escort service*), pengiriman uang (*money transfer*). Semakin banyaknya peminat pengguna layanan tersebut semakin banyak pula opini yang didapatkan dari masyarakat, baik itu tentang fasilitas ataupun lainnya.

Media sosial telah menjadi ruang publik tanpa batas ruang dan waktu yang ditandai dengan banyaknya penggunaan *smartphone* terutama Twitter. Netizen menggunakan sebagai sarana komunikasi yang efektif karena sangat mudah dalam penggunaannya. Sekali klik, jutaan netizen yang terhubung dapat melihat dengan mudah tanpa batas . Hasil opini masyarakat tersebut dapat digunakan sebagai data untuk analisa sentimen yang berupa komentar dan dijadikan bahan untuk analisa teks mining. Dengan data teks komentar tersebut terdapat berbagai macam opini yang diberikan oleh masyarakat melalui media sosial contohnya seperti di media sosial twitter. PT Tiki Jalur Nugraha Eka Kurir (JNE) memproyeksikan pertumbuhan pengiriman paket hingga akhir tahun 2022 dapat konsisten sekitar 30%-40%. Optimisme ini didukung oleh perkembangan *e-commerce* yang begitu signifikan, sehingga dapat turut meningkatkan permintaan pengiriman paket ke depan .

JNE mempunyai akun resmi di twitter yaitu bernama JNE Express, dalam akun tersebut terdapat berbagai *tweet* dan komentar masyarakat yang telah diberikan, sehingga perlu adanya teknologi yang tepat untuk melakukan analisa sentimen. Metode yang dapat digunakan untuk analisa sentimen adalah metode KNN. Keuntungan menggunakan metode ini yaitu tangguh terhadap training data yang noisy dan efektif apabila datanya besar serta metode yang mudah dipahami dan mudah diimplementasikan. Penelitian ini bertujuan untuk mempermudah dalam memperoleh informasi yang lebih akurat tentang *mood* media massa, dan hasil informasi tersebut lebih akurat jika layanan yang diberikan negatif, positif dan netral.

Berdasarkan latar belakang diatas, penulis mengusulkan penelitian yang akan dibahas yaitu tentang Analis Sentimen Komentar Terhadap JNE Menggunakan Metode K-Nearest Neighbor (KNN) pada Media Sosial Twitter menggunakan 100 data.

PT. Tiki Jalur Nugraha Eka Kurir didirikan oleh H. Soeprapto Suparno pada tanggal 26 November 1990. Perusahaan ini merupakan divisi dari PT Citra van Titipan Kilat (TIKI) yang bertugas untuk mengelola jaringan kurir internasional. Berawal dari delapan orang dan modal 100 juta rupiah, JNE memulai usahanya yang terpusat pada penanganan kegiatan kepabeanan, impor kiriman barang, dokumen serta pengantarannya dari luar negeri ke Indonesia .

Twitter merupakan sebuah layanan jejaring sosial dan bisa disebut juga *microblog daring* yang memungkinkan penggunanya untuk berkirim dan membaca pesan berbasis teks yang tidak lebih dari 280 karakter dalam sebuah tweet .

Analisa sentimen bagian dari opinion mining merupakan proses memahami, mengestrak dan mengelolah data teksual yang secara otomatis untuk memperoleh sebuah informasi. Analisa sentimen sering disebut juga opinion mining (penambangan opini) adalah koputasi untuk mengenali dan mengekspresikan opini, sentimen evaluasi, sikap, emosi, subjektifitas, penilaian serta pandangan dalam suatu teks [6].

Database atau basis data adalah kumpulan dari beberapa informasi yang disimpan di dalam komputer secara sistematis sehingga dapat diperiksa menggunakan suatu program komputer untuk memperoleh informasi dari basis data tersebut [7].

Flowchart merupakan penggambaran secara grafik dari langkah-langkah dan urutan prosedur suatu program. Biasanya mempengaruhi penyelesaian masalah yang khususnya perlu dipelajari dan dievaluasi lebih lanjut [8].

Python merupakan salah satu bahasa pemrograman yang banyak digunakan oleh perusahaan besar maupun para developer untuk mengembangkan berbagai macam aplikasi berbasis desktop, web dan mobile. Python diciptakan oleh Guido van Rossum di Belanda pada tahun 1990 dan namanya diambil dari acara televisi kesukaan Guido Monty Python's Flying Circus. Van Rossum mengembangkan Python sebagai hobi, kemudian Python menjadi bahasa pemrograman yang dipakai secara luas dalam industri dan pendidikan karena sederhana, ringkas, sintak sintuitif dan memiliki pustaka yang luas [9].

## JNE

PT. Tiki Jalur Nugraha Eka Kurir didirikan oleh H. Soeprapto Suparno pada tanggal 26 November 1990. Perusahaan ini merupakan divisi dari PT Citra van Titipan Kilat (TIKI) yang bertugas untuk mengelola jaringan kurir internasional. Berawal dari delapan orang dan modal 100 juta rupiah, JNE memulai usahanya yang terpusat pada penanganan kegiatan kepabeanan, impor kiriman barang, dokumen serta pengantarannya dari luar negeri ke

Indonesia [4].

## **Twitter**

Twitter merupakan sebuah layanan jejaring sosial dan bisa disebut juga *microblog daring* yang memungkinkan penggunanya untuk berkirim dan membaca pesan berbasis teks yang tidak lebih dari 280 karakter dalam sebuah tweet [5].

## **Analisa Sentimen**

Analisa sentimen bagian dari opinion mining merupakan proses memahami, mengestrak dan mengelolah data tekstual yang secara otomatis untuk memperoleh sebuah informasi. Analisa sentimen sering disebut juga opinion mining (penambangan opini) adalah koputasi untuk mengenali dan mengekspresikan opini, sentimen evaluasi, sikap, emosi, subjektifitas, penilaian serta pandangan dalam suatu teks [6]

## **Database**

Database atau basis data adalah kumpulan dari beberapa informasi yang disimpan di dalam komputer secara sistematis sehingga dapat diperiksa menggunakan suatu program komputer untuk memperoleh informasi dari basis data tersebut [7]

## **Flowchart**

Flowchart merupakan penggambaran secara grafik dari langkah-langkah dan urutan prosedur suatu program. Biasanya mempengaruhi penyelesaian masalah yang khususnya perlu dipelajari dan dievaluasi lebih lanjut [8].

## **Python**

Python merupakan salah satu bahasa pemrograman yang banyak digunakan oleh perusahaan besar maupun para developer untuk mengembangkan berbagai macam aplikasi berbasis desktop, web dan mobile. Python diciptakan oleh Guido van Rossum di Belanda pada tahun 1990 dan namanya diambil dari acara televisi kesukaan Guido Monty Python's Flying Circus. Van Rossum mengembangkan Python sebagai hobi, kemudian Python menjadi bahasa pemrograman yang dipakai secara luas dalam industri dan pendidikan karena sederhana, ringkas, sintak sintuitif dan memiliki pustaka yang luas [9].

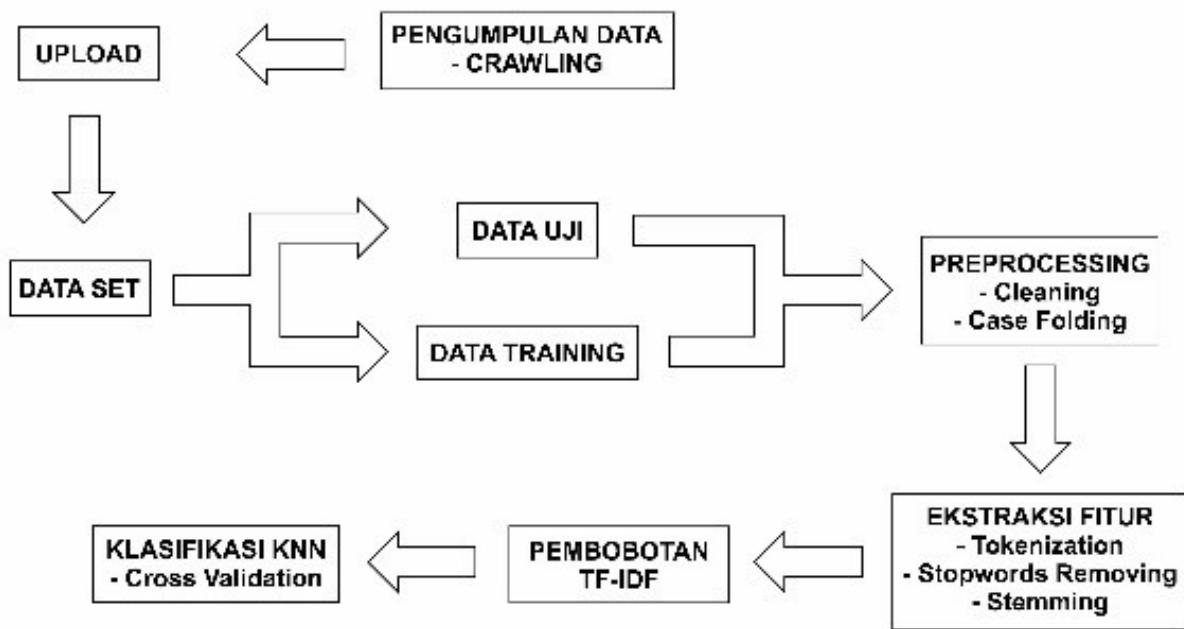
## **Text Mining**

Text mining adalah sebuah proses yang digunakan untuk mengeksplorasi data berupa teks, dimana sumber data tersebut biasanya didapatkan dari dokumen [10]. Text Mining bertujuan untuk menemukan informasi berharga yang tersembunyi baik dari sumber informasi terstruktur dan tidak terstruktur. Text Mining merupakan suatu penemuan baru yang sebelumnya informasinya tidak diketahui. Informasi yang diekstrak dari berbagai sumber daya tertulis dilakukan secara otomatis. Elemen kuncinya adalah menghubungkan beberapa informasi yang diekstraksi menjadi satu sehingga dapat membentuk fakta baru atau hipotesis baru untuk dieksplorasi lebih lanjut [11]. Tugas text mining antara lain, pengkategorian teks (*text categorization*) dan pengelompokan teks (*text clustering*).

## **Metodologi Penelitian**

Pengklasifikasi K-Nearest Neighbor (KNN) dapat diklasifikasikan dengan mempelajari "kumpulan dari model contoh dokumen yang sebelumnya diklasifikasikan". Dengan pembobotan TF-IDF dalam mengklasifikasikan komentar pada sebuah teks bahasa Indonesia dan mengetahui seberapa besar akurasi yang didapatkan.

Langkah-langkah proses analisis metode sentiment:



**Figure 1.** Proses Analisis Sentimen

Data berupa teks sebanyak 100 komentar yang diambil dari akun resmi PT. Tiki Jalur Nugraha Eka Kurir (JNE) yang bersumber dari media sosial twitter.

Dalam table dibawah ini merupakan 10 contoh dokumen yang telah didapatkan peneliti dari twitter. Data ini telah peneliti klasifikasi secara manual sebagai contoh perhitungan manual:

No	Text	Klasifikasi
1.	ngirim pake jne trucking lama banget dah	Negatif
2.	kami melayani kiriman anda sampai ke penjuru nusantara dan pelosok desa... percayakan pada JNE Juanda Depok saja yukk.	Positif
3.	@JNECare Iya parah tuh, punya saya juga dr jakarta kr bandung, dr hari sdh 8 hari gk nyampe	Negatif
4.	@JNECare Iya respon yanh menyatakan 3-6 hari? Padahal paket saya udah dari tanggal 5? Ini tanggalan JNE yang beda	Negatif
5.	Manfaatkan layanan jemput gratis dari JNE Kalimantan Blitar, Mudah dan Praktis.#JNEBLITAR	Positif
6.	Packing dirusak kek gini maksudnya apa ya? Mana barangnya diambil sendiri di gudang, dah 5hari ga diantar @JNE_ID	Negatif
7.	@Audioford1 selamat siang kak, maaf atas ketidaknyamanannya. terkait kendala yang terjadi dibantu tim @JNECare ya kak	Positif
8.	JNE lama dan tidak membantu, paket sudah 14 hari tidak bergerak tapi setiap hubungi CS selalu diminta	Negatif

	tunggu	
9.	Bila tim pick up JNE Juanda Depok kurang berkenan atau kurang sopan dalam pelayanannya bisa mention kami utk ditindak lanjuti,terima kasih	Positif
10.	@JNECare tolong liat pesan saya ini email sy gk drespon,paket saya salah alamt krn jne salah input alamat ini gmnaa!!	Negatif

**Table 1.** Contoh Klasifikasi

## Preprocessing

Tahap ini merupakan tahap membersihkan kata sebelum diolah lebih lanjut pada tahap berikutnya. Tahap preprocessing sangat penting, karena pada tahap ini semua perubahan kata akan memiliki nilai sendiri yang mempengaruhi akurasi sistem yang akan dibuat nantinya. Berikut tahapan Preprocessing:

### Case Folding

Tahapan Case Folding merupakan sebuah proses yang digunakan untuk pengubahan teks yang berhuruf kapital menjadi huruf kecil dan hanya digunakan untuk huruf 'a' sampai dengan 'z' yang diterima. Selain karakter huruf dihilangkan dan dianggap pembatas.

No	Input	Output
1.	ngirim pake jne trucking lama banget dah	ngirim pake jne trucking lama banget dah
2.	kami melayani kiriman anda sampai ke penjuru nusantara dan pelosok desa... percayakan pada JNE Juanda Depok saja yukk.	kami melayani kiriman anda sampai ke penjuru nusantara dan pelosok desa... percayakan pada jne juanda depok saja yukk.
3.	@JNECare Iya parah tuh, punya saya juga dr jakarta kr bandung, dr hari sdh 8 hari gk nyampe	@jnecare iya parah tuh, punya saya juga dr jakarta kr bandung, dr hari sdh 8 hari gk nyampe
4.	@JNECare Iya respon yanh menyatakan 3-6 hari? Padahal paket saya udah dari tanggal 5? Ini tanggalan JNE yang beda	@jnecare iya respon yanh menyatakan 3-6 hari? padahal paket saya udah dari tanggal 5? ini tanggalan jne yang beda
5.	Manfaatkan layanan jemput gratis dari JNE Kalimantan Blitar, Mudah dan Praktis.#JNEBLITAR	manfaatkan layanan jemput gratis dari jne kalimantan blitar, mudah dan praktis.#jneblitar
6.	Packing dirusak kek gini maksudnya apa ya? Mana barangnya diambil sendiri di gudang, dah 5hari ga diantar @JNE_ID	packing dirusak kek gini maksudnya apa ya? mana barangnya diambil sendiri di gudang, dah 5hari ga diantar @jne_id
7.	@Audioford1 selamat siang kak, maaf atas ketidaknyamanannya. terkait kendala yang terjadi dibantu tim @JNECare ya kak	@audioford1 selamat siang kak, maaf atas ketidaknyamanannya. terkait kendala yang terjadi dibantu tim @jnecare ya kak
8.	JNE lama dan tidak membantu, paket sudah 14 hari tidak bergerak tapi setiap hubungi CS selalu diminta tunggu	jne lama dan tidak membantu, paket sudah 14 hari tidak bergerak tapi setiap hubungi cs selalu diminta tunggu
9.	Bila tim pick up JNE Juanda Depok kurang berkenan atau kurang sopan dalam pelayanannya bisa mention kami utk ditindak lanjuti,terima kasih	bila tim pick up jne juanda depok kurang berkenan atau kurang sopan dalam pelayanannya bisa mention kami utk ditindak lanjuti,terima kasih
10.	@JNECare tolong liat pesan saya ini email sy gk drespon,paket saya salah alamt krn jne salah input alamat ini gmnaa!!	@jnecare tolong liat pesan saya ini email sy gk drespon,paket saya salah alamt krn jne salah input alamat ini gmnaa!!

**Table 2.** Proses Case Folding

# Indonesian Journal of Innovation Studies

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

## Tokenizing

Tahapan tokenizing merupakan tahapan proses untuk menguraikan deskripsi yang semulanya berupa kalimat-kalimat menjadi kata-kata dan menghilangkan delimiter seperti tanda titik (.), koma (,) dan spasi pada karakter kata pada sebuah kalimat tersebut:

No	Input	Output
1.	ngirim pake jne trucking lama banget dah	ngirim pake jne trucking lama banget dah
2.	kami melayani kiriman anda sampai ke penjuru nusantara dan pelosok desa... percayakan pada jne juanda depok saja yukk.	kami melayani kiriman anda sampai ke penjuru nusantara dan pelosok desa percayakan pada jne juanda depok saja yukk
3.	@jnecare iya parah tuh punya saya juga dr jakarta kr bandung, dr hari sdh 8 hari gk nyampe	jnecare iya parah tuh punya saya juga dr jakarta kr bandung dr hari sdh 8 hari gk nyampe
4.	@jnecare iya respon yanh menyatakan 3-6 hari? padahal paket saya udah dari tanggal 5? ini tanggalan jne yang beda	jnecare iya respon yanh menyatakan 36 hari padahal paket saya udah dari tanggal 5 ini tanggalan jne yang beda
5.	manfaatkan layanan jemput gratis dari jne kalimantan blitar, mudah dan praktis.#jneblitar	manfaatkan layanan jemput gratis dari jne kalimantan blitar mudah dan praktis jneblitar
6.	packing dirusak kek gini maksudnya apa ya? mana barangnya diambil sendiri di gudang, dah 5hari ga diantar @jne_id	packing dirusak kek gini maksudnya apa ya mana barangnya diambil sendiri di gudang dah 5hari ga diantar jneid
7.	@audioford1 selamat siang kak, maaf atas ketidaknyamanannya. terkait kendala yang terjadi dibantu tim @jnecare ya kak	audioford1 selamat siang kak maaf atas ketidaknyamanannya terkait kendala yang terjadi dibantu tim jnecare ya kak
8.	jne lama dan tidak membantu, paket sudah 14 hari tidak bergerak tapi setiap hubungi cs selalu diminta tunggu	jne lama dan tidak membantu paket sudah 14 hari tidak bergerak tapi setiap hubungi cs selalu diminta tunggu
9.	bila tim pick up jne juanda depok kurang berkenan atau kurang sopan dalam pelayanannya bisa mention kami utk ditindak lanjuti,terima kasih	bila tim pick up jne juanda depok kurang berkenan atau kurang sopan dalam pelayanannya bisa mention kami utk ditindak lanjuti terima kasih
10.	@jnecare tolong liat pesan saya ini email sy gk drespon,paket saya salah alamt krn jne salah input alamat ini gmnaa!!	jnecare tolong liat pesan saya ini email sy gk dresponpaket saya salah alamt krn jne salah input alamat ini gmnaa

Table 3. Proses Tokenizing

## Filtering / Stopword Removal

Stopword Removal adalah proses menghilangkan kata yang tidak penting, guna merampingkan perhitungan terhadap suatu dokumen. Penghilangan kata pada tahap ini misalnya pada kata 'di', 'ada', 'ketika' dll. Karena dokumen pada twitter merupakan data yang sangat tidak terstruktur, sehingga untuk memproses bagian ini harus ditambahkan list kata tambahan. Berikut merupakan contoh table luaran proses stopword removal:

No	Input	Output
1.	ngirim pake jne trucking lama banget dah	ngirim pake jne trucking lama banget
2.	kami melayani kiriman anda sampai ke penjuru nusantara dan pelosok desa percayakan pada jne juanda depok saja yukk	kami melayani kiriman anda sampai penjuru nusantara pelosok desa percayakan jne juanda depok saja yukk
3.	jnecare iya parah tuh punya saya juga dr jakarta kr bandung, dr hari sdh 8 hari gk nyampe	jnecare parah tuh punya saya juga dr jakarta kr bandung dr hari sdh 8 hari gk nyampe
4.	jnecare iya respon yanh menyatakan 36	jnecare respon yanh menyatakan 36

# Indonesian Journal of Innovation Studies

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

	36 hari padahal paket saya udah dari tanggal 5 ini tanggalan jne yang beda	hari padahal paket saya udah tanggal 5 tanggalan jne beda
5.	manfaatkan layanan jemput gratis dari jne kalimantan blitar mudah dan praktis jneblitar	manfaatkan layanan jemput gratis jne kalimantan blitar mudah praktis jneblitar
6.	packing dirusak kek gini maksudnya apa ya mana barangnya diambil sendiri di gudang dah 5hari ga diantar jneid	packing dirusak kek gini maksudnya mana barangnya diambil sendiri gudang dah 5hari ga diantar jneid
7.	audioford1 selamat siang kak maaf atas ketidaknyamanannya terkait kendala yang terjadi dibantu tim jnecare ya kak	audioford1 selamat siang kak maaf atas ketidaknyamanannya terkait kendala terjadi dibantu tim jnecare kak
8.	jne lama dan tidak membantu paket sudah 14 hari tidak bergerak tapi setiap hubungi cs selalu diminta tunggu	jne lama membantu paket sudah 14 hari bergerak tapi setiap hubungi cs selalu diminta tunggu
9.	bila tim pick up jne juanda depok kurang berkenan atau kurang sopan dalam pelayanannya bisa mention kami utk ditindak lanjuti,terima kasih	bila tim pick up jne juanda depok kurang berkenan kurang sopan dalam pelayanannya bisa mention kami utk ditindak lanjuti terima kasih
10.	jnecare tolong liat pesan saya ini email sy gk dresponpaket saya salah alamt krn jne salah input alamat ini gmnaa	jnecare tolong liat pesan saya email sy gk dresponpaket saya salah alamt krn jne salah input alamat gmnaa

**Table 4.** Proses Filtering/ Stopword Removal

## Stemming

Proses Stemming merupakan mengubah kata yang berimbahan menjadi kata dasar, pada tahap ini peneliti menggunakan library dari sastrawi untuk melakukan perubahan pada kata berikut:

No	Input	Output
1.	ngirim pake jne trucking lama banget	kirim pake jne trucking lama banget
2.	kami melayani kiriman anda sampai penjuru nusantara pelosok desa percayakan jne juanda depok saja yukk	kami layan kirim anda sampai penjuru nusantara pelosok desa percaya jne juanda depok saja yukk
3.	jnecare iya parah tuh punya saya juga dr jakarta kr bandung dr hari sdh 8 hari gk nyampe	jnecare parah tuh punya saya juga dr jakarta kr bandung dr hari sdh 8 hari gk nyampe
4.	jnecare iya respon yanh menyatakan 36 hari padahal paket saya udah tanggal 5 tanggalan jne beda	jnecare respon yanh nyata 36 hari padahal paket saya udah tanggal 5 tanggalan jne beda
5.	manfaatkan layanan jemput gratis jne kalimantan blitar mudah praktis jneblitar	manfaat layan jemput gratis jne kalimantan blitar mudah praktis jneblitar
6.	packing dirusak kek gini maksudnya mana barangnya diambil sendiri gudang dah 5hari ga diantar jneid	packing rusak kek gini maksud mana barang ambil sendiri gudang dah 5hari ga antar jneid
7.	audioford1 selamat siang kak maaf atas ketidaknyamanannya terkait kendala terjadi dibantu tim jnecare kak	audioford1 selamat siang kak maaf atas tidak nyaman kait kendala jadi bantu tim jnecare kak
8.	jne lama membantu paket sudah 14 hari bergerak tapi setiap hubungi cs selalu diminta tunggu	jne lama bantu paket sudah 14 hari gerak tapi tiap hubung cs selalu minta tunggu
9.	bila tim pick up jne juanda depok kurang berkenan kurang sopan dalam pelayanannya bisa mention kami utk ditindak lanjuti terima kasih	bila tim pick up jne juanda depok kurang kenan kurang sopan dalam layan bisa mention kami utk tindak lanjut terima kasih
10.	jnecare tolong liat pesan saya email	jnecare tolong liat pesan saya email

# Indonesian Journal of Innovation Studies

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

	sy gk dresponpaket saya salah alamt krn jne salah input alamat gmnaa	sy gk dresponpaket saya salah alamt krn jne salah input alamat gmnaa
--	---	---

**Table 5.** Proses Stemming

## Pembobotan TF/IDF

Pembobotan kata yang peneliti gunakan disini adalah pembobotan TF-IDF. Menentukan kata kunci terlebih dahulu sebelum melakukan perhitungan. Berikut ini merupakan contoh penerapan TF pada data testing di dokumen.

D9	bila tim pick up jne juanda depok kurang kenan kurang sopan dalam layan bisa mention kami utk tindak lanjut terima kasih
----	--

**Table 6.** Data Testing

Berikut ini merupakan bentuk pencarian TF pada data testing dengan mencari kata yang muncul dalam dokumen.

bila	1
tim	1
pick	1
up	1
jne	1
juanda	1
depok	1
kurang	1
sopan	1
dalam	1
layan	1
bisa	1
mention	1
kami	1
utk	1
tindak	1
lanjut	1
terima	1
kasih	1

**Table 7.** Pencarian TF Pada Data Testing

Berikut ini adalah pencarian nilai TF dan DF pada data training,. nilai TF merupakan kata yang sering muncul pada dokumen sedangkan DF merupakan jumlah kata yang muncul dari dokumen. Di bawah ini contoh pencarian TF dan DF pada data training.

# Indonesian Journal of Innovation Studies

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

NO	Term	TF										DF
		D9	D1	D2	D3	D4	D5	D6	D7	D8	D10	
1	bila	1	0	0	0	0	0	0	0	0	0	1
2	tim	1	0	0	0	0	0	0	0	0	0	1
3	pick	1	0	0	0	0	0	0	0	0	0	1
4	up	1	0	0	0	0	0	0	0	0	0	1
5	jne	1	1	1	0	0	1	0	0	1	1	6
6	juanda	1	0	1	0	0	0	0	0	0	0	2
7	depok	1	0	1	0	0	0	0	0	0	0	1
8	kurang	1	0	0	0	0	0	0	0	0	0	1
9	sopan	1	0	0	0	0	0	0	0	0	0	1
10	dalam	1	0	0	0	0	0	0	0	0	0	1
11	layan	1	0	1	0	0	1	0	0	0	0	3
12	bisa	1	0	0	0	0	0	0	0	0	0	1
13	mention	1	0	0	0	0	0	0	0	0	0	1
14	kami	1	0	1	0	0	0	0	0	0	0	2
15	utk	1	0	0	0	0	0	0	0	0	0	1
16	tindak	1	0	0	0	0	0	0	0	0	0	1
17	lanjut	1	0	0	0	0	0	0	0	0	0	1
18	terima	1	0	0	0	0	0	0	0	0	0	1
19	kasih	1	0	0	0	0	0	0	0	0	0	1

**Figure 2.** Pencarian TF dan DF

Terdapat sebagai contoh kata “dapat” muncul 2 kali pada D2 berarti TF D2 = 2, sedangkan nilai DF = 1 di karenakan DF hanya akan menghitung jumlah kata yang mewakili setiap dokumen, buatkan menghitung jumlah kata yang muncul dalam dokumen.

## Menghitung Nilai IDF

Proses perhitungan idf menggunakan rumus (2.1). Dimana D merupakan jumlah dokumen dan DF merupakan dokumen frekuensi atau jumlah kata yang akan muncul pada setiap dokumen. Untuk mencari IDF bisa menggunakan rumus  $\text{Log}(D/DF)$ . Contoh pada kata “terima” yang memiliki adalah D9 dan DF = 1. Maka bisa dihitung seperti berikut :

$$\begin{aligned} \text{Misalkan IDF bayar} &= \text{Log} \left( \frac{D}{DF} \right) \\ &= \text{Log} \left( \frac{10}{1} \right) \\ &= \text{Log } 10 = 1.000 \end{aligned}$$

**Figure 3.** Rumus IDF

# Indonesian Journal of Innovation Studies

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

NO	Term	TF										DF	IDF
		D9	D1	D2	D3	D4	D5	D6	D7	D8	D10		
1	bila	1	0	0	0	0	0	0	0	0	0	1	1,000
2	tim	1	0	0	0	0	0	0	0	0	0	1	1,000
3	pick	1	0	0	0	0	0	0	0	0	0	1	1,000
4	up	1	0	0	0	0	0	0	0	0	0	1	1,000
5	jne	1	1	1	0	0	1	0	0	1	1	6	0,222
6	juanda	1	0	1	0	0	0	0	0	0	0	2	0,699
7	depok	1	0	1	0	0	0	0	0	0	0	1	1,000
8	kurang	1	0	0	0	0	0	0	0	0	0	1	1,000
9	sopan	1	0	0	0	0	0	0	0	0	0	1	1,000
10	dalam	1	0	0	0	0	0	0	0	0	0	1	1,000
11	layan	1	0	1	0	0	1	0	0	0	0	3	0,523
12	bisa	1	0	0	0	0	0	0	0	0	0	1	1,000
13	mention	1	0	0	0	0	0	0	0	0	0	1	1,000
14	kami	1	0	1	0	0	0	0	0	0	0	2	0,699
15	utk	1	0	0	0	0	0	0	0	0	0	1	1,000
16	tindak	1	0	0	0	0	0	0	0	0	0	1	1,000
17	lanjut	1	0	0	0	0	0	0	0	0	0	1	1,000
18	terima	1	0	0	0	0	0	0	0	0	0	1	1,000
19	kasih	1	0	0	0	0	0	0	0	0	0	1	1,000

Figure 4. Pencarian TF dan DF

wdt=tf.idf										
D9	D1	D2	D3	D4	D5	D6	D7	D8	D10	
1,000	0	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
0,222	0,22184875	0,22184875	0	0	0,22184875	0	0	0,22184875	0,22184875	
0,699	0	0,698970004	0	0	0	0	0	0	0	
1,000	0	1,000	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
0,523	0	0,523	0	0	0,522878745	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
0,699	0	0,699	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0	0

Figure 5. Pencarian TF IDF Proses ini pencarian nilai tf idf dengan mengalikan nilai tf dengan nilai idf.

## Tahap Klasifikasi

Berdasarkan tabel di atas untuk menentukan nilai tf idf dengan cara mengalikan nilai tf dengan nilai idf. Berikut ini sebagai contoh

Tahap ini merupakan tahap lanjutan dari tahap pembobotan yaitu tahapan nilai similaritas atau nilai kemiripan antara data testing dengan data training menggunakan algoritma KNN. Algoritma ini harus menentukan nilai k dan mengambil nilai kemiripan yang paling tinggi. Kemudian menyesuaikan dengan hasil kemiripan yang sesuai dengan nilai k. Berikut ini contoh perkalian data testing dengan data training

# Indonesian Journal of Innovation Studies

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

WD9*Wdi									
D1	D2	D3	D4	D5	D6	D7	D8	D10	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0,049217	0,049217	0	0	0,049217	0	0	0,049217	0,049217	
0	0,488559	0	0	0	0	0	0	0	
0	1,000	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0,273	0	0	0,273402	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0,489	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	
0,0492	2,2997	0,0000	0,0000	0,3226	0,0000	0,0000	0,0492	0,0492	

**Figure 6.** Perkalian jumlah data testing dengan data training

Tahap selanjutnya yaitu mencari penjang vektor dengan cara mengkalikan akar jumlah data testing sudah dikudratkan dan akar jumlah data training yang sudah dikuadratkan.

Panjang vector									
D9	D1	D2	D3	D4	D5	D6	D7	D8	D10
1,000	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
0,049	0,049217	0,049217	0	0	0,049217	0	0	0,049217	0,049217
0,489	0	0,488559	0	0	0	0	0	0	0
1,000	0	1,000	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
0,273	0	0,273	0	0	0,273402	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
0,489	0	0,489	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
1,000	0	0	0	0	0	0	0	0	0
16,2997	0,0492	2,2997	0,0000	0,0000	0,3226	0,0000	0,0000	0,0492	0,0492
4,0373	0,2218	1,5165	0,0000	0,0000	0,5680	0,0000	0,0000	0,2218	0,2218

**Figure 7.** Penjumlahan Vektor

Selanjutnya yaitu mencari nilai jarak kemiripan, seperti contoh dibawah ini:

Hasil Perhitungan Tabel								
D1	D2	D3	D4	D5	D6	D7	D8	D10
0,0549	0,3756	0	0	0,140687	0	0	0,05495	0

**Figure 8.** Hasil Consine Similarity

Berdasarkan tabel di atas untuk perhitungan Consine Similarity bisa dilihat seperti contoh di bawah ini:

Setelah perhitungan Consine Similarity selesai, selanjutnya mengurutkan dokumen dari nilai terbesar ke terkecil. lalu menentukan nilai  $k$ . Disini peneliti menggunakan nilai  $k = 3$  nilai yang paling tinggi kemiripan dengan data testing, berikut penjelasan hasil klasifikasi:

Hasil Klasifikasi									
1	2	3	4	6	7	8	9	10	
D2		D5	D8	D1					

**Table 8.** Hasil Klasifikasi

Dengan hasil kemiripan tersebut dapat diklasifikasikan bahwa D9 termasuk kelas (positif). Berikut ini percobaan dengan menambahkan nilai  $k$  dengan batasan angka ganjil.

Jika  $k = 1$ , maka D9 masuk ke kelas C1 (positif), diwakili oleh dokumen D2

Jika  $k = 3$ , maka D9 masuk ke kelas C2 (negatif), diwakili oleh dokumen D2, D5 dan D8

Jika  $k=5$ , maka D9 masuk ke kelas C1 (positif), diwakili oleh dokumen D2,D5,D8, dan D1.

## Hasil dan Pembahasan

Proses pengambilan data dari API Twitter sekaligus menampilkan hasil sentimen dari komentar netizen.

index	Tanggal	Pengguna	Tweet	Sentimen
0	2023-03-01 03:13:57	harisN0	_uncover_ID _i love , cuma masih banyak seller yg malah nolak pakai pos. gtau kenapa	Positif
1	2023-03-01 03:13:48	Bilrondahood	RT _uncover: hallo _ID saya kirim hp oppo di garansi nokia, memang tidak ada video buka paket hp. bapak saya gantek, bapa	Netral
2	2023-03-01 03:13:29	Prabaya_Gilang	_Anton_miring _uncover _ID _ Kalau ini jangan bang hahaha Tidak aaminnn	Netral
3	2023-03-01 03:13:16	faridessultana	RT _uncover: hallo _ID saya kirim hp oppo di garansi nokia, memang tidak ada video buka paket hp. bapak saya gantek, bapa	Netral
4	2023-03-01 03:12:57	JNEDEPOK_ID	bingung mau kirim paket barang / dokumen kemana ? Ke JNE Depok aja Kami siap melayani Anda :)	Positif
5	2023-03-01 03:12:56	ratting	RT _uncover: hallo _ID saya kirim hp oppo di garansi nokia, memang tidak ada video buka paket hp. bapak saya gantek, bapa...	Netral
6	2023-03-01 03:12:05	AndiPandi	RT _uncover: hallo _ID saya kirim hp oppo di garansi nokia, memang tidak ada video buka paket hp. bapak saya gantek, bapa...	Netral
7	2023-03-01 03:11:43	happineszperson	RT _uncover: hallo _ID saya kirim hp oppo di garansi nokia, memang tidak ada video buka paket hp. bapak saya gantek, bapa...	Netral
8	2023-03-01 03:11:28	agumwicaksana	RT _uncover: hallo _ID saya kirim hp oppo di garansi nokia, memang tidak ada video buka paket hp. bapak saya gantek, bapa...	Netral
9	2023-03-01 03:11:14	AnungLuhung	_ _ID Kandani mending nggo si merah pengusa tajem kok, leres mboten pakde 🌸	Netral
10	2023-03-01 03:10:37	JNE_ID	Meski Kak, yuk langsung dicek! 😊	Netral

**Figure 9.** Data API Twitter

Penambahan kolom untuk pelabelan difungsikan untuk mempermudah mengkategorikan positif, netral, negatif.

## **Indonesian Journal of Innovation Studies**

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

Unnamed: 0	Tanggal	Pengguna	Tweet	Sentimen	label	
0	0	2023-03-01 03:13:57	harisN0	_uncover_ID_ i love , cuma masih bryk seller...	Positif	1
1	1	2023-03-01 03:13:48	Bilfromdahood	RT _uncover: hallo_ID saya kirim hp oppo di g...	Netral	0
2	2	2023-03-01 03:13:29	Prahaya_Gilang	_Anton_miring_uncover_ID_ Kalau ini jangan...	Netral	0
3	3	2023-03-01 03:13:16	fandassutisna	RT _uncover: hallo_ID saya kirim hp oppo di g...	Netral	0
4	4	2023-03-01 03:12:57	JNEDEPOK_ID	bingung mau kirim paket barang / dokumen keman...	Positif	1
5	5	2023-03-01 03:12:56	rativng	RT _uncover: hallo_ID saya kirim hp oppo di g...	Netral	0
6	6	2023-03-01 03:12:05	AndiParuli	RT _uncover: hallo_ID saya kirim hp oppo di g...	Netral	0
7	7	2023-03-01 03:11:43	happiesz1person	RT _uncover: hallo_ID saya kirim hp oppo di g...	Netral	0
8	8	2023-03-01 03:11:28	agumwicaksana	RT _uncover: hallo_ID saya kirim hp oppo di g...	Netral	0
9	9	2023-03-01 03:11:14	AnungLuhung	_ID Kandani mending nggo si merah penguasa t...	Netral	0
10	10	2023-03-01 03:10:37	JNE_ID	Masih Kak, yuk langsung dicek! 😊	Netral	0

**Figure 10.** Penambahan Kolom

Data training diambil beberapa sebagai contoh untuk menguji data tersebut.

**Figure 11.** *Data Training*

Pengambilan data testing sebanyak 20% dari total data keseluruhan.

```
array(['RT _hallo saya kirim paket hp android oppo di ganti sama hp nokia _ID https://t.co/Cj94M4Q9eb',
       'RT _uncover: hallo _ID saya kirim hp oppo di ganti nokia, menang tidak ada video buka paket hp. bapak saya gaptek,
       bapak_',
       '_Anton_miring _uncover _ID _Kalo ini jangan bang hehehe Tidak aslinya',
       'RT : _uncover _ID _ mulalish pakai kirin macbook tinggal masukin sana doschnya...',
       '_uncover _ID _ Saya udah lebih dari 5x kirim barang elektronik pake JNE. Nggak pernah pakai asuransi, paling
       mentok minta packing kayu. Pertama kalinya says ngirin hp secon. Udah di dusun dari rumah. Nyampe counter JNE dibuka lagi
       dusunya. Apa betul isinya hp atau bukan. Count',
       '__ ID awneuh pindah berarti han wkwk',
       '_mate (2/2) Harap lebih berhati-hati dengan penipuan yang mengatasnamakan JNE ya kak. Jika masih ada yang ingin
       ditanyakan atau kurang jelas silahkan hubungi kami kembali melalui media social kami di twitter: . Selamat beraktivitas
       kembalii :) Thanks - Ally -',
       'RT _hallo _ID saya kirim hp oppo di ganti nokia, menang tidak ada video buka paket hp. bapak saya gaptek,
       bapak_',
       'RT _uncover: hallo _ID saya kirim hp oppo di ganti nokia, menang tidak ada video buka paket hp. bapak saya gaptek,
       bapak_',
       'RT _uncover: hallo _ID saya kirim hp oppo di ganti nokia, menang tidak ada video buka paket hp. bapak saya gaptek,
       bapak_',
       '_uncover _ID _ Itu pengalaman saya pribadi. Tidak tau apakah prosedur itu masih ada atau tidak.',
       'RT : _uncover _ID _ mulalish pakai kirin macbook tinggal masukin sana doschnya...',
       'RT _uncover: hallo _ID saya kirim hp oppo di ganti nokia, menang tidak ada video buka paket hp. bapak saya gaptek,
       bapak_',
       'RT _uncover: hallo _ID saya kirim hp oppo di ganti nokia, menang tidak ada video buka paket hp. bapak saya gaptek,
```

**Figure 12.** Data Testing

Tahap ini merupakan tahap membersihkan kata sebelum diolah lebih lanjut pada tahap berikutnya. Tahap preprocessing sangat penting, karena pada tahap ini semua perubahan kata akan memiliki nilai sendiri yang mempengaruhi akurasi sistem yang akan dibuat nantinya.

# Indonesian Journal of Innovation Studies

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

```
✓ 0s  csv_data['Tweet']

 0   _uncover _ID _ i love , cuma masih bnyk seller...
 1   RT _uncover: hallo _ID saya kirim hp oppo di g...
 2   _Anton _miring _uncover _ID _ Kalau ini jangan...
 3   RT _uncover: hallo _ID saya kirim hp oppo di g...
 4   bingung mau kirim paket barang / dokumen keman...
 ...
 95  RT _uncover: hallo _ID saya kirim hp oppo di g...
 96  RT _uncover: hallo _ID saya kirim hp oppo di g...
 97  RT _: hallo _ID saya kirim hp oppo di ganti no...
 98  RT _uncover: hallo _ID saya kirim hp oppo di g...
 99  RT _uncover: hallo _ID saya kirim hp oppo di g...

Name: Tweet, Length: 100, dtype: object
```

**Figure 13.** Preprocessing

Tahapan tokenizing merupakan tahapan proses untuk menguraikan deskripsi yang awalnya berupa kalimat-kalimat menjadi kata-kata dan menghilangkan delimiter seperti underscore (\_), titik dua (:), dan spasi pada karakter kata pada sebuah kalimat.

```
✓ 0s  def identify_tekens(row) :
    text = row['Tweet']
    tokens = nltk.word_tokenize(text)
    token_words = [w for w in tokens if w.isalpha()]
    return token_words

csv_data['Tweet'] = csv_data.apply(identify_tekens, axis=1)
csv_data.Tweet

 0   [uncover, id, i, love, cuma, masih, bnyk, sell...
 1   [rt, uncover, hallo, id, saya, kirim, hp, oppo...
 2   [anton, miring, uncover, id, kalau, ini, jangan...
 3   [rt, uncover, hallo, id, saya, kirim, hp, oppo...
 4   [bingung, mau, kirim, paket, barang, dokumen, ...

 ...
 95  [rt, uncover, hallo, id, saya, kirim, hp, oppo...
 96  [rt, uncover, hallo, id, saya, kirim, hp, oppo...
 97  [rt, hallo, id, saya, kirim, hp, oppo, di, gan...
 98  [rt, uncover, hallo, id, saya, kirim, hp, oppo...
 99  [rt, uncover, hallo, id, saya, kirim, hp, oppo...

Name: Tweet, Length: 100, dtype: object
```

**Figure 14.** Tokenizing

Stemming mengubah kata yang berimbuhan menjadi kata dasar, pada tahap ini menggunakan library dari sastrawi untuk melakukan perubahan kalimat.

```
✓ 0s  def stem_list(row) :
    text = row['Tweet']
    stem = [stemming.stem(word) for word in text]
    return(stem)

csv_data['Tweet'] = csv_data.apply(stem_list, axis=1)
csv_data.Tweet

 0   [uncov, id, i, love, cuma, masih, bnyk, seller...
 1   [rt, uncov, hallo, id, saya, kirim, hp, oppo, ...
 2   [anton, mire, uncov, id, kalau, ini, jangan, b...
 3   [rt, uncov, hallo, id, saya, kirim, hp, oppo, ...
 4   [bingung, mau, kirim, paket, barang, dokumen, ...

 ...
 95  [rt, uncov, hallo, id, saya, kirim, hp, oppo, ...
 96  [rt, uncov, hallo, id, saya, kirim, hp, oppo, ...
 97  [rt, hallo, id, saya, kirim, hp, oppo, di, gan...
 98  [rt, uncov, hallo, id, saya, kirim, hp, oppo, ...
 99  [rt, uncov, hallo, id, saya, kirim, hp, oppo, ...

Name: Tweet, Length: 100, dtype: object
```

**Figure 15.** Stemming

# Indonesian Journal of Innovation Studies

Vol. 23 (2023): July

DOI: 10.21070/ijins.v22i.883 . Article type: (Innovation in Computer Science)

Proses menghilangkan kata yang tidak penting, guna merampingkan perhitungan terhadap suatu dokumen.

Unnamed: 0			Tweet	label
0	0	[uncov, id, i, love, cuma, masih, bnyk, seller...		1
1	1	[rt, uncov, hallo, id, saya, kirim, hp, oppo, ...		0
2	2	[anton, mire, uncov, id, kalaupun, ini, jangan, b...		0
3	3	[rt, uncov, hallo, id, saya, kirim, hp, oppo, ...		0
4	4	[bingung, mau, kirim, paket, barang, dokumen, ...		1
5	5	[rt, uncov, hallo, id, saya, kirim, hp, oppo, ...		0
6	6	[rt, uncov, hallo, id, saya, kirim, hp, oppo, ...		0
7	7	[rt, uncov, hallo, id, saya, kirim, hp, oppo, ...		0
8	8	[rt, uncov, hallo, id, saya, kirim, hp, oppo, ...		0
9	9	[id, kandani, mend, nggo, si, merah, penguasa, ...		0
10	10	[masih, kak, yuk, langsung, dicek]		0
11	11	[rt, uncov, hallo, id, saya, kirim, hp, oppo, ...		0
12	12	[rt, uncov, hallo, id, saya, kirim, hp, oppo, ...		0
13	13	[rt, uncov, hallo, id, saya, kirim, hp, oppo, ...		0
14	14	[rt, uncov, hallo, id, saya, kirim, hp, oppo, ...		0

Figure 16. Stopword

Proses menghitung jumlah kata yang akan muncul pada setiap dokumen.

TF-IDF Vectorizer															\
	0	1	2	3	4	5	6	7	8	9	...	429	430	431	\
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
5	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	
96	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
97	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
98	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
99	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
100	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
	432	433	434	435	436	437	438								
1	0	0	0	0	0	1	0								
2	0	0	0	0	0	0	0								
3	0	0	0	0	0	0	0								
4	0	0	0	0	0	0	0								
5	0	0	0	0	0	0	0								
..	..	..	..	..	..	..	..								
96	0	0	0	0	0	0	0								
97	0	0	0	0	0	0	0								
98	0	0	0	0	0	0	0								
99	0	0	0	0	0	0	0								
100	0	0	0	0	0	0	0								

Figure 17. Perhitungan TF-IDF

Klasifikasi

```
from sklearn.metrics import classification_report, confusion_matrix
cm = confusion_matrix(y_test, y_pred)
cr = classification_report(y_test, y_pred)
print(cm)
print(cr)

[[18  0]
 [ 2  0]]
      precision    recall  f1-score   support

          0       0.90      1.00     0.95      18
          1       0.00      0.00     0.00       2

   accuracy                           0.90      20
  macro avg       0.45      0.50     0.47      20
weighted avg       0.81      0.90     0.85      20
```

Figure 18. Klasifikasi

## Kesimpulan

Berdasarkan hasil penelitian yang telah diuraikan, dapat disimpulkan bahwa metode Klasifikasi K-Nearest Neighbor (KNN) dengan pembobotan TF-IDF mampu menghasilkan analisis sentimen terhadap komentar-komentar di Twitter terkait pelayanan JNE dengan akurasi yang tinggi, mencapai 90%. Temuan ini membuktikan bahwa penggunaan data komentar dari media sosial dapat digunakan sebagai sumber informasi yang relevan untuk memahami pandangan publik terhadap layanan suatu perusahaan. Implikasi dari penelitian ini adalah penerapan metode analisis sentimen pada platform media sosial dapat menjadi alat yang efektif bagi perusahaan seperti JNE untuk memantau dan mengevaluasi persepsi pelanggan serta meningkatkan kualitas pelayanan berdasarkan umpan balik yang diterima. Namun, penelitian ini juga memperlihatkan adanya potensi untuk peningkatan lebih lanjut, misalnya dengan menggunakan metode-metode klasifikasi lainnya, memperluas sampel data, atau mempertimbangkan konteks budaya dalam analisis sentimen. Oleh karena itu, penelitian lebih lanjut di bidang ini diharapkan dapat mengoptimalkan penggunaan analisis sentimen untuk kepentingan perusahaan dan meningkatkan pemahaman tentang opini publik melalui media sosial.

## References

1. S. Paleon, H. Harcici, and Y. Septian, "Pengiriman Kargo Untuk Mendukung Transaksi Pengiriman Barang Untuk Peningkatan Transaksi Biaya Pengiriman Lebih Efisien," db, vol. 1, no. 1, pp. 7-14, May 2020, doi: 10.46484/db.v1i1.189.
2. M. Christin and L. Rahayu, "Kritik Sosial Pada Media Sosial (Analisis Semiotika Pada Youtube 'TV, Jasamu Tiada...')," J-IKA : Jurnal Ilmu Komunikasi Fakultas Ilmu Komunikasi Universitas BSI Bandung, vol. 4, no. 1, pp. 36-46, Apr. 2017, doi: 10.31294/kom.v4i1.1789.
3. F. A. Prasetyo, "Analisis Perbandingan Service Quality Antara JNE dan J&T Ekpress," Journal:eArticle, Petra Christian University, 2017.
4. "JNE," Wikipedia bahasa Indonesia, ensiklopedia bebas. Jul. 20, 2023. Accessed: Jul. 25, 2023. [Online]. Available: <https://id.wikipedia.org/w/index.php?title=JNE&oldid=23878471>.
5. A. Fadly, "Evaluative Language in the Discourse of Cebong Vs Kampret ('Tadpole Vs Microbats') on Twitter," Bahtera, vol. 19, no. 1, pp. 1-14, Jan. 2020, doi: 10.21009/bahtera.191.01.
6. M. A. Ramdhani and O. N. Rahim, "Analisis sentimen untuk mengukur popularitas tokoh publik berdasar data pada media sosial twitter menggunakan algoritma data mining dengan teknik klasifikasi," Informasi, vol. VI, no. 2, pp. 1-15, 2014.
7. E. Helmud, "Optimasi Basis Data Oracle Menggunakan Complex View Studi Kasus : PT. Berkat Optimis Sejahtera (PT.BOS) Pangkalpinang," J. Informatika, vol. 7, no. 1, pp. 80-86, 2021.
8. F. F. Mailoa, "Analisis Sentimen Data Twitter Menggunakan Metode Text Mining Tentang Masalah Obesitas di Indonesia," JISPH, vol. 6, no. 1, pp. 44-51, 2021, doi: 10.22146/jisph.44455.
9. M. Romzi and B. Kurniawan, "Pembelajaran Pemrograman Python Dengan Pendekatan Logika Algoritma," JTIM J. Tek. Inform. Mahakarya, vol. 03, no. 2, pp. 37-44, 2020.
10. H. Hartanto, "Text mining dan sentimen analisis twitter," J. Psikol. Ilm., vol. 9, no. 1, pp. 18-25, 2017.
11. A. Firdaus and W. I. Firdaus, "Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan)," J. JUPITER, vol. 13, no. 1, p. 66, 2021.