



**IMPLEMENTASI MULTINOMIAL NAIVE BAYES UNTUK
KLASIFIKASI SENTIMEN TERHADAP PELAYANAN PERUSAHAAN
OTOBUS MENGGUNAKAN DATA FACEBOOK**

TUGAS AKHIR

DONY JACARRIA PANGESTU
41516010127

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022



**IMPLEMENTASI MULTINOMIAL NAIVE BAYES UNTUK
KLASIFIKASI SENTIMEN TERHADAP PELAYANAN PERUSAHAAN
OTOBUS MENGGUNAKAN DATA FACEBOOK**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:
DONY JACARRIA PANGESTU
41516010127

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022

LEMBAR PERNYATAAN ORISINALITAS

LEMBAR PERNYATAAN ORISINALITAS

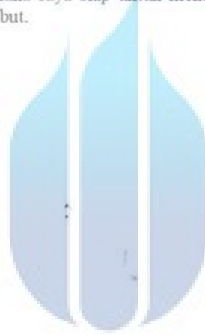
Yang bertanda tangan dibawah ini:

NIM : 41516010127

Nama : Dony Jacarria Pangestu

Judul Tugas Akhir : Implementasi Multinomial Naive Bayes untuk klasifikasi sentimen terhadap pelayanan perusahaan otobus menggunakan data facebook

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.



Tangerang, 25 Juni 2022



Dony Jacarria Pangestu

UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Dony Jacarria Pangestu
NIM : 41516010127
Judul Tugas Akhir : Implementasi Multinomial Naive Bayes untuk klasifikasi sentimen terhadap pelayanan perusahaan otobus menggunakan data facebook

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai peneliti/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS
MERCU BUANA

Tangerang, 25 Juni 2022



DONY JACARRIA PANGESTU

SURAT PERNYATAAN LUARAN TUGAS AKHIR

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Dony Jacarria Pangestu
NIM : 41516010127
Judul Tugas Akhir : Implementasi Multinomial Naive Bayes untuk klasifikasi sentimen terhadap pelayanan perusahaan otobus menggunakan data facebook

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis	Status
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi	
		Jurnal Nasional Terakreditasi	✓
		Jurnal International Tidak Bereputasi	
		Jurnal International Bereputasi	Diterima
	Disubmit/dipublikasikan di :	Nama Jurnal : Jurnal Pengembangan IT	
	ISSN : 2477 - 5126		
	Link Jurnal :		
	Link File :		
	Jurnal Jika Sudah di Publish :		

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan.

Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 25 Juni 2022



DONY JACARRIA PANGESTU

LEMBAR PERSETUJUAN PENGUJI

LEMBAR PERSETUJUAN PENGUJI

NIM : 41516010127
Nama : Dony Jacarria Pangestu
Judul Tugas Akhir : Implementasi Multinomial Naive Bayes untuk
klasifikasi sentimen terhadap pelayanan perusahaan
otobus menggunakan data facebook

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 27 Juli 2022



(Ir. Emil R. Kaburuan, Ph.D., IPM)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41516010127
Nama : Dony Jacarria Pangestu
Judul Tugas Akhir : Implementasi Multinomial Naive Bayes untuk
klasifikasi sentimen terhadap pelayanan perusahaan
otobus menggunakan data facebook

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 27 Juli 2022



(Dwiki Jatikusumo, S.Kom,M.Kom)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

LEMBAR PERSETUJUAN PENGUJI

NIM : 41516010127
Nama : Dony Jacarria Pangestu
Judul Tugas Akhir : Implementasi Multinomial Naive Bayes untuk
klasifikasi sentimen terhadap pelayanan perusahaan
otobus menggunakan data facebook

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 27 Juli 2022



(Rahmat Rian Hidayat, ST, MMSI)

UNIVERSITAS
MERCU BUANA


LEMBAR PENGESAHAN

NIM : 41516010127
Nama : Dony Jacarria Pangestu
Judul Tugas Akhir : Implementasi Multinomial Naive Bayes untuk klasifikasi sentimen terhadap pelayanan perusahaan otobus menggunakan data facebook

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.


Jakarta, 27 Juli 2022

Menyetujui,




(Drs. ACHMAD KODAR, MT)
Dosen Pembimbing

Mengetahui,



(Wawan Gunawan, S.Kom, MT)
Koord. Tugas Akhir Teknik Informatika



(Ir. Emil K. Kaburuan, Ph.D., IPM.)
Ka. Prodi Teknik Informatika

KATA PENGANTAR

Puji syukur kita panjatkan kehadiran Allah SWT. yang telah memberikan nikmatNya, terutama nikmat kesehatan dan kesempatan sehingga proses pembuatan Tugas Akhir ini dapat peneliti laksanakan dengan baik. Begitupun dengan rahmat Allah SWT. tugas akhir dengan judul “Implementasi Multinomial Naive Bayes untuk klasifikasi sentimen terhadap pelayanan perusahaan otobus menggunakan data facebook” dapat peneliti selesaikan dengan baik. Peneliti menyadari bahwa tanpa bantuan dan bimbingan banyak pihak yang membantu dan berkontribusi dalam terselesaikanya tugas akhir ini. Segala bentuk bantuan, baik berupa dukungan moril dan material sangat membantu peneliti dalam mengumpulkan semangat dan keinginan untuk menyelesaikan studi. Oleh karena itu, peneliti mengucapkan terima kasih dengan ketulusan hati kepada pihak-pihak yang telah membantu selama pembuatan tugas akhir ini, yakni kepada:

1. Ir. Emil R. Kaburuan, Ph.D., IPM (Ka. Prodi Teknik Informatika)
2. Wawan Gunawan, S.Kom, MT (Koord. Tugas Akhir Teknik Informatika)
3. Drs. ACHMAD KODAR, MT (Dosen Pembimbing Tugas Akhir)
4. Harwikarya Dr.MT (Dosen Pembimbing Akademik)
5. Kedua Orangtua beserta Keluarga Besar
6. Fauziah Ayu (Calon Istri)
7. Dan teman-teman

Akhir kata, peneliti berharap semoga skripsi ini dapat bermanfaat khususnya bagi semua pihak yang berkepentingan. Dan semoga Allah SWT membalas semua amal dan kebaikan kepada pihak-pihak yang telah membantu menyelesaikan Tugas Akhir ini. Amiin Wassalamu’alaikum Warahmatullahi Wabarakatuh

Tangerang, 25 Juni 2022
Peneliti



Dony Jacarria Pangestu

DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN JUDUL	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR... ..	iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR.....	iv
LEMBAR PERSETUJUAN PENGUJI	v
ABSTRAK	ix
ABSTRACT	x
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xii
NASKAH JURNAL	1
KERTAS KERJA.....	6
BAB 1. LITERATUR REVIEW	7
BAB 2. ANALISIS DAN PERANCANGAN.....	14
BAB 3. SOURCE CODE	25
BAB 4. DATASET.....	45
BAB 5. TAHAPAN EKSPERIMEN.....	46
BAB 6. HASIL SEMUA EKSPERIMEN.....	47
DAFTAR PUSTAKA	52
LAMPIRAN DOKUMEN HAKI.....	54
LAMPIRAN KORESPONDENSI	56

NASKAH JURNAL

Jurnal Informatika: Jurnal pengembangan IT (JPIT), Vol.xx, No.xx, Bulan 2018

ISSN: 2477-5126
e-ISSN: 2548-9356

Implementasi Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Terhadap Pelayanan Perusahaan Otobus Menggunakan Data Facebook (Studi Kasus: Grup Facebook Murni Jaya Lovers)

Dony Jacarria Pangestu^{1*)}, Achmad Kodar²

^{1,2}Jurusan Informatika, Fakultas Ilmu Komputer, Universitas Mercubuana, Jakarta
^{1,2}Jl.Meruya Selatan No.1 Kembangan, Kota Jakarta Barat, DKI Jakarta, 11650, Indonesia
email: ¹donyjacariapangestu@gmail.com, ²achmad.kodar@mercubuana.ac.id

Abstract – Customer satisfaction can be seen from the good and bad quality of service. many people use the social network facebook to expressive opinions. Sentiment analysis is a blend of data mining and text mining. Therefore, a collection of Facebook posts in the form of text is collected for processing so that it can be used for sentiment analysis. Use of multinomial Naïve Bayes algorithm to optimize sentiment data classification results. There are three classes used, namely negative, neutral, and positive. Before being classified, raw data goes through preprocessing stage first to normalize words in data. use of TF-IDF method helps for data weighting, it is necessary to share data in the classification stage, namely test data and training data. Which is where the test data is 20%, and the training data is 80%. Testing this data using a confusion matrix. percentage of sentiment data on otobus company services obtained from facebook. with a percentage of 23% negative sentiment, 53% neutral sentiment, and 25% positive sentiment. Can the key word that service bus company is still considered neutral by custommers. From the results of the data test, value accuracy is 95%, precision is 95%, and recall is 95%. This shows that the method used has a fairly good level.

Abstrak – Kepuasan pelanggan dapat dilihat dari baik buruknya kualitas pelayanan. banyak masyarakat yang menggunakan jejaring sosial facebook untuk mengekspresikan opini. Analisis sentimen adalah paduan dari data mining dan teks mining. Oleh karena itu kumpulan postingan facebook yang berupa teks dikumpulkan untuk selanjutnya di olah agar bisa digunakan untuk analisis sentimen. Penggunaan algoritma multinomial Naïve Bayes agar hasil klasifikasi data sentimen lebih optimal. Ada tiga kelas yang digunakan, yaitu negative, netral, dan positif. Sebelum di klasifikasi, data mentah melewati tahap preprocessing terlebih dahulu untuk menormalisasi kata yang ada pada data. penggunaan metode TF-IDF membantu untuk pembobotan data. diperlukan pembagian data dalam tahap klasifikasi, Yang dimana data uji sebesar 20%, dan data latih sebesar 80%. Pengujian data ini menggunakan confusion matrix. jumlah persentase data sentimen terhadap pelayanan perusahaan otobus yang didapatkan dari facebook. dengan persentase 23% sentiment negatif, 53% sentiment netral, dan 25% sentiment Positif. Bisa disimpulkan bahwa pelayanan perusahaan otobus tersebut masih di anggap netral oleh penggunaannya. Dari hasil uji data diperoleh nilai akurasinya 95%, precision 95%, dan recall 95%. Penelitian ini menunjukkan bahwa metode yang digunakan memiliki tingkatan cukup baik.

Kata Kunci – sentimen, opini, multinomial naïve bayes, facebook, universitas mercubuana.

I. PENDAHULUAN

Jasa angkutan bis antar kota antar provinsi (AKAP) merupakan angkutan umum yang memiliki mobilitas tinggi. Sehingga persaingan antar pelaku bisnis tersebut menjadi sangat banyak dan ketat dalam menjalankan bisnisnya. Kenyamanan dan keamanan penumpang menjadi tolak ukur untuk pertimbangan suatu perusahaan otobus [1]. Peningkatan pelayanan merupakan salah satu upaya agar pencapaian kepuasan dapat bisa selesai segera [2].

Pada era saat ini, untuk mengetahui tingkat kepuasan pelanggan bisa dilakukan dengan komputer. tingkat perhitungan kepuasan pelanggan semakin mendekati akurat dengan adanya dukungan proses pembelajaran yang dilakukan oleh komputer. Salah satunya menggunakan proses data mining dan text mining. [4]. Salah satunya menggunakan media sosial. Opini tersebut diberikan untuk sebuah produk jasa, layanan, maupun intansi. Dalam analisis sentimen terdapat 3 jenis opini, yaitu opini *negative*, *neutral*, dan *positive* [5].

Dalam mengekspresikan opini, masyarakat menggunakan media sosial, salah satunya adalah *facebook*. Awalnya *facebook* hanya dipergunakan untuk media pertemanan virtual saja, seiring perkembangan zaman dan waktu berjalan, *facebook* mengaalami sedikit perubahan sebagai sarana interaaksi untuk mempengaruhi seseorang . Pada umumnya postingan seseorang yang ada pada *facebook* hanya digunakan untuk sebuah perihal dari para penggunaannya, dan berbagai banyak informasi, serta penyampaian berupa berita dan dapat juga mengungkap perasaan dan isi hati dari para penggunaannya [6]. Oleh sebab itu membutuhkan sebuah proses analisis yang tepat dan akurat agar bisa menyuguhkan isi dari informasi yang sangat berharga mengenai semua opini seseorang, yang selanjutnya akan dikumpulkan untuk proses analisis sentimen [7].

Dari paparan tren dan masalah yang ada, peneliti ingin melakukan penelitian mengenai Prediksi sentimen pelayanan perusahaan otobus menggunakan algoritma *Multinomial Naïve Bayes* dengan teks mining data *facebook*. Multinomial Naïve Bayes adalah model yang dikembangkan dari algoritma Naïve Bayes dan umumnya digunakan untuk klasifikasi teks dan dokumen. [8].

*) penulis korespondensi: Dony Jacarria Pangestu
Email: donyjacariapangestu@gmail.com

Penulis Pertama: Empat Kata Judul Atas ... (ditambah tanda titik tiga setelah judul)

1

II. PENELITIAN YANG TERKAIT

Beberapa penelitian terkait yang peneliti temukan yaitu yang pertama dari Wiyanto yang melakukan penelitian terkait analisis kepuasan pelanggan perusahaan otobus XYZ menggunakan metode "algoritma *Naive Bayes*." Hasil dari penelitian tersebut menghasilkan akurasi algoritma *Naive Bayes* sebesar 94%, dan 88% puas terhadap pelayanan servisnya [9].

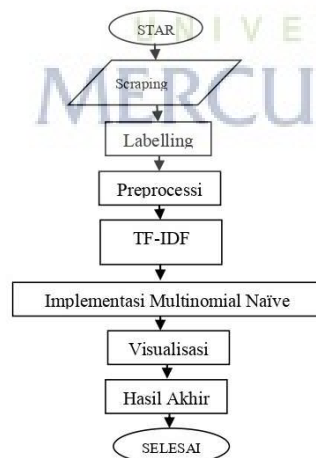
Selanjutnya penelitian dari Anggita Safitri Febrianti dan Erna Zuni Astuti melakukan penelitian mengenai prediksi kepuasan penumpang *Bus Rapid Transit* menggunakan algoritma C4.5. dari hasil tiga kali percobaan bahwa nilai tertinggi adalah akurasi 95%, *precision* 96,67%, dan *recall* 97,87% [10].

Selanjutnya jurnal dari Intl. *Journal on ICT Vol.2*, yang melakukan penelitian mengenai analisis sentiment twitter pada public transportasi menggunakan *Support vector machine*. Hasil Analisa menggunakan SVM menampilkan hasil dengan akurasi 78,12% pada dataset yang di klasifikasi menggunakan SVM [11].

III. METODE PENELITIAN

3.1 Alur Sistem

Agar klasifikasi teks lebih terarah, peneliti membagi dalam lima tahapan, yaitu yang pertama peneliti melakukan pengumpulan data menggunakan bahasa pemrograman *python* dengan library *facebook scraper*. Yang kedua yaitu tahapan *preprocessing*, merupakan proses mempersiapkan Mengubah teks tidak terstruktur menjadi teks terstruktur yang selanjutnya siap untuk di proses. Pada tahapan *preprocessing* terdapat tahapan-tahapan lain seperti *case folding*, *cleansing*, *tokenizing*, *filtering*, dan *stemming* [12]. Tahap ketiga yaitu menerapkan metode TF-IDF untuk proses pembobotan masing-masing kata dalam kalimat. Tahapan keempat yaitu tahapan klasifikasi Multinomial Naive Bayes untuk melihat sentiment negative, netral, dan positif. Tahapan kelima atau tahapan terakhir adalah melakukan proses pengujian dan Latihan kalsifikasi menggunakan *confusion matrix* [13].



Gambar 1. Flowchart Sistem

3.2 Sumber Data

Penggunaan data dalam penelitian ini merupakan data yang diambil dari kumpulan postingan grup *facebook* Mumi Jaya Lovers yang berangota 62000 pengguna. Data postingan diperoleh dengan membuat program *scraping* yang menggunakan bahasa pemrograman *python*. Dataset mentah yang didapatkan peneliti berjumlah 25 dataset berformat *.csv* yang masing masing dataset berisi 25 record data. data-data itu selanjutnya melalui proses penggabungan data. hasil penggabungan data berhasil menciptakan satu dataset yang berisi 504 record data.

3.3 Multinomial Naive Bayes

Multinomial Naive Bayes adalah model yang dikembangkan dari algoritma Naive Bayes dan cocok untuk klasifikasi teks dan dokumen. Multinomial Naive bayes merupakan pembelajaran probabilistic yang berdasarkan teorema Bayes dan digunakan dalam "pemrosesan bahasa alami atau *Natural Language Processing (NLP)*" [14]. Multinomial Naive Bayes bekerja dengan konsep *term frequency* untuk menghitung seberapa banyak sebuah kata tersebut muncul dalam kalimat atau dokumen. dan menjelaskan apa kata tersebut muncul dalam kalimat atau tidak serta jumlah frekuensi kemunculanya. *Multinomial Naive Bayes* dapat di formulakan sebagai berikut [15].

$$P(p|n) \propto P(p) \prod_{1 < k < n} P(t_k|p)$$

"P(tpk) merupakan probabilitas yang ada pada dokumen teks (tk). n merupakan jumlah dokumen dan p merupakan polaritas." Untuk penghitungan polaritas yang memiliki kesamaan yang mirip di formulakan sebagai berikut,

$$P(t_k|p) = \frac{\text{count}(tk|p)+1}{\text{count}(t_p)+|V|}$$

"(tk | p) merupakan jumlah teks yang ada pada dokumen teks yang memiliki polaritas p dan (tp) berarti jumlah token yang ada pada teks postingan dengan polaritas p."

3.4 Validasi Model

Untuk menjalankan pengujian data, perlu mengevaluasi hasil prediksi. Peneliti menggunakan "*confusion matrix*" untuk menguji hasil dengan mencari akurasi, recall, dan presisi. Untuk hitung presisi terhadap "*confusion matrix*" dengan persamaan 5. "Presisi = TP/(TP+FP)" Sehingga *weighted average precision* = jumlah total aktual.

Aktual dibagi Jumlah data**precision*. Untuk menghitung recall, bagilah jumlah klasifikasi yang benar dengan jumlah sebenarnya. Selanjutnya untuk menghitung akurasi dari ketiga kelas dari seluruh data uji peneliti gunakan persamaan 3 [16].

"*Accuracy* = TP + TN / TP + TN + FP + FN
TP adalah *true positive*, FN adalah *false negative*, FP adalah *false positive*, dan TN adalah *true negative*."

IV. HASIL DAN PEMBAHASAN

4.1 Preprocessing

Peneliti memperoleh data hasil dari *scraping* berupa data mentah. Melalui proses *preprocessing*, data-data mentah tersebut perlu di transformasi ke format yang dapat di pahami [17]. Tahapan-tahapan *preprocessing* data yaitu, yang pertama *case folding*, merupakan proses mengubah semua huruf kapital (*uppercase*) di dalam data menjadi huruf kecil (*lowercase*). Dalam penelitian ini, peneliti mengambil sample berikut, tabel 1.

Tabel 1. *case folding*

Sebelum <i>case folding</i>	Sesudah <i>case folding</i>
Mas Mbak mohon info ada gak Murni jaya dari ciputat - ke gombong kebumen ?	mas mbak mohon info ada gak murni jaya dari ciputat - ke gombong kebumen ?

Kedua yaitu *Text cleaning*, proses penghapusan karakter text yang tidak terpakai berupa emoticon, whitespace (\n atau \t), dan beberapa symbol baca , seperti tanda (.) titik, (?) tanda tanya, dan lainnya [18]. Contoh sampel *text cleaning* ada pada tabel 2.

Tabel 2. *Text cleaning*

Sebelum <i>text cleaning</i>	Sesudah <i>text cleaning</i>
mas mbak mohon info ada gak murni jaya dari ciputat - ke gombong kebumen ?	mas mbak mohon info ada gak murni jaya dari ciputat ke gombong kebumen

Tahap ketiga adalah *Normalization*, merupakan tahapan untuk menyeragamkan term yang memiliki makna sama namun penulisiannya berbeda, bisa diakibatkan oleh kesalahan dalam penulisan (*typo*), penyingkiran kata, ataupun bahasa gaul. Tahapan selanjutnya yaitu *tokenizing*, yaitu proses membagi teks kalimat atau paragraph untuk menjadi bagian bagian tertentu [19]. Contohnya ada pada tabel 3.

Tabel 3. *Tokenizing*

Sebelum <i>tokenizing</i>	Sesudah <i>tokenizing</i>
mas mbak mohon info ada gak murni jaya dari ciputat - ke gombong kebumen ?	["mas", "mbak", "mohon", "info", "ada", "gak", "murni", "jaya", "dari", "ciputat", "ke", "gombong", "kebumen"]

Tahap terakhir yaitu *stemming*, merupakan sebuah proses pengubah kalimat dan kata menjadi sebuah kata dasar. Kata-kata tersebut diambil bentuk dasarnya dengan menghilangkan awalan atau akhiran dari kata-kata tersebut. Setiap kata dalam postingan akan diperiksa secara otomatis dari awal sampai akhir kata pada kalimat tersebut [20].

4.2 Hasil Pengujian Multinomial Naïve Bayes

Sebelum proses pengujian data, data-data diolah terlebih dahulu menggunakan TF-IDF, merupakan sebuah frekuensi (*Term Frequency*) dari kemunculan term dalam kalimat yang

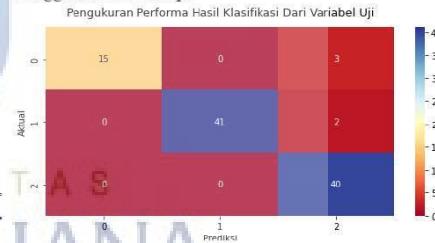
bersangkutan. Semakin besar kemunculan term, maka semakin besar juga bobotnya. Untuk IDF (*inverse Document Frequency*) merupakan perhitungan dari bagaimana term didistribusikan luas koleksi dokumen yang bersangkutan. Semakin dikit jumlah dokumen yang mengandung term, maka nilai IDF semakin besar [21]. Selanjutnya peneliti menerapkan library SMOTE untuk mengatasi label data yang tidak seimbang / *imbalance*.

Dalam pengujian ini peneliti menggunakan total 504 data yang dimana akan di bagi menjadi data uji dan data latih. mengolah data uji sebanyak 101 postingan *facebook* yang terdapat pada Grup *Facebook Murni Jaya Lovers*. Kemudian peneliti mengklasifikasikan dalam tiga kelas, yaitu kelas negatif ,netral, dan positif, diperlukan pembagian data dalam tahap klasifikasi, yaitu data uji dan data latih. Yang dimana data uji 20%, dan data latih 80%. Pengujian data ini menggunakan *confusion matrix*. peneliti menggunakan pengukuran terhadap *precision*, *recall*, dan *accuracy*.

	precision	recall	f1-score	support
negatif	1.00	0.83	0.91	18
netral	1.00	0.95	0.98	43
positif	0.89	1.00	0.94	40
accuracy		0.95		101

Gambar 2. Pengujian terhadap akurasi, *precision*, dan *recall*

Dari hasil uji algoritma *Multinomial Naïve Bayes*, bahwa performanya memiliki akurasi sebesar 95%. Terdapat tiga kelas yang menjadi output dalam penelitian ini, sehingga pengujian ini menggunakan "*confusion matrix* skala 3x3." Berikut merupakan gambar grafik *confusion matrix system* menggunakan *Heatmap*.



Gambar 3. Pengukuran Data Uji

Pada gambar 2, "kelas negative dilambangkan angka 0, kelas netral dilambangkan dengan angka 1, kelas positif dilambangkan dengan angka 2. Peneliti melakukan penghitungan *akurasi*, *precision*, dan *recall* secara manual." Bisa dilihat pada tabel 4.

Tabel 4. Menghitung Confusion Matrix

Aktual \ Prediksi	Prediksi Negatif	Prediksi Netral	Prediksi Positif	Total Aktual
Negatif	15	0	3	18
Netral	0	41	2	43
Positif	0	0	40	40
Total Prediksi	15	41	45	101

Pengujian menggunakan *confusion matrix* berskala 3x3 memperoleh nilai TP 15 untuk negatif, TP 41 untuk netral, dan TP 40 untuk positif. Proses selanjutnya menghitung nilai *precision* dengan menguji seberapa persen nilai dari ketiga kelas itu yang sudah benar-benar negatif, netral, dan positif. Penghitungan presisi menggunakan persamaan 5.

Tabel 5. Perhitungan Precision

Kelas	Perhitungan	Hasil
Negatif	15/15	1.00
Netral	41/41	1.00
Positif	40/45	0,889

Weighted average precision = Jumlah total aktual Aktual / Jumlah data * *precision*
 $(18/101*1) + (43/101*1) + (40/101*0,889) = 0,9560$
 Untuk bisa menghitung *precision* yaitu membagi antara jumlah akurasi yang benar (aktual) dengan total keseluruhan prediksi. Setelah perhitungan dilakukan, maka mendapatkan hasil 0,9560 atau dalam persentase 95%.

Setelah berhasil menghitung *precision*, peneliti melanjutkan dengan menghitung *recall*. Pengujian ini untuk mengetahui seberapa persen ketiga sentiment yang benar-benar negatif, netral, dan positif dari seluruh kelas sebenarnya.

Tabel 6. Perhitungan Recall

Kelas	Perhitungan	Hasil
Negatif	15/18	0,8333
Netral	41/43	0,9534
Positif	40/40	1.00

Untuk bisa menghitung sebuah *recall* bisa dengan cara membagi seluruh jumlah klasifikasi yang benar dengan seluruh total aktual. Sehingga diperoleh.

$$(18/101*0,8333) + (43/101*0,9534) + (40/101*1) = 0,9504$$

Maka hasil yang didapatkan sebesar 0,9504 atau dalam persentase adalah 95%. Selanjutnya peneliti menghitung akurasi dari ketiga kelas sentiment tersebut yang benar benar negatif, netral, dan positif dari seluruh data.

$$\text{Akurasi} = \text{Total keseluruhan klasifikasi benar}$$

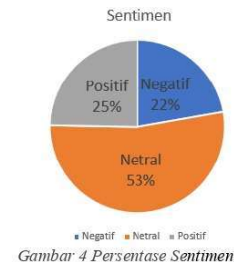
$$15 + 41 + 40 = 96$$

Kemudian, Hasil dibagi dengan Total data

$96/101 = 0,9504$ atau dalam persentase sebesar 95%.

V. KESIMPULAN

Berdasarkan analisis diatas, peneliti menyimpulkan bahwa system bekerja sangat baik untuk melakukan proses kategorisasi sentiment teks didalam kelas negatif, netral, dan positif terhadap kondisi pelayanan Perusahaan Otobus. Berdasarkan analisis diatas, Berikut hasil presentasi kelas sentiment terhadap pelayanan perusahaan otobus menggunakan data facebook.



Berdasarkan gambar 4, terlihat jumlah persentase data sentiment terhadap pelayanan perusahaan otobus yang didapatkan dari facebook. dengan persentase 23% sentiment negatif, 53% sentiment netral, dan 25% sentiment Positif. Bisa disimpulkan bahwa pelayanan perusahaan otobus tersebut masih di anggap netral oleh penggunanya.

Peneliti melakukan analisis dengan Algoritma Multinomial Naïve Bayes menghasilkan akurasi sebesar 95%, untuk *precision* sebesar 95%, dan *recall* sebesar 95%. Menandakan bahwa algoritma Multinomial Naïve Bayes memiliki tingkat *fair classification* dan nilai diagnostic cukup baik. Untuk penelitian selanjutnya, peneliti akan mempertimbangkan penggunaan sosial media lain. Dan perlu menambahkan algoritma yang lain untuk bisa membantu kinerja sebuah algoritma Multinomial Naïve Bayes didalam sebuah proses pengklasifikasian sentiment.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada pihak-pihak yang telah memberikan dukungan dan dukungan sehubungan dengan penelitian yang dilakukan, seperti bantuan fasilitas penelitian, bantuan ilmu, maupun bantuan lainnya yang tidak bisa peneliti sebutkan satu per satu.

DAFTAR PUSTAKA

- [1] C. A. Saptomo and Y. P. Krisnadi, "Pengaruh Kepuasan Konsumen dan Insentif terhadap Perilaku WOM (Word of Mouth) Konsumen Jasa Angkutan Penumpang Bus Antar Kota Propinsi Kelas Eksekutif di Yogyakarta * Yehizkia Putra Krisnadi , Chrsientianus Abdi Saptomo," *J. Bisnis dan Akunt.*, vol. 12, no. 2, pp. 63–78, 2018, [Online]. Available: <http://www.e-jurnal.ukrimuniversity.ac.id/file/5.Yehizkia & Chrsientianus.pdf>
- [2] D. Saidah, "Kualitas Pelayanan Commuter Line," *J. Manaj. Transp. Dan Logistik*, vol. 4, no. 1, p. 51, 2017, doi: 10.25292/j.mtl.v4i1.47.
- [3] I. B. N. Sanditya Hardaya, A. Dhini, and I. Surjandari, "Application of text mining for classification of community complaints and proposals," *Proceeding - 2017 3rd Int. Conf. Sci. Inf. Technol. Theory Appl. IT Educ. Ind. Soc. Big Data Era, ICSITech 2017*, vol. 2018-January, pp. 144–149, 2017, doi: 10.1109/ICSITech.2017.8257100.
- [4] S. Ernawati and R. Wati, "Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Review Agen Travel," *J. Khatulistiwa Inform.*, vol. 6, no. 1, pp. 64–69, 2018.
- [5] A. Salam, J. Zeniarja, and R. S. U. Khasanah, "Analisis Sentimen Data Komentar Sosial Media Facebook Dengan K-Nearest Neighbor (Studi Kasus Pada Akun Jasa Ekspedisi Barang J&T Ekspres Indonesia)," *Pros. SINTAK*, pp. 480–486, 2018.
- [6] M. Rifauddin and A. N. Halida, "Waspada Cybercrime dan Informasi Hoax pada Media Sosial Facebook," *Khazanah al-Hikmah J. Ilmu Perpustakaan, Informasi, dan Kearsipan*, vol. 6, no. 2, p. 98, 2018, doi: 10.24252/kah.v6i2a2.
- [7] E. B. Santoso and A. Nugroho, "Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook," *Eksplora Inform.*, vol. 9, no. 1, pp. 60–69, 2019, doi: 10.30864/eksplora.v9i1.254.
- [8] 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.
- [9] Wiyanto, "Analisa Tingkat Kepuasan Pelanggan Terhadap Pelayanan Perusahaan Otobus XYZ Menggunakan Metode Naive Bayes," *J. Pelita Teknol.*, vol. 15, no. 1, pp. 56–67, 2020.
- [10] A. S. Febriarini and E. Z. Astuti, "Penerapan Algoritma C4.5 untuk Prediksi Kepuasan Penumpang Bus Rapid Transit (BRT) Trans Semarang," *Eksplora Inform.*, vol. 8, no. 2, pp. 95–103, 2019, doi: 10.30864/eksplora.v8i2.156.
- [11] V. Effendy, "Sentiment Analysis on Twitter about the Use of City Public Transportation Using Support Vector Machine Method," *Int. J. Inf. Commun. Technol.*, vol. 2, no. 1, p. 57, 2016, doi: 10.21108/ijoiect.2016.21.85.
- [12] M. A. Rosid, A. S. Fitriani, I. R. I. Astutik, N. I. Mulloh, and H. A. Gozali, "Improving Text Preprocessing for Student Complaint Document Classification Using Sastrawi," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 874, no. 1, 2020, doi: 10.1088/1757-899X/874/1/012017.
- [13] Yuyun, Nunul Hidayah, and Supriadi Sahibu, "Algoritma Multinomial Naive Bayes Untuk Klasifikasi Sentimen Pemerintah Terhadap Penanganan Covid-19 Menggunakan Data Twitter," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 820–826, 2021, doi: 10.29207/resti.v5i4.3146.
- [14] M. P. Munthe, A. S. R. Ansori, and ..., "Analisis Sentimen Komentar Pada Saluran Youtube Food Vlogger Berbahasa Indonesia Menggunakan Algoritma Naive Bayes," *eProceedings Eng.*, vol. 8, no. 6, pp. 11909–11916, 2021, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/16897>
- [15] G. Singh, B. Kumar, L. Gaur, and A. Tyagi, "Comparison between Multinomial and Bernoulli Naive Bayes for Text Classification," *2019 Int. Conf. Autom. Comput. Technol. Manag. ICACTM 2019*, pp. 593–596, 2019, doi: 10.1109/ICACTM.2019.8776800.
- [16] M. Y. H. Setyawan, R. M. Awangga, and S. R. Efendi, "Comparison Of Multinomial Naive Bayes Algorithm And Logistic Regression For Intent Classification In Chatbot," *Proc. 2018 Int. Conf. Appl. Eng. ICAE 2018*, pp. 1–5, 2018, doi: 10.1109/INCAE.2018.8579372.
- [17] M. Sholehuddin, M. Fauzi Ali, and S. Adinugroho, "Implementasi Metode Text Mining dan K-Means Clustering untuk Pengelompokan Dokumen Skripsi (Studi Kasus : Universitas Brawijaya)," vol. 2, no. 11, pp. 5518–5524, 2018.
- [18] M. N. Randhika, J. C. Young, A. Suryadibrata, and H. Mandala, "Implementasi Algoritma Complement dan Multinomial Naive Bayes Classifier Pada Klasifikasi Kategori Berita Media Online," *Ultim. J. Tek. Inform.*, vol. 13, no. 1, pp. 19–25, 2021, doi: 10.31937/ti.v13i1.1921.
- [19] Ratino, N. Hafidz, S. Anggraeni, and W. Gata, "Sentimen Analisis Informasi Covid-19 menggunakan Support Vector Machine dan Naive Bayes," *J. JUPITER*, vol. 12, no. 2, pp. 1–11, 2020.
- [20] Rianto, A. B. Mutiara, E. P. Wibowo, and P. I. Santosa, "Improving the accuracy of text classification using stemming method, a case of non-formal Indonesian conversation," *J. Big Data*, vol. 8, no. 1, pp. 1–16, 2021, doi: 10.1186/s40537-021-00413-1.
- [21] H. Zhou, "Research of Text Classification Based on TF-IDF and CNN-LSTM," *J. Phys. Conf. Ser.*, vol. 2171, no. 1, pp. 218–222, 2022, doi: 10.1088/1742-6596/2171/1/012021.

KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan tempat pengumpulan bahan material kelengkapan pembuatan sebuah artikel jurnal dengan judul Implementasi Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Terhadap Pelayanan Perusahaan Otobus Menggunakan Data Facebook (Studi Kasus: Grup Facebook Murni Jaya Lovers). Kertas kerja ini berisikan semua material – material hasil penelitian Tugas Akhir yang tidak dimuat atau disertakan pada artikel jurnal. Pada kertas kerja ini terdapat material seperti Literatur Review, Analisis perancangan, Source Code, Dataset, Tahapan eksperimen, serta Hasil Eksperimen secara keseluruhan.

Hasil penelitian ini merupakan hasil akurasi analisis terhadap pelayanan perusahaan otobus yang dilakukan dengan proses scrraping pada Grup Facebook Murni Jaya Lovers. Algoritma yang digunakan adalah Multinomial Naïve Bayes. Diharapkan hasil penelitian ini dapat menjadi bahan evaluasi pelayanan kepada Perusahaan Otobus PT. Murni Anugrah Jaya Usaha.

