

Analisis Sentimen Kepuasan Pelanggan Perusahaan Telekomunikasi Seluler Telkomsel di Twitter

Melia Haerunnissa ^{*1}, Agung Priyanto², Choerun Asnawi³, Nafisa Alfi Sa'diya⁴

^{1,2,3}Teknik Informatika, FTI Unjani, Yogyakarta, Indonesia

⁴Teknologi Informasi, FTI Unjani, Yogyakarta, Indonesia

e-mail: ¹[*1meliahhaerunnisa@gmail.com](mailto:meliahhaerunnisa@gmail.com), ²agungprie2010@gmail.com, ³c.asnawi@gmail.com,
⁴nafisaalfi204@gmail.com

Abstract - Telkomsel, the largest operator in Indonesia with the most users, collects a significant amount of tweet data on Twitter, containing both positive and negative feedback about their internet service. Analyzing this data can provide valuable insights and accurate information about Telkomsel's internet services based on user tweets, retweets, and comments. The aim is to build a sentiment analysis model to extract relevant information from Telkomsel users' tweets on Twitter, serving as feedback for service evaluation and an educational tool for users. The sentiment analysis process involves data retrieval, preprocessing, training, testing, classification, and visualization using Python programming with the Flask framework. Analysis of customer satisfaction sentiment reveals that Telkomsel has a negative sentiment, with an accuracy of 81.7% for training data and 84% for testing data. The sentiment analysis model was built using the Naive Bayes Classification method.

Keywords - Telkomsel, Naive Bayes Classification, WebApp, Sentiment Analysis, Python, Flask.

Abstrak - Telkomsel adalah operator terbesar di Indonesia yang masih memiliki jumlah pengguna terbanyak. Dengan peningkatan jumlah pengguna dan pengikut media sosial, Telkomsel memiliki banyak data tweet mengenai layanan internetnya di media sosial Twitter, yang belum optimal dalam penggunaannya. Data tersebut berupa tweet, retweet, dan komentar dari pengguna Telkomsel di Twitter dapat digunakan sebagai bahan penelitian untuk analisis dan identifikasi kebenaran guna mendapatkan informasi yang tepat mengenai layanan internet yang disediakan berdasarkan tweet, retweet, dan komentar dari pengguna Telkomsel di media sosial Twitter. Tujuan penelitian ini adalah membangun model analisis sentimen mengenai Telkomsel untuk mendapatkan informasi relevan berdasarkan tweet, retweet, dan komentar dari pengguna Telkomsel di media sosial Twitter, yang nantinya dapat digunakan sebagai masukan dalam evaluasi produk layanan agar lebih baik serta sebagai alat edukasi bagi pengguna Telkomsel. Analisis sentimen dengan kata kunci "Telkomsel" menggunakan 13.239 tweet yang diambil dari tanggal 20 April hingga 30 Juli. Proses analisis sentimen melibatkan beberapa tahap, antara lain pengambilan data, preprocessing, pelatihan,

pengujian, klasifikasi, dan visualisasi dalam bentuk diagram batang menggunakan bahasa pemrograman Python dengan kerangka kerja Flask. Berdasarkan hasil analisis sentimen kepuasan pelanggan terhadap perusahaan telekomunikasi seluler Telkomsel di Twitter, ditemukan bahwa nilai akurasi sebesar 81,7% untuk data pelatihan dan 84% untuk data pengujian. Analisis sentimen kepuasan pelanggan terhadap perusahaan telekomunikasi seluler Telkomsel di Twitter dibangun menggunakan bahasa pemrograman Python, kerangka kerja Flask, dengan metode Klasifikasi Naive Bayes, dan menunjukkan akurasi sebesar 81,7% untuk data pelatihan dan 84% untuk data pengujian. Berdasarkan hasil klasifikasi, Telkomsel memiliki sentimen negatif.

Kata kunci - Telkomsel, Naive Bayes Classification, WebApp, Analisis Sentimen, Python, Flask.

I. PENDAHULUAN

Layanan Internet sangat penting di era revolusi industri 4.0, dengan jumlah pengguna internet di seluruh dunia mencapai 4,95 miliar orang pada Januari 2022. Di Indonesia, pengguna internet terus meningkat dari tahun ke tahun, mencapai 202,6 juta pengguna pada Maret 2021, dan total pengguna per Januari 2022 mencapai 204,7 juta orang [1]. Kebutuhan akan jasa layanan komunikasi, terutama operator telekomunikasi seluler seperti Telkomsel, sangat tinggi. Telkomsel merupakan salah satu operator terbesar di Indonesia dengan 173,5 juta pelanggan pada tahun 2021 [2]. Namun, Telkomsel menghadapi beberapa kendala dalam layanan internet, seperti keterbatasan sinyal, penggunaan kuota yang tidak sesuai, dan keluhan terkait pulsa. Oleh karena itu, penting untuk melakukan analisis sentimen di media sosial Twitter guna mengetahui respons pengguna terhadap layanan Telkomsel berdasarkan tweet, retweet, dan komentar negatif yang ada [3].

Berdasarkan permasalahan tersebut, penelitian ini dilakukan untuk menganalisis sentimen Telkomsel di media sosial Twitter, dengan tujuan untuk memahami respons pengguna terhadap layanan internet Telkomsel [4]. Penelitian ini didasarkan pada landasan teori yang diambil dari beberapa jurnal terkait, seperti penelitian oleh [5]. mengenai analisis sentimen masyarakat terhadap virus corona berdasarkan opini dari Twitter berbasis web scraper,. Penelitian oleh [6]

mengenai sistem analisis sentimen pada ulasan produk menggunakan metode *Naive Bayes Classification*, serta Penelitian oleh [7] mengenai optimasi *TextBlob* menggunakan *Support Vector Machine* untuk analisis sentimen layanan Telkomsel. Tujuan utama penelitian ini adalah menciptakan model analisis sentimen dengan metode *Naive Bayes Classification* yang dapat digunakan untuk menganalisis sentimen di media sosial *Twitter* terkait dengan Telkomsel [8].

II. METODE PENELITIAN

Penelitian ini merupakan analisis sentimen menggunakan data dari *Twitter*, dengan metode *Naive Bayes Classification*. Penelitian ini dimulai dengan latar belakang permasalahan, pemetaan proses, mencari sumber permasalahan, dan menganalisis data mengenai Telkomsel di media sosial *Twitter*. Data *tweet* yang diperlukan diperoleh melalui *Twitter* yang berkaitan dengan Telkomsel [9]. Penelitian ini menggunakan bahan, alat, dan metode tertentu, serta melibatkan langkah-langkah penelitian untuk menyelesaikan analisis sentimen dan merancang sistem untuk menampilkan hasil analisis yang diperoleh.

A. Bahan dan Alat Penelitian

Penelitian ini akan mengidentifikasi data dan informasi dari berbagai sumber, termasuk tweet terkait Telkomsel di media sosial *Twitter*, buku analisis sentimen, jurnal, penelitian sebelumnya, dan informasi internet. Alat yang digunakan adalah komputer dengan spesifikasi yang memadai dan koneksi internet. Sistem operasi dan program aplikasi yang akan digunakan dalam pengembangan aplikasi penelitian ini akan ditentukan lebih lanjut.

Tabel 1. Alat dan Bahan

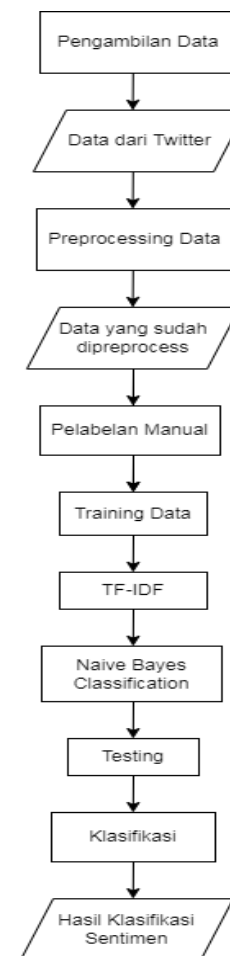
Processor	Intel(R) Celeron(R) CPU 1005M @ 1.90GHz (2CPUs), ~1.9GHz
RAM	10 GB
Sistem Operasi	Windows 7
Bahasa Pemrograman	Python 3.7.3
Framework	Flask, Bootstrap 4
Text editor	Sublime Text, Jupyter Notebook

Alat dan Bahan yang digunakan dalam penelitian ini seperti yang terdapat pada Tabel 1. Yang terdiri dari komputer dengan spesifikasi yang cukup untuk menjalankan sistem operasi dan pengembangan perangkat lunak serta koneksi Internet.

B. Jalan Penelitian

Naive Bayes Classification merupakan metode machine learning untuk klasifikasi data teks yang menggunakan perhitungan probabilitas dan statistik dengan menggabungkan frekuensi dan nilai-nilai

dataset [10]. Metode ini melibatkan langkah-langkah yang dieksekusi secara berurutan. Penelitian ini akan mengikuti alur penelitian dan langkah-langkah metode *Naive Bayes Classification* seperti yang ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

1. Tahap Pengambilan data melibatkan pengambilan tweet tentang Telkomsel dari media sosial *Twitter* menggunakan *Anaconda Prompt* dan dieksekusi pada *Jupyter Notebook*. Data tweet yang diperoleh dalam format file CSV dapat dibuka di *Microsoft Excel*.
2. Tahap *Preprocessing* data melibatkan pengolahan data teks yang belum terstruktur dengan beberapa tahapan, termasuk *tokenization*, *case folding*, *stopword removal*, *number removal*, *punctuation removal*, *whitespaces removal*, *normalization*, dan *stemming*.
3. Tahap Pelabelan manual melibatkan proses pelabelan manual untuk memberikan sentimen positif, netral, atau negatif pada kata-kata dalam dokumen agar dapat dianalisis.
4. Tahap *Training* data melibatkan proses ekstraksi fitur menggunakan *term frequency-inverse document frequency* (TF-IDF) pada

data teks. Proses training data digunakan untuk melakukan klasifikasi sentimen secara otomatis.

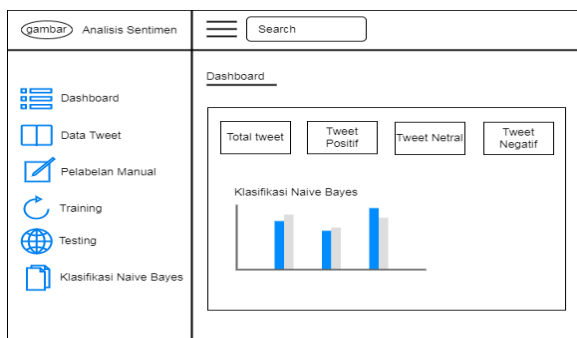
5. Tahap *Testing* melibatkan proses pengujian untuk mengukur akurasi model yang telah dibangun pada tahap training. Hasil pengujian akan diprediksi kelas dan labelnya, dan *confusion matrix* digunakan untuk menghitung persentase dari setiap pengujian.
6. Tahap Klasifikasi melibatkan proses di mana hasil prediksi kelas dan label dari data tweet akan ditampilkan secara otomatis dalam bentuk dashboard menggunakan *framework Flask* dengan menggunakan bahasa pemrograman *Python*.

C. Desain Interface

Interface merupakan gambaran yang berhubungan langsung dengan pengguna sistem. Perancangan antar muka dibuat agar sistem yang dihasilkan lebih menarik dan mudah dipahami pada saat perancangan. Berikut ini adalah deskripsi atau desain-desain *interface* pada Analisis Sentimen Kepuasan Pelanggan Perusahaan Telekomunikasi Seluler Telkomsel di Twitter. *Preprocessing* pada sistem analisis sentimen merupakan proses pengolahan data tweet yang ada dengan melakukan langkah-langkah untuk menyempurnakan data *tweet* yang masih memiliki komponen sehingga menjadi data yang diinginkan.

1. Dashboard Analisis Sentimen

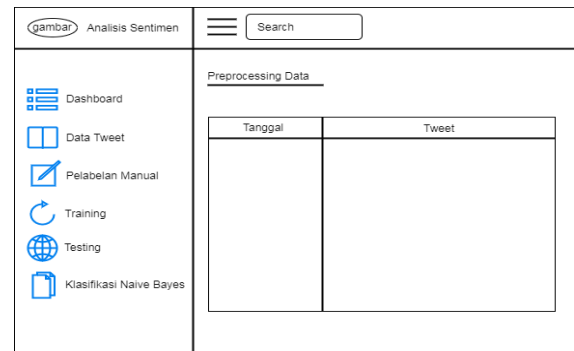
Pada sistem analisis sentimen dapat mengetahui data *tweet* yang telah didapatkan kelas dan labelnya berdasarkan prediksi yang sudah dibuat pada model *training* dan ditampilkan pada menu *dashboard*. Tampilan menu *dashboard* ditunjukkan pada Gambar 2.



Gambar 2. Desain Interface Halaman *Dashboard*

2. Preprocessing

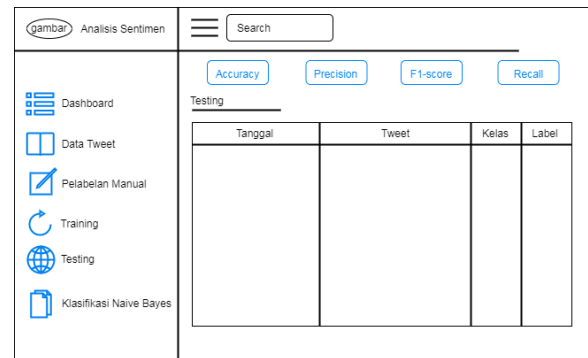
Preprocessing pada sistem analisis sentimen merupakan proses pengolahan data *tweet* yang ada dengan melakukan langkah-langkah untuk menyempurnakan data *tweet* yang masih memiliki komponen sehingga menjadi data yang diinginkan. Tampilan menu *preprocessing* data ditunjukkan pada Gambar 3.



Gambar 3. Desain Interface Halaman *Preprocessing* Data

3. Pelabelan Manual

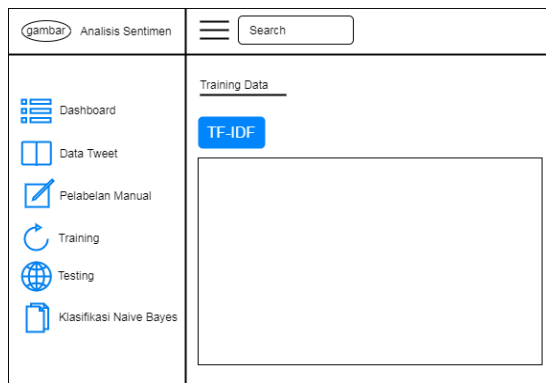
Pelabelan manual pada sistem analisis sentimen merupakan proses untuk menampilkan data *tweet* yang sudah diberi kelas dan label terhadap kalimat yang ada pada data *tweet* secara manual. Proses pelabelan manual dilakukan pada *Microsoft Excel*. Tampilan menu pelabelan manual ditunjukkan pada Gambar 4.



Gambar 4. Desain Interface Halaman Pelabelan Manual

4. Training

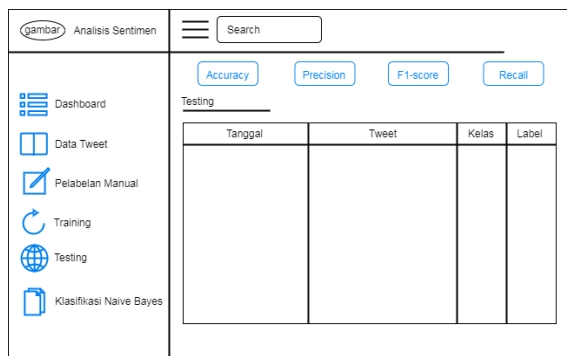
Training pada sistem analisis sentimen menggunakan ekstraksi *Term Frequency-Invers Document Frequency* (TF-IDF) pada data teks untuk menghasilkan model klasifikasi yang dapat digunakan untuk menampilkan proses klasifikasi sentimen secara otomatis. Tampilan menu *training* data ditunjukkan pada Gambar 5.



Gambar 5. Desain Interface Halaman Training

5. Testing

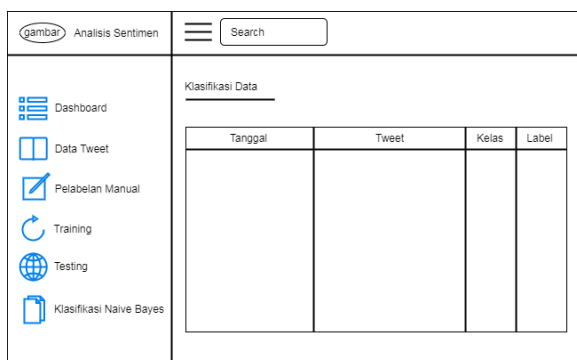
Testing pada sistem analisis sentimen dilakukan untuk mengetahui seberapa akurat model yang dibangun dalam *training* data digunakan untuk memprediksi kelas atau label dari data uji yang tersedia. Tampilan menu *training* data ditunjukkan pada Gambar 6.



Gambar 6. Desain Interface Halaman Testing

6. Klasifikasi Naive Bayes

Klasifikasi *Naive Bayes* pada sistem analisis sentimen dilakukan untuk menampilkan data *tweet* yang akan diprediksi kelas, label yang sudah dibangun pada *training* data. Tampilan menu klasifikasi *Naive Bayes* ditunjukkan pada Gambar 7.



Gambar 7. Desain Interface Klasifikasi Naive Bayes

III. HASIL DAN PEMBAHASAN

A. Ringkasan Hasil Penelitian

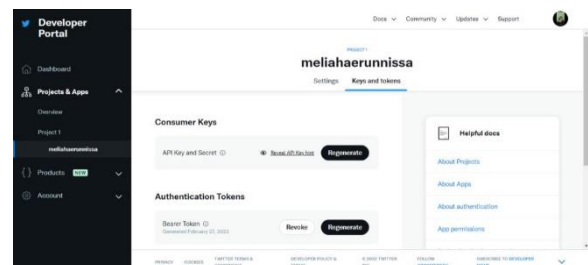
Penelitian ini menggunakan metode *Naive Bayes Classification* untuk menganalisis 13.239 data tweet berbahasa Indonesia tentang "Telkomsel" dari periode 20 April hingga 30 Juli 2022. Data tersebut terdiri dari 600 data *training* dan 336 data *testing*. Analisis sentimen dilakukan dengan bahasa pemrograman Python menggunakan *framework* Flask. Penelitian ini memberikan informasi tentang sentimen pengguna Telkomsel di Twitter terkait dengan perusahaan tersebut.

B. Pengambilan Data Twitter

Pengambilan data dilakukan di media sosial Twitter untuk mendapatkan data tweet yang berkaitan dengan kata kunci "Telkomsel". Data *tweet* yang diambil mulai dari tanggal 20 Mei – 30 Juli 2022 dengan 13.239 data *tweet*, *re-tweet* dan komentar. Proses pengambilan data tweet menggunakan API Twitter dilakukan di Anaconda Prompt dan dieksekusi di *Jupyter Notebook*. Data tweet yang dapat diambil berformat file *Comma Separated Values* (CSV) sehingga dapat dibuka di *Microsoft Excel*.

1. Autentikasi API Twitter

Autentikasi API Twitter memungkinkan akses terbatas dan aman melalui pengisian data di portal pengembang Twitter. *Access token*, *access token secret*, *consumer key*, dan *consumer secret* API digunakan untuk mengambil data tweet sesuai kebutuhan. Gambar 8 menunjukkan tampilan autentikasi API Twitter.



Gambar 8. Tampilan Autentikasi API Twitter

2. Pengumpulan Data Twitter

Pengumpulan data Twitter dilakukan dengan bahasa pemrograman Python. Penelitian ini menggunakan *library tweepy* dengan melakukan akses pada API Twitter untuk mendapatkan data *tweet* berdasarkan kata kunci "Telkomsel". Contoh data *tweet* yang sudah diambil ditunjukkan pada Tabel 2.

No	Data Tweet
1	my telkomsel ganggu kah
2	kenapa tidak masuk langgan disney lagi
3	app my telkomsel error mau panjang paket gimana ini
4	siap kak putra kami akan cek dm dari kakak tunggu konfirmasi lanjut di dm ya kak makasih gundi
5	memang lagi ganggu apa gimana sih kok tiba tiba sinyal hilang
6	baik kak jika belum kakak sudah pernah kirim data yang minta belum yuk siapa kami di dm
7	aku dapat bulan khusus disney di telkomsel
8	mati lampu mati juga lah telkomsel
9	sih telkomsel mahal banget kouta nya boros pula
10	maaf banget ya kak jeje jadi tidak nyaman kena keluh tidak bisa akses apps mytelkomsel yuk infoin nomor

Tabel 2. Data Tweet

Tabel 2 menunjukkan hasil data *tweet* yang sudah dilakukan proses *preprocessing*. Data *tweet* tersebut menjadi lebih terstruktur agar dapat digunakan untuk melakukan perhitungan ditahap selanjutnya.

3. Pelabelan manual

Pelabelan manual merupakan proses memberikan label positif, negatif dan netral terhadap data *tweet* yang sudah dilakukan proses *preprocessing*. Proses pelabelan manual dilakukan di *Microsoft Excel* secara manual agar data *tweet* dapat dianalisis. Data *tweet* yang sudah diberi label didapatkan 600 data *tweet* yang digunakan untuk *training* dengan data berlabel positif 200 data *tweet*, berlabel netral 200 data *tweet* dan berlabel negatif 200 data *tweet*. Contoh data *tweet* yang sudah dilakukan pelabelan manual ditunjukkan pada Tabel 3.

No	Tweet	Kelas	Label
1	ngebut banget asli si paling ngebut sudah pokok	Positif	1
2	ada cashback nampol sampai dari mytelkomsel yang paling bisa bagi bagi promo ntransaksi minimal rp ribu di mytelkomsel	Positif	1
3	hai kak maaf ya untuk keluh akses internet lambat yuk infoin nomor handphone tanggal jadi loka	Netral	0
4	tenang kak gilang kalau yang kakak maksud kendala akses aplikasi mytelkomsel yuk infoin nomor handphone tanggal	Netral	0
5	jelek banget kalau mati lampu sinyal ikut ngilang	Negatif	-1
6	telkomsel memang asa boros banget apa gimana sih pakai bentar doang sudah habis banyak	Negatif	-1
7	hai kak maaf ya jadi tidak nyaman inara akan bantu kendala kakak yuk dm nomor handphone kakak lokasi detail	Netral	0
8	harga naik disney hotstar yang murah banget kalau pakai telkomsel cuma rb bulan	Positif	1
9	telkomsel legal gift dari mytelkomsel proses cepat langgan disney bulan	Positif	1

Tabel 3. Tabel Pelabelan Manual

4. Hasil Evaluasi Klasifikasi

Hasil evaluasi klasifikasi untuk data *testing* dilakukan dengan menggunakan perhitungan *confusion matrix* untuk menghitung nilai akurasi untuk menentukan perbedaan antara data *training* dan *testing*. Data *testing* akan dihitung sebanyak 336 data *tweet* yang telah diberi kelas positif, negatif dan netral masing-masing 112 label data *tweet*. Hasil perhitungan klasifikasi pada data *testing* ditunjukkan pada Tabel 4.

Tabel 4. Hasil Perhitungan Klasifikasi

Accuracy	0,84
F1-Score	0,84
Precision	0,84
Recall	0,83

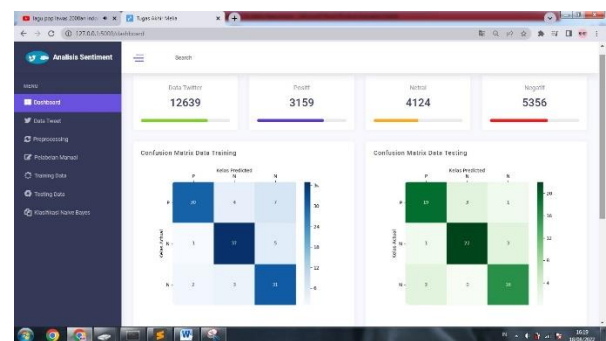
Tabel 4. menunjukkan hasil perhitungan klasifikasi memiliki akurasi yang cukup baik yaitu 0,84, selama pengujian menggunakan *confusion matrix* untuk mengetahui nilai aktual dan prediksi dalam data *testing*.

C. Implementasi Desain Interface

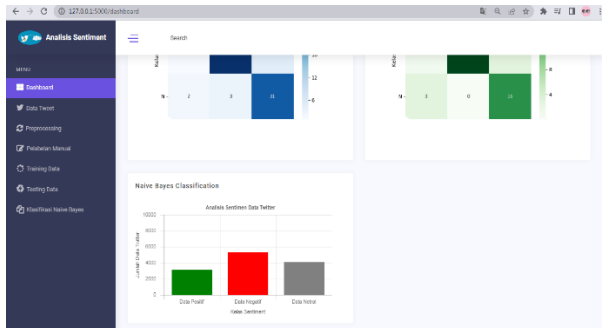
Implementasi desain *interface* merupakan implementasi desain tampilan sistem. Setiap desain yang telah dilakukan sebelumnya di realisasikan menggunakan bahasa pemrograman Python. Dalam Analisis sentimen kepuasan pelanggan perusahaan telekomunikasi seluler Telkomsel di Twitter ini yang digunakan adalah Python dengan memanfaatkan *framework* Flask. Dibawah ini adalah beberapa halaman dan contoh kode program yang termasuk dalam sistem analisis sentimen kepuasan pelanggan perusahaan telekomunikasi seluler Telkomsel di Twitter.

1. Halaman Dashboard Analisis Sentimen

Halaman *dashboard* memuat beberapa informasi berdasarkan hasil dari perhitungan *confusion matrix data training*, *testing* dan hasil data *tweet* yang didapat seperti banyaknya data *Twitter* berdasarkan komentar positif, netral, dan negatif. Implementasi halaman *dashboard* ditunjukkan pada Gambar 9.



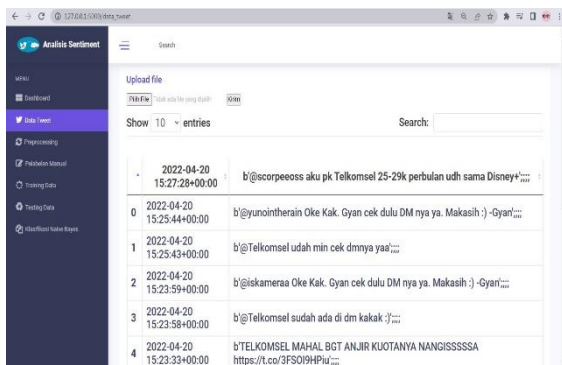
Gambar 9. Implementasi Halaman Dashboard Confusion Matrix



Gambar 10. Implementasi Halaman Dashboard Grafik

2. Halaman Data Twitter

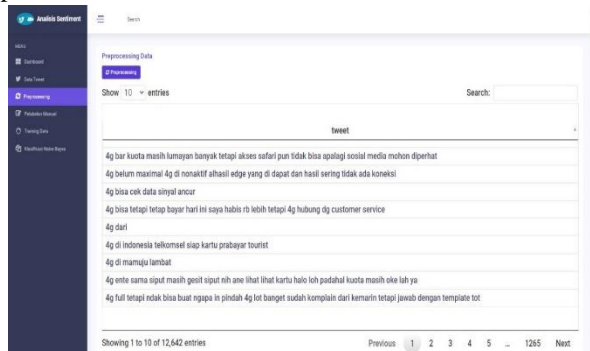
Halaman data *Twitter* digunakan untuk melakukan upload file dan menampilkan file CSV yang sudah terupload ke dalam directory agar memudahkan proses *preprocessing*, *training*, *testing* dan *klasifikasi*. Implementasi halaman data *Twitter* ditunjukkan pada Gambar 11.



Gambar 11. Implementasi Halaman Data Tweet

3. Halaman Preprocessing Data

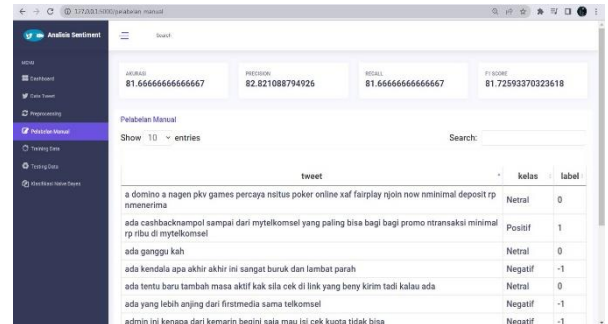
Halaman *preprocessing* berfungsi untuk menampilkan dan melakukan proses pengolahan data teks yang sudah tersedia. *Preprocessing* berfungsi untuk memperbaiki data *tweet* dengan melakukan tahapan-tahapan agar menjadi data yang terstruktur. Implementasi halaman *preprocessing* ditunjukkan pada Gambar 12.



Gambar 12. Implementasi Halaman Preprocessing Data

4. Halaman Pelabelan Manual

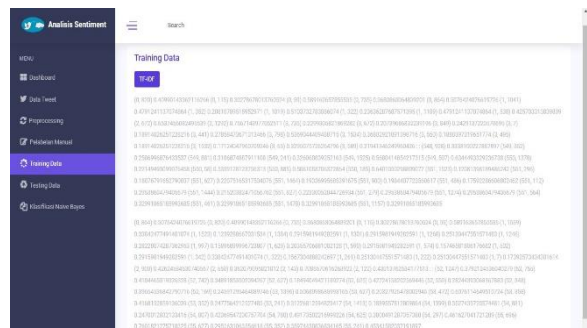
Halaman pelabelan manual digunakan untuk menampilkan hasil *Accuracy*, *Precision*, *Recall* dan *F1-score* berdasarkan data *tweet* yang sudah diberi kelas, label terhadap kalimat yang ada pada data *tweet* secara manual. Proses pelabelan manual dilakukan menggunakan *Microsoft Excel*. Implementasi halaman pelabelan manual ditunjukkan pada Gambar 13.



Gambar 13. Implementasi Halaman Pelabelan Manual

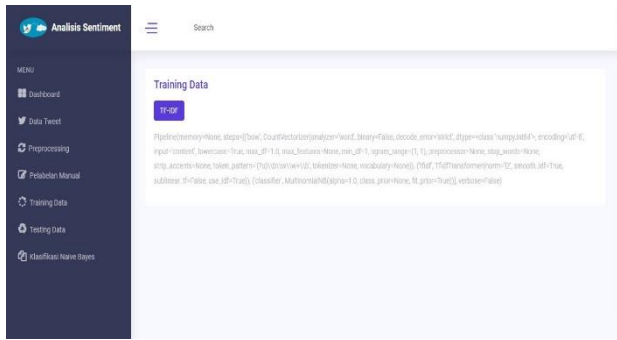
5. Halaman Training Data

Halaman *training* data digunakan untuk membuat model klasifikasi yang digunakan untuk mengklasifikasi sentimen secara otomatis. Langkah-langkah dalam *training* data menggunakan metode *Naive Bayes Classification* dan *library Multinomial Naive Bayes* dengan fitur ekstraksi menggunakan *term frequency- invers document frequency* (TF-IDF) pada data teks. Implementasi halaman *training* data ditunjukkan pada Gambar 14.



Gambar 14. Implementasi Halaman Training Data

Setelah melakukan perhitungan *term frequency- invers document frequency* (TF-IDF) selanjutnya dilakukan pembuatan model klasifikasi yang dibuat dalam sebuah fungsi untuk memudahkan dalam pemanggilan. Tombol *Button* TF-IDF berfungsi untuk melakukan pembuatan model yang akan disimpan dalam bentuk file *.pickle* yang diberi nama variabel *pipeline*. Implementasi halaman pembuatan model klasifikasi dapat dilihat pada Gambar 15.



Gambar 15. Pembuatan Model Klasifikasi

6. Halaman Testing Data

Halaman *testing* digunakan untuk menampilkan hasil *Accuracy*, *Precision*, *Recall* dan *F1-score* berdasarkan data *tweet* yang sudah diprediksi kelas dan labelnya secara otomatis, untuk menentukan seberapa akurat model yang dibangun pada *training* data. Implementasi halaman *testing* data ditunjukkan pada Gambar 16.

AKURASI	PRECISION	RECALL	F1 SCORE
83.82352941176471	83.94117647058823	83.82352941176471	83.8545825570047

tweet	label	kelas
ada di telkomsel goto	1	Positif
ada kendala apa akhir akhir ini sangat buruk dan lambat parah	-1	Negatif
ada yang mau trf wifi jan sampe mati okey	-1	Negatif
admin cek dm dong	1	Positif
admin fio sayang dong sama semua langgan setia telkomsel jadi kalau kakak butuh bantu jangan sungka	-1	Negatif
admin kenapa simcard aku ni di android tidak ke baca dan tidak keluar jaringannya gilir di los ada tetapi niga	-1	Negatif
admin kenapa sudah isi pulsa masa aktif tidak pernah tambah barusan isi pulsa masa aktif sama kara itu mbaw akhir	-1	Negatif

Gambar 16. Implementasi Halaman Testing

7. Halaman klasifikasi Naive Bayes Classification

Halaman klasifikasi Naive Bayes Classification digunakan untuk menampilkan data *tweet* yang sudah diprediksi kelas dan labelnya secara otomatis berdasarkan model yang sudah dibangun pada tahap *training*. Implementasi halaman klasifikasi Naive Bayes Classification ditunjukkan pada Gambar 17.

tweet	label	kelas
4g bar kuota masih lumayan banyak tetapi akses safari pun tidak bisa apalagi sosial media mohon diperhat	-1	Negatif
4g belum maksimal 4g di nonaktif alhasil edge yang di dapat dan hasil sering tidak ada koneksi	-1	Negatif
4g bisa cek data sinyal ancur	-1	Negatif
4g bisa tetapi tetap bayar hari ini saya habis rb lebih tetapi 4g hubung dg customer service	-1	Negatif
4g dari	-1	Negatif
4g di indonesia telkomsel siap kartu Prabayar tourist	1	Positif
4g di mamuju lambat	-1	Negatif

Gambar 17. Implementasi Halaman Klasifikasi Naive Bayes Classification

D. Pembahasan

Analisis sentimen dilakukan menggunakan data *tweet* sebanyak 13.239 dengan fokus pada kata kunci "Telkomsel". Dari data tersebut, 600 *tweet* diberi label

manual dengan akurasi rata-rata 81,7%. Selain itu, terdapat 336 *tweet* yang diprediksi secara otomatis menggunakan model yang dibangun pada tahap *training* dengan akurasi rata-rata 84%. Hasil prediksi pada 12.639 *tweet* menunjukkan bahwa terdapat 3.159 *tweet* berlabel positif, 4.124 *tweet* berlabel netral, dan 5.356 *tweet* berlabel negatif.

Analisis sentimen kepuasan pelanggan perusahaan telekomunikasi seluler di *Twitter* dilakukan menggunakan bahasa pemrograman *Python* dengan *framework Flask*. Sistem ini memiliki beberapa menu, termasuk *dashboard*, data *tweet*, *preprocessing*, pelabelan manual, *training* data, *testing*, dan klasifikasi menggunakan metode *Naive Bayes Classification*. Grafik bar digunakan untuk menampilkan jumlah *tweet* yang diprediksi berdasarkan kelas positif, negatif, dan netral. Sistem ini masih memiliki beberapa kekurangan, seperti ketidakmampuan dalam melakukan pemanggilan file secara otomatis.

IV. KESIMPULAN

Analisis sentimen dilakukan menggunakan data *tweet* sebanyak 13.239 dengan fokus pada kata kunci "Telkomsel". Dari data tersebut, 600 *tweet* diberi label manual dengan akurasi rata-rata 81,7%. Selain itu, terdapat 336 *tweet* yang diprediksi secara otomatis menggunakan model yang dibangun pada tahap *training* dengan akurasi rata-rata 84%. Hasil prediksi pada 12.639 *tweet* menunjukkan bahwa terdapat 3.159 *tweet* berlabel positif, 4.124 *tweet* berlabel netral, dan 5.356 *tweet* berlabel negatif.

Analisis sentimen kepuasan pelanggan perusahaan telekomunikasi seluler di *Twitter* dilakukan menggunakan bahasa pemrograman *Python* dengan *framework Flask*. Sistem ini memiliki beberapa menu, termasuk *dashboard*, data *tweet*, *preprocessing*, pelabelan manual, *training* data, *testing*, dan klasifikasi menggunakan metode *Naive Bayes Classification*. Grafik bar digunakan untuk menampilkan jumlah *tweet* yang diprediksi berdasarkan kelas positif, negatif, dan netral. Sistem ini masih memiliki beberapa kekurangan, seperti ketidakmampuan dalam melakukan pemanggilan file secara otomatis.

DAFTAR PUSTAKA

- [1] P. Agustini, "Warganet Meningkat, Indonesia Perlu Tingkatkan Nilai Budaya di Internet.," <https://aptika.kominfo.go.id/2021/09/warganet-meningkat-indonesia-perlu-tingkatkan-nilai-budaya-di-internet/>, 2021.
- [2] R. Pahlevi, "Pengguna Internet di Dunia Capai 4,95 Miliar Orang Per Januari 2022.," <https://databoks.katadata.co.id/datapublish/2022/02/07/pengguna-internet-di-dunia-capai-495-miliar-orang-per-januari-2022>, Jan. 2022.
- [3] I. Asror, E. Darwiyanto, and A. Nisa, "Analisis Sentimen Menggunakan Naive

- Bayes Classifier dengan Chi-Square Feature Selection Terhadap Penyedia Layanan Telekomunikasi,” vol. 6 No.2, pp. 8650–8659, 2019.
- [4] H. Abubakar, A. M. A. Khaidar, and N. Siskawati, “Pengaruh Kualitas Layanan Terhadap Kepuasan Pelanggan Pada PT Telekomunikasi Selular Indonesia Grapari Telkomsel Makassar (Studi Kasus: Aplikasi My Telkomsel).,” *Economics Bosowa*, vol. 4 No.1, pp. 56–56, 2018.
- [5] R. Kurniawan And A. Apriliani, “Analisis Sentimen Masyarakat Terhadap Virus Corona Berdasarkan Opini Dari Twitter Berbasis Web Scraper,” *Jurnal INSTEK (Informatika Sains dan Teknologi)*, vol. 5, no. 1, p. 67, Apr. 2020, doi: 10.24252/instek.v5i1.13686.
- [6] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, “Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes,” *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 4, no. 2, p. 113, Dec. 2018, doi: 10.26418/jp.v4i2.27526.
- [7] Kevin Perdana, Titania Pricillia, and Z. Zulfachmi, “Optimasi Textblob Menggunakan Support Vector Machine Untuk Analisis Sentimen (Studi Kasus Layanan Telkomsel),” *Jurnal Bangkit Indonesia*, vol. 10, no. 1, pp. 13–15, Mar. 2021, doi: 10.52771/bangkitindonesia.v10i1.120.
- [8] S. Samsir, A. Ambiyar, U. Verawardina, F. Edi, And R. Watrianthos, “Analisis Sentimen Pembelajaran Daring Pada Twitter Di Masa Pandemi Covid-19 Menggunakan Metode Naive Bayes,” *Jurnal Media Informatika Budidarma*, Vol. 5, No. 1, P. 157, Jan. 2021, doi: 10.30865/mib.v5i1.2580.
- [9] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, “Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification,” *SISTEMASI*, vol. 10, no. 1, p. 222, Jan. 2021, doi: 10.32520/stmsi.v10i1.1179.
- [10] R. N. Devita, H. W. Herwanto, and A. P. Wibawa, “Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa indonesia,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 4, p. 427, Oct. 2018, doi: 10.25126/jtiik.201854773.