

**ANALISIS SENTIMEN APLIKASI WHATSAPP BERDASARKAN ULASAN DI PLAY
STORE BERBASIS *NATURAL LANGUAGE PROCESSING***

SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Mendapatkan
Gelar Sarjana Komputer (S.Kom) Program Studi Informatika



SYAMRILLA DEWI

105841102720

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH MAKASSAR

2024

**ANALISIS SENTIMEN APLIKASI WHATSAPP BERDASARKAN ULASAN DI PLAY STORE
BERBASIS *NATURAL LANGUAGE PROCESSING***

SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Mendapatkan
Gelar Sarjana Komputer (S.Kom) Program Studi Informatika



SYAMRILLA DEWI

105841102720

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH MAKASSAR

2024



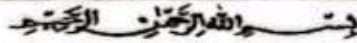
FAKULTAS TEKNIK

GEDUNG MENARA IQRA LT. 3

Jl. Sultan Alauddin No. 259 Telp. (0411) 866 972 Fax (0411) 865 588 Makassar 90221

Website: www.unismuh.ac.id, e_mail: unismuh@gmail.com

Website: <http://teknik.unismuh.makassar.ac.id>



PENGESAHAN

Skripsi atas nama Syamrilla Dewi dengan nomor induk Mahasiswa 105 84 11027 20, dinyatakan diterima dan disahkan oleh Panitia Ujian Tugas Akhir/Skripsi sesuai dengan Surat Keputusan Dekan Fakultas Teknik Universitas Muhammadiyah Makassar Nomor : 0008/SK-Y/55202/091004/2024, sebagai salah satu syarat guna memperoleh gelar Sarjana Komputer pada Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar pada hari Sabtu tanggal 26 Agustus 2024

Panitia Ujian :

Makassar, 26 Muharram 1446 H
26 Agustus 2024 M

1. Pengawas Umum

a. Rektor Universitas Muhammadiyah Makassar

Dr. Ir. H. Abd. Rakhim Nanda, ST., MT., IPU

b. Dekan Fakultas Teknik Universitas Muhammadiyah Makassar

Prof. Dr. Eng. Muhammad Isran Ramli, ST., MT.

2. Penguji

a. Ketua : Dr. Ir. Zahr Zaihuudin, M.Sc

b. Sekretaris : Dr. Ir. Ridwang, S.Kom., MT

3. Anggota

1. Muhyiddin A.M. Hayati, S.Kom., MT

2. Fahrin Irfhamna Rahman S.Kom., MT

3. Lukman Anas, S.Kom., MT

Mengetahui :

Pembimbing I

Pembimbing II


Rizki Yusliana Bakti ST., MT


Titin Wahyuni S.Pd., M.T

Dekan


Dr. Ir. H. Nuharway, ST., MT., IPM.
NBM 795/108



FAKULTAS TEKNIK

GEDUNG MENARA IQRA LT. 3

Jl. Sultan Alauddin No. 259 Telp. (0411) 866 972 Fax (0411) 865 588 Makassar 90221

Website: www.unismuh.ac.id, e_mail: unismuh@gmail.com

Website: <http://teknik.unismuh.makassar.ac.id>

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan untuk memenuhi syarat ujian guna memperoleh gelar Sarjana Komputer (S.Kom) Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar.

Judul Skripsi : **ANALISIS SENTIMEN APLIKASI WHATSAPP BERDASARKAN ULASAN DI PLAYSTORE BERBASIS NATURAL LANGUAGE PROCESSING**


Nama : Syamrilla Dewi
Stambuk : 105 84 11027 20


Makassar, 26 Agustus 2024

Telah Diperiksa dan Disetujui
Oleh Dosen Pembimbing;


Pembimbing I

Pembimbing II


Rizki Yusliana Bakti ST., MT


Titin Wahyuni S.Pd., M.T

Mengetahui,
Ketua Program Studi Informatika


Muhyiddin A M Nawayat, S.Kom., MT.
NBM : 1504 577

ABSTRAK

Perkembangan teknologi yang cepat, terutama di bidang internet, telah mengubah cara komunikasi individu dan kelompok dalam berbagai aspek kehidupan. Internet telah melahirkan berbagai platform digital, termasuk media sosial seperti WhatsApp. WhatsApp adalah aplikasi pesan instan yang memungkinkan penggunaanya mengirim teks, pesan suara, gambar, video, dokumen, serta melakukan panggilan suara dan video melalui koneksi internet. Pemrosesan Bahasa Alami (*Natural Language Processing*) telah mengalami perkembangan pesat dalam beberapa dekade terakhir, terutama terkait dengan analisis sentimen. Metode naive bayes dipilih agar dapat mengklasifikasikan ulasan yang bersentimen positif dan negatif serta netral agar memudahkan masyarakat dalam menentukan aplikasi terbaik berdasarkan nilai akurasi yang telah diteliti. Berdasarkan hasil yang diperoleh dengan menggunakan naive bayes dan perbandingan 80:20 untuk data training dan data testing, menghasilkan akurasi yang cukup tinggi, yaitu sebesar 85,51%, dalam klasifikasi sentimen ulasan. Hal ini menunjukkan bahwa pendekatan NLP yang diterapkan dalam penelitian ini telah berhasil meningkatkan kinerja Naive Bayes Classifier dalam analisis sentimen ulasan pengguna di platform Google Play Store.

Kata Kunci : Analisis Sentimen, WhatsApp, Naive Bayes, NLP

ABSTRACT

Rapid technological developments, especially in the internet sector, have changed the way individuals and groups communicate in various aspects of life. The internet has given birth to various digital platforms, including social media such as WhatsApp. WhatsApp is an instant messaging application that allows users to send texts, voice messages, images, videos, documents, as well as make voice and video calls via an internet connection. Natural Language Processing (Natural Language Processing) has experienced rapid development in the last few decades, especially related to sentiment analysis. The Naive Bayes method was chosen to be able to classify reviews with positive, negative and neutral sentiments to make it easier for the public to determine the best application based on the accuracy values that have been researched. Based on the results obtained using naive Bayes and a ratio of 80:20 for training data and testing data, it produces quite high accuracy, namely 85.51%, in reviewing sentiment classification. This shows that the NLP approach applied in this research has succeeded in improving the performance of the Naive Bayes Classifier in sentiment analysis of user reviews on the Google Play Store platform.

Keywords: *Sentiment Analysis, WhatsApp, Naive Bayes, NLP*

KATA PENGANTAR

Puji Syukur Alhamdulillah penulis ucapkan kepada Allah swt, atas rahmat dan ridho-nya penulis dapat menyelesaikan skripsi yang berjudul “Analisis *Sentimen Aplikasi WhatsApp Berdasarkan Ulasan Di Play Store Berbasis Natural Language Processing*”. Penulisan skripsi ini diajukan untuk memenuhi syarat guna memperoleh gelar Sarjana Komputer (S.Kom) dari Jurusan Informatika Fakultas Teknik Universitas Muhammadiyah Makassar. Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna, oleh karena itu penulis sangat berharap kritik dan saran dari pembaca yang bersifat membangun, sehingga peneliti dapat melakukan perbaikan untuk karya selanjutnya.

Penyusunan dan penulisan skripsi ini tidak terlepas dari dukungan, bantuan serta doa dari banyak pihak khususnya kedua orang tua penulis yang tercinta. **Bapak Syamsuddin dan Ibu Ramlah**, Beliau sangat berperan penting dalam menyelesaikan program study penulis, yang selalu bekerja keras untuk memenuhi kebutuhan penulis, merawat, mendidik serta senantiasa mendoakan penulis dengan sepenuh hati. Semoga panjang umur dan selalu di lindungi oleh Allah swt, kapan pun dan dimana pun berada. Skripsi ini penulis persembahkan untuk kalian.

Penulis juga mengucapkan terima kasih yang sebesar-besarnya kepada seluruh pihak yang telah membantu dengan sepenuh hati, baik berupa ide, kritik, saran dan dukungan semangat, doa, ataupun material sehingga skripsi ini dapat diselesaikan. Oleh karena itu, pada kesempatan ini penulis juga ingin mengucapkan terima kasih kepada :

1. Bapak Dr. Ir. H. Abd. Rakhim Nanda. S.T., M.T., IPU, Sebagai Rektor Perguruan Tinggi Universitas Muhammadiyah Makassar.
2. Ibu Dr. Ir. Hj. Nurnawaty, ST.MT., IPM. Sebagai Dekan Fakultas Teknik Universitas Muhammadiyah Makassar.
3. Bapak Muhyiddin A M Hayat, S.Kom, M.T. Selaku Ketua Program Studi Informatika Universitas Muhammadiyah Makassar

4. Ibu Rizki Yusliana Bakti ST., MT. Selaku Dosen Pembimbing I dan ibu Titin Wahyuni S.Pd.,MT Selaku Dosen Pembimbing II yang senantiasa meluangkan waktunya membimbing dan mengarahkan penulis sejak awal sampai terselesaikannya penulisan skripsi ini.
5. Seluruh Dosen Fakultas Teknik Universitas Muhammadiyah Makassar yang telah memberikan bekal dan ilmu pengetahuan kepada peneliti selama mengikuti pendidikan.
6. Seluruh Staff Administrasi Dan Akademik Lingkungan Fakultas Teknik Universitas Muhammadiyah Makassar yang telah memberikan pelayanan yang baik kepada peneliti selama ini.
7. Semua pihak yang telah banyak membantu penulisan skripsi ini. Peneliti mengucapkan terima kasih yang tak terhingga. Peneliti berharap semoga skripsi ini bisa bermanfaat dan memberikan tambahan pengetahuan dan wawasan yang semakin luas kepada pembaca.

“Billahi fii Sabilil Haq Fastabiqul Khaerat”

Makassar , 20 Maret 2024

Syamrilla Dewi

DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERSETUJUAN.....	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	ix
DAFTAR TABEL.....	x
DAFTAR LAMPIRAN	xi
DAFTAR ISTILAH	vii
BAB I PENDAHULUAN	1
A. Latar Belakang	1
B. Rumusan Masalah.....	2
C. Tujuan Penelitian	3
D. Manfaat Penelitian	3
E. Ruang Lingkup Penelitian	4
F. Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
A. Landasan Teori	5
1. Analisis Sentimen	5
2. <i>Natural Language Processing (NLP)</i>	5
3. WhatsApp	5
4. <i>Naive Bayes Classifier</i>	6
5. Faktor-faktor	6
B. Penelitian Terkait.....	7
C. Kerangka Pikir	9
BAB III METODE PENELITIAN	10
A. Tempat dan Waktu Penelitian.....	10

B.	Alat dan Bahan	10
C.	Perancangan Sistem	10
D.	Teknik Pengujian Sistem	16
E.	Teknik Analisis Data	17
BAB IV HASIL DAN PEMBAHASAN.....		19
A.	Pengambilan Data	19
B.	Pelabelan Data	20
C.	Tahap Preprocessing	32
D.	Spiliting Data	36
E.	Pembobotan TF-IDF	36
F.	Klasifikasi	37
G.	Hasil Pengujian	38
H.	Confusion Matrix	42
BAB V PENUTUP.....		48
A.	Kesimpulan	48
B.	Saran	49
DAFTAR PUSTAKA		50
LAMPIRAN		53

DAFTAR GAMBAR

Gambar 1. Kerangka Fikir Penelitian	9
Gambar 2. <i>Flowchart System</i>	11
Gambar 3. Pelabelan Kelas	13
Gambar 4. Tahap Proses Processing	14
Gambar 5. Hasil Jumlah Pengumpulan Data.....	19
Gambar 6. Hasil Data Sebelum Di Filter Dan Diubah Menjadi <i>Data Frame</i>	19
Gambar 7. Hasil data sesudah di filter dan diubah menjadi <i>Data Frame</i>	20
Gambar 8. Proses Pelabelan Data.....	20
Gambar 9. Proses Pelabelan Data 2.....	21
Gambar 10. Data Ulasan WhatsApp Setelah Dilakukan Cleaning.....	32
Gambar 11. Hasil Data Ulasan WhatsApp Sebelum Proses Case Folding.....	33
Gambar 12. Hasil Data Ulasan WhatsApp Setelah Proses Case Folding.....	33
Gambar 13. Hasil Tahap Pembobotan TF-IDF	37
Gambar 14. Hasil Proses Tahap Klasifikasi Naïve Bayes.....	38
Gambar 15. Hasil Proses Tahap Klasifikasi Naïve Bayes.....	39
Gambar 16. Confusion Matrix Data ulasan WhatsApp.....	41

DAFTAR TABEL

Tabel 1. Hasil Tahap Pelabelan Data.....	22
Tabel 2. Hasil Data Positif.....	24
Tabel 3. Hasil Tabel Data Negatif.....	26
Tabel 4. Hasil Data Netral.....	28
Tabel 5. Tabel Data Ulasan Yang Tidak Sesuai Score.....	29
Tabel 6. Hasil Data Ulasan WhatsApp Setelah Tahap Stopword Removal.....	34
Tabel 7. Hasil Data Ulasan WhatsApp Setelah Tahap Tokenizing.....	34
Tabel 8. Hasil Data Ulasan WhatsApp Setelah Tahap Stemming.....	35
Tabel 9. Jumlah Data.....	36
Tabel 10. Hasil Tahap Akhir Pelabelan Klasifikasi Data True Positif.....	41
Tabel 11. Hasil Tahap Akhir Pelabelan Klasifikasi Data True Negatif.....	43
Tabel 12. Hasil Tahap Akhir Pelabelan Klasifikasi Data True Netral.....	45



DAFTAR LAMPIRAN

Lampiran 1. Proses Pelabelan Data 1	54
Lampiran 2. Hasil data sesudah di filter dan diubah menjadi Data Frame.....	54
Lampiran 3. Hasil Data Sebelum Di Filter Dan Diubah Menjadi Data Frame.....	54
Lampiran 4. Proses Pelabelan Data 2	54
Lampiran 5. Data Ulasan WhatsApp Setelah Dilakukan Cleaning	55
Lampiran 6. Hasil Data Ulasan WhatsApp Sebelum Proses Case Folding	55
Lampiran 7. Hasil Data Ulasan WhatsApp Setelah Proses Case Folding	55
Lampiran 8. Hasil Tahap Pembobotan TF-IDF.....	55
Lampiran 9. Hasil Proses Tahap Klasifikasi Naïve Bayes	56
Lampiran 10. Confusion Matrix Data ulasan WhatsApp	56



DAFTAR ISTILAH

<i>Sentimen</i>	adalah proses memahami dan mengklasifikasikan emosi yang diekspresikan dalam teks, seperti positif, negatif, atau netral.
<i>Whatsapp</i>	Aplikasi pesan instan yang populer, sering digunakan untuk komunikasi pribadi dan bisnis.
<i>Playstore</i>	Toko aplikasi resmi untuk perangkat Android yang dikelola oleh Google.
<i>Natural language processing (nlp)</i>	Cabang dari kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia, termasuk pemahaman, interpretasi, dan pembuatan bahasa alami.
<i>Naive Bayes Classifier</i>	Algoritma klasifikasi berbasis teorema Bayes yang sering digunakan dalam pemrosesan bahasa alami dan pengklasifikasian teks.
<i>Enkripsi end-to-end</i>	Metode enkripsi di mana hanya pengirim dan penerima yang dapat membaca pesan. Pesan dienkripsi di perangkat pengirim dan hanya dapat didekripsi di perangkat penerima.

<i>Yahoo</i>	Perusahaan teknologi yang dikenal dengan layanan email, mesin pencari, dan berita online.
<i>Facebook</i>	Platform media sosial terbesar yang memungkinkan pengguna untuk berbagi konten, berinteraksi dengan teman, dan berpartisipasi dalam komunitas online.
<i>Probabilitas posterior</i>	Probabilitas suatu hipotesis setelah mempertimbangkan bukti baru.
<i>Probabilitas prior</i>	Probabilitas awal suatu hipotesis sebelum mempertimbangkan bukti baru.
<i>Logistic Regression</i>	Model statistik yang digunakan untuk prediksi probabilitas hasil biner, seperti true/false atau yes/no.
<i>Preprocessing</i>	Langkah-langkah yang dilakukan untuk mempersiapkan data mentah sebelum dianalisis atau digunakan dalam model pembelajaran mesin.
<i>Recall</i>	Ukuran dalam evaluasi kinerja model

yang menunjukkan seberapa baik model mendeteksi semua contoh positif.

Flowchart

Diagram yang digunakan untuk mewakili langkah-langkah atau proses dalam bentuk diagram alir.

Web scraping

Teknik untuk mengekstraksi data dari situs web secara otomatis.

Cleaning

Proses membersihkan data dari kesalahan, duplikasi, atau informasi yang tidak relevan.

Case Folding

Mengubah semua huruf dalam teks menjadi huruf kecil untuk memastikan konsistensi dalam pemrosesan teks.

Stopword Removal

Menghapus kata-kata umum yang tidak memiliki makna signifikan dalam analisis teks, seperti "dan", "atau", "di".

Tokenizing

Memecah teks menjadi unit-unit kecil, seperti kata atau kalimat.

<i>Splitting Data</i>	Membagi data menjadi set pelatihan dan set pengujian untuk mengevaluasi kinerja model.
<i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	Teknik untuk mengukur pentingnya suatu kata dalam sebuah dokumen relatif terhadap seluruh korpus.
<i>Machine Learning</i>	Cabang dari kecerdasan buatan yang berfokus pada pengembangan algoritma yang memungkinkan komputer belajar dari dan membuat prediksi berdasarkan data.
<i>Confusion Matrix</i>	Tabel yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan menunjukkan jumlah prediksi benar dan salah untuk setiap kelas.
<i>Lowercase</i>	Mengubah semua huruf dalam teks menjadi huruf kecil untuk konsistensi.
<i>Classification Report</i>	Laporan yang menampilkan metrik evaluasi seperti precision, recall, f1-score, dan akurasi untuk model klasifikasi.

F1-score

Ukuran kinerja model klasifikasi yang menggabungkan precision dan recall menjadi satu metrik.

Precision

Ukuran dalam evaluasi kinerja model yang menunjukkan seberapa baik model mengidentifikasi contoh positif dari semua prediksi positif.



BAB I

PENDAHULUAN

A. Latar Belakang

Perkembangan teknologi yang cepat, terutama di bidang internet, telah mengubah cara komunikasi individu dan kelompok dalam berbagai aspek kehidupan. Internet telah melahirkan berbagai platform digital, termasuk media sosial seperti WhatsApp. WhatsApp adalah aplikasi pesan instan yang memungkinkan penggunanya mengirim teks, pesan suara, gambar, video, dokumen, serta melakukan panggilan suara dan video melalui koneksi internet.

Penelitian ini memanfaatkan data dari ulasan pengguna WhatsApp yang diambil dari platform Google PlayStore. Platform tersebut menyediakan ruang bagi pengguna untuk memberikan ulasan, komentar, dan penilaian terhadap aplikasi WhatsApp yang mereka gunakan. Berbeda dengan penelitian sebelumnya yang hanya fokus pada penilaian skor, penelitian ini lebih menekankan analisis ulasan pengguna serta skornya, karena penilaian skor tidak selalu mencerminkan isi ulasan. Dengan menganalisis ulasan teks, penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam tentang sentimen pengguna terhadap aplikasi WhatsApp.

Untuk memudahkan pengolahan data ulasan, penelitian ini menggunakan analisis sentimen. Analisis sentimen adalah proses menganalisis teks digital untuk menentukan opini yang terkandung di dalamnya, yang bisa berupa positif, negatif, atau netral. Menurut Adhi Putra dalam Maulana (2023), tujuan dari analisis sentimen adalah untuk mengidentifikasi opini positif atau negatif seseorang yang kemudian dapat digunakan dalam pengambilan keputusan. Dengan analisis sentimen ini, peneliti berharap dapat memberikan rekomendasi kepada masyarakat

mengenai apakah aplikasi ini layak digunakan, serta memberikan masukan kepada pengembang aplikasi untuk meningkatkan kualitas layanannya.

Penelitian di bidang Pemrosesan Bahasa Alami (Natural Language Processing) telah mengalami perkembangan pesat dalam beberapa dekade terakhir, terutama terkait dengan analisis sentimen. Analisis sentimen, sebagai salah satu aplikasi utama dari Natural Language Processing, menghasilkan hasil yang lebih akurat dan bermakna dalam berbagai konteks.

Dalam penelitian ini, penulis menggunakan algoritma Naive Bayes untuk melakukan analisis sentimen pada aplikasi WhatsApp. Algoritma *Naive Bayes* dipilih karena merupakan metode klasifikasi dalam penambangan teks yang banyak digunakan dalam analisis sentimen. Selain itu, *Naive Bayes* dikenal sebagai metode yang sederhana namun efektif, serta mampu menghasilkan akurasi tinggi dalam analisis sentimen untuk mengklasifikasikan ulasan pengguna ke dalam kategori opini positif, negatif, atau netral (Indrayuni, 2020).

Latar belakang inilah yang mendorong peneliti untuk meneliti ulasan-ulasan di Google Play Store dari pengguna aplikasi WhatsApp agar pengguna baru dapat mempertimbangkan penggunaan aplikasi ini. Penelitian ini bertujuan untuk melakukan analisis sentimen pada aplikasi WhatsApp berdasarkan ulasan di Play Store dengan pendekatan *Natural Language Processing* menggunakan metode *Naive Bayes Classifier*. Selain itu, penelitian ini juga berupaya untuk mendapatkan pemahaman yang lebih baik tentang kinerja metode ini berdasarkan penelitian sebelumnya.

B. Rumusan Masalah

Berdasarkan latar belakang diatas, maka rumusan masalah yang dapat disampaikan antara lain:

1. Bagaimana cara menganalisis *sentimen* ulasan pengguna *WhatsApp* di *Play Store* menggunakan metode *Naive bayes Classifier*?

2. Apa saja faktor-faktor yang memengaruhi *sentimen* pengguna dalam ulasan mereka terhadap aplikasi *WhatsApp* di *Play Store*?
3. Bagaimana klasifikasi *sentimen* ulasan pengguna *WhatsApp* ke dalam kategori positif, negatif, atau netral dapat dilakukan secara efektif menggunakan pendekatan *Natural Language Processing*?

C. Tujuan Penelitian

1. Menganalisis sentimen ulasan pengguna *WhatsApp* di *Google Play Store* menggunakan metode *Naive Bayes Classifier*.
2. Mengidentifikasi dan menganalisis faktor-faktor yang memengaruhi sentimen pengguna dalam ulasan mereka terhadap aplikasi *WhatsApp* di *Google Play Store*.
3. Mengklasifikasikan sentimen ulasan pengguna *WhatsApp* ke dalam kategori positif, negatif, atau netral secara efektif dengan menggunakan pendekatan *Natural Language Processing (NLP)* yang diintegrasikan dengan metode *Naive Bayes Classifier*.

D. Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Bagi penulis, penelitian ini merupakan sebuah eksplorasi teori-teori yang selama ini dipelajari, serta menambah wawasan, ilmu pengetahuan, dan pengalaman terhadap analisis *sentimen*
2. Bagi Universitas, sebagai tolak ukur pengetahuan mahasiswa dalam menguasai ilmu sudah dipelajari dan sebagai referensi untuk penelitian selanjutnya.
3. Bagi pembaca, memberikan informasi mengenai *sentimen* terhadap kepuasan pengguna aplikasi *WhatsApp* dan bermanfaat untuk referensi penelitian analisis *sentimen*

E. Ruang Lingkup Penelitian

Penelitian ini dibuat ruang lingkup penelitian agar penelitian tetap pada jalur yang telah ditentukan di antaranya yaitu:

1. Data yang diambil berasal dari ulasan para pengguna aplikasi *WhatsApp* yang ada di *google Play Store*.
2. Klasifikasi kelas *sentimen* yang dipakai hanya ulasan positif dan ulasan negatif dan netral.

F. Sistematika Penulisan

Struktur penulisan penelitian ini terbagi menjadi beberapa bab yang diuraikan sebagai berikut:

BAB I : PENDAHULUAN, Bab ini mencakup latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan penelitian ini..

BAB II : TINJAUAN PUSTAKA, Bab ini menguraikan pembahasan mengenai penelitian-penelitian terkait, landasan teori, dan kerangka pikir yang mendasari penelitian ini..

BAB III : METODE PENELITIAN, Bab ini memberikan penjelasan tentang metode yang digunakan dalam penelitian, serta rincian tentang pelaksanaan penelitian ini.

BAB IV : HASIL DAN PEMBAHASAN, Bab ini memaparkan hasil penelitian, diskusi mengenai hasil yang diperoleh, dan hasil implementasi sistem.

BAB V : PENUTUP, Bab ini menyajikan kesimpulan dari penelitian serta saran-saran berdasarkan hasil penelitian yang telah dilakukan.

BAB II

TINJAUAN PUSTAKA

A. Landasan Teori

Pada bagian ini berisikan teori–teori yang akan digunakan dalam penelitian ini. Peneliti meninjau teori–teori dari buku, jurnal, dan dari penelitian terdahulu guna mendapatkan informasi yang sesuai dengan penelitian ini.

1. Analisis Sentimen

Analisis sentimen adalah proses yang digunakan untuk menganalisis teks digital guna menentukan opini yang terkandung di dalamnya, apakah itu positif, negatif, atau netral. Menurut Adhi Putra dalam Maulana, tujuan dari analisis sentimen adalah mengidentifikasi opini seseorang, baik itu positif maupun negatif, yang kemudian dapat dijadikan dasar pengambilan keputusan (Maulana, 2023). Secara sederhana, analisis sentimen bertujuan untuk menilai apakah suatu ulasan atau opini terhadap produk, layanan, atau aplikasi bersifat positif atau negatif.

2. *Natural Language Processing (NLP)*

Natural Language Processing (NLP) adalah bidang dalam ilmu komputer yang menggabungkan pembelajaran mesin dan linguistik, bertujuan untuk membuat komputer memahami bahasa manusia yang tertulis. NLP bertujuan untuk memudahkan interaksi antara manusia dan komputer menggunakan bahasa alami, mengingat tidak semua pengguna fasih dalam bahasa pemrograman tertentu. NLP memungkinkan pengguna untuk berkomunikasi dengan komputer menggunakan bahasa yang mereka kenal.

3. WhatsApp

WhatsApp adalah aplikasi pesan yang sederhana, aman, dan andal, yang dapat diunduh secara gratis di ponsel di seluruh dunia. Nama "WhatsApp" merupakan permainan kata dari frasa "What's Up."

Aplikasi ini digunakan oleh lebih dari 2 miliar orang di lebih dari 180 negara untuk tetap terhubung dengan teman dan keluarga. WhatsApp memungkinkan pengguna mengirim pesan teks, foto, video, dokumen, dan lokasi, serta melakukan panggilan suara. WhatsApp didirikan oleh Jan Koum dan Brian Acton, yang sebelumnya bekerja di Yahoo, dan bergabung dengan Facebook pada tahun 2014. Meskipun bergabung dengan Facebook, WhatsApp terus beroperasi sebagai aplikasi terpisah dengan fokus utama pada kecepatan dan keandalan pesan.

4. *Naive Bayes Classifier*

Naive Bayes adalah metode klasifikasi yang populer karena kesederhanaannya dan akurasinya yang tinggi. Metode ini didasarkan pada probabilitas sederhana dan mengasumsikan adanya independensi antara variabel penjelas. Keunggulan dari *Naive Bayes* adalah kemampuannya dalam menghasilkan akurasi tinggi, terutama saat diterapkan pada dataset yang besar. Algoritma ini sangat efektif dalam analisis sentimen, terutama dalam mengklasifikasikan ulasan pengguna ke dalam kategori positif, negatif, atau netral.

5. **Faktor-faktor**

Faktor yang memengaruhi sentimen pengguna dalam ulasan mereka terhadap aplikasi WhatsApp meliputi

- kualitas layanan, fitur aplikasi, Pembaruan dan Perubahan, Keamanan dan Privasi, Dukungan Pengguna dan pengalaman pengguna secara keseluruhan. Analisis terhadap ulasan pengguna menunjukkan bahwa sentimen positif umumnya dipengaruhi oleh kepuasan terhadap fitur-fitur yang ditawarkan oleh WhatsApp, sementara sentimen negatif sering kali muncul karena masalah teknis atau kekurangan tertentu dalam aplikasi.

B. Penelitian Terkait

Penelitian ini bertujuan untuk mengklasifikasikan sentimen dalam ulasan aplikasi WhatsApp menggunakan metode *Naive Bayes*. Dalam penelitian ini, opini pengguna dibagi menjadi dua kategori, yaitu positif dan negatif, yang kemudian diklasifikasikan menggunakan algoritma *Naive Bayes Classifier*.

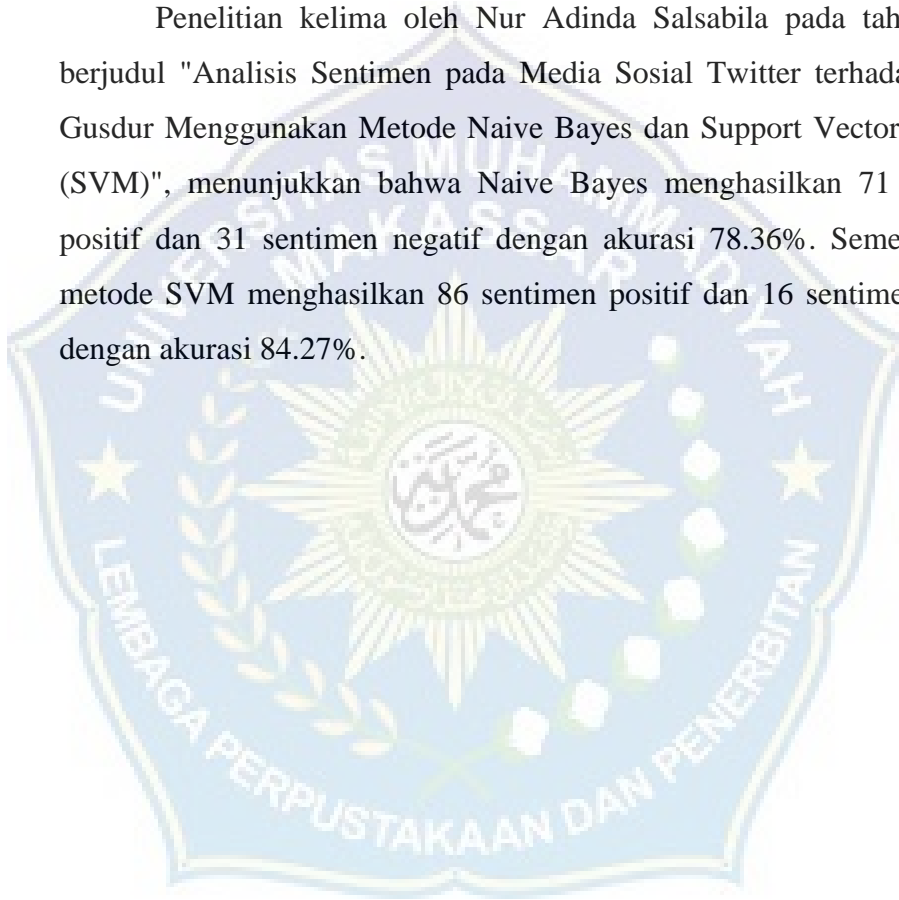
Penelitian pertama yang relevan adalah jurnal yang ditulis oleh Marsha Auriel Prasetya, Miftakhul Wulandari, dan Siti Alvi Nikmah pada tahun 2024, yang berjudul "Implementasi NLP (Natural Language Processing) Dasar pada Analisis Sentimen Review Spotify". Penelitian ini menggunakan algoritma Naive Bayes dan Logistic Regression, dengan hasil terbaik diperoleh melalui Logistic Regression, yang mencapai akurasi sebesar 79%, sedangkan Naive Bayes mencapai akurasi 76%.

Penelitian kedua dilakukan oleh Dhodi Surya Sayogo, Bambang Irawan, dan Agus Bahtiar pada tahun 2023, dengan judul "Analisis Sentimen Ulasan Instagram di Google Play Store Menggunakan Algoritma Naive Bayes". Dari 1100 ulasan, setelah preprocessing data, diperoleh 978 ulasan yang diberi label, terdiri dari 152 ulasan positif dan 826 ulasan negatif. Penelitian ini menunjukkan bahwa ulasan terhadap Instagram cenderung negatif, dengan hasil akurasi Naive Bayes mencapai 89%, precision 90%, dan recall 98%.

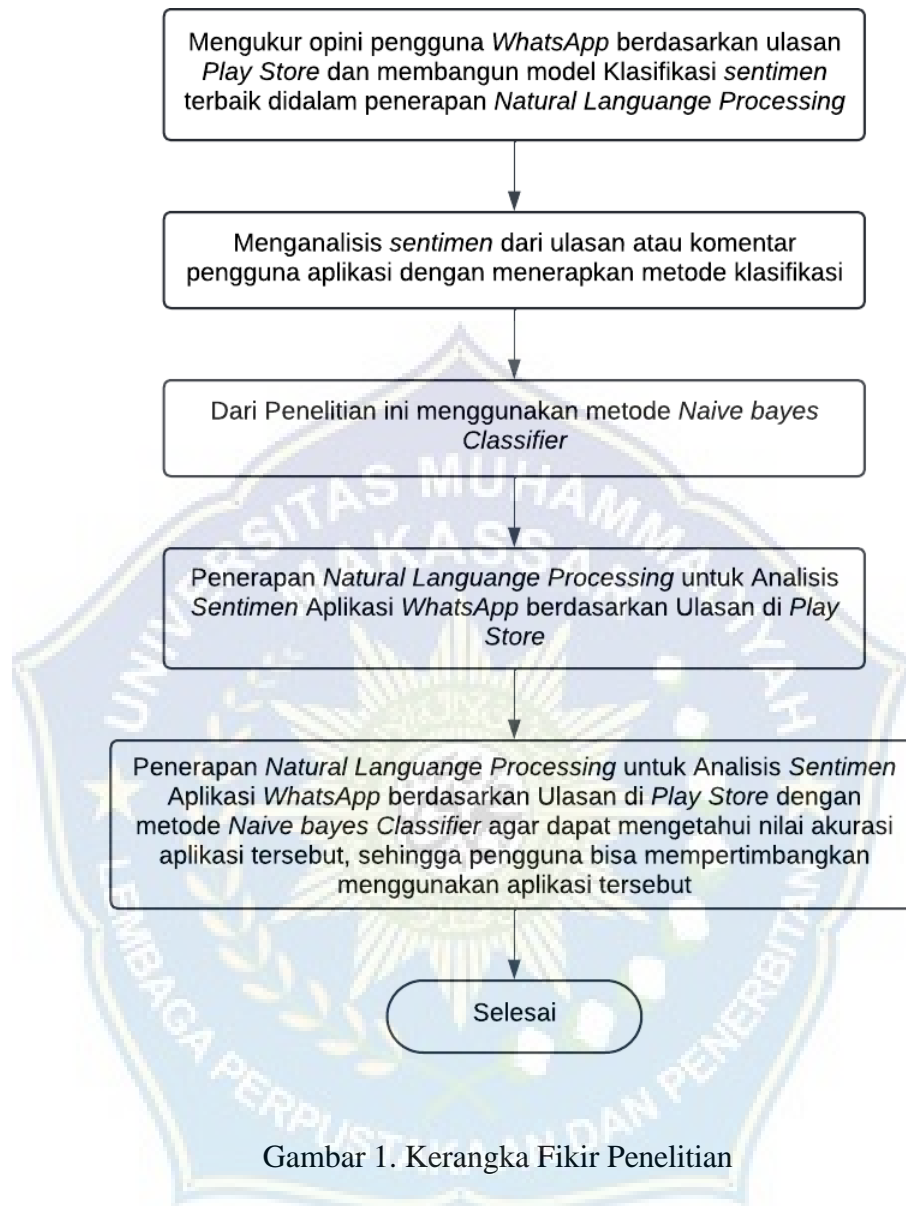
Penelitian ketiga oleh Mhd. Furqan pada tahun 2023, berjudul "Chatbot Telegram Menggunakan Natural Language Processing", menunjukkan bahwa penerapan NLP memungkinkan proses penyaringan data yang cepat dan tepat, dengan hasil UAT menunjukkan bahwa pengguna puas dengan chatbot tersebut, yang dianggap efektif dengan persentase kepuasan sebesar 94%.

Penelitian keempat oleh Angga Aditya Permana pada tahun 2023, berjudul "Implementasi Algoritma *Naive Bayes* terhadap Review Aplikasi KFCKU", menemukan bahwa algoritma ini efektif dalam menganalisis sentimen dengan akurasi 85%, precision 82%, recall 99%, dan F1 Score 90%. Fitur stemming digunakan untuk mereduksi kata berimbuhan menjadi kata dasar, sehingga analisis sentimen menjadi lebih fokus dan relevan.

Penelitian kelima oleh Nur Adinda Salsabila pada tahun 2022, berjudul "Analisis Sentimen pada Media Sosial Twitter terhadap Tokoh Gusdur Menggunakan Metode Naive Bayes dan Support Vector Machine (SVM)", menunjukkan bahwa Naive Bayes menghasilkan 71 sentimen positif dan 31 sentimen negatif dengan akurasi 78.36%. Sementara itu, metode SVM menghasilkan 86 sentimen positif dan 16 sentimen negatif dengan akurasi 84.27%.



C. Kerangka Pikir



Gambar 1. Kerangka Fikir Penelitian

BAB III

METODE PENELITIAN

A. Tempat dan Waktu Penelitian

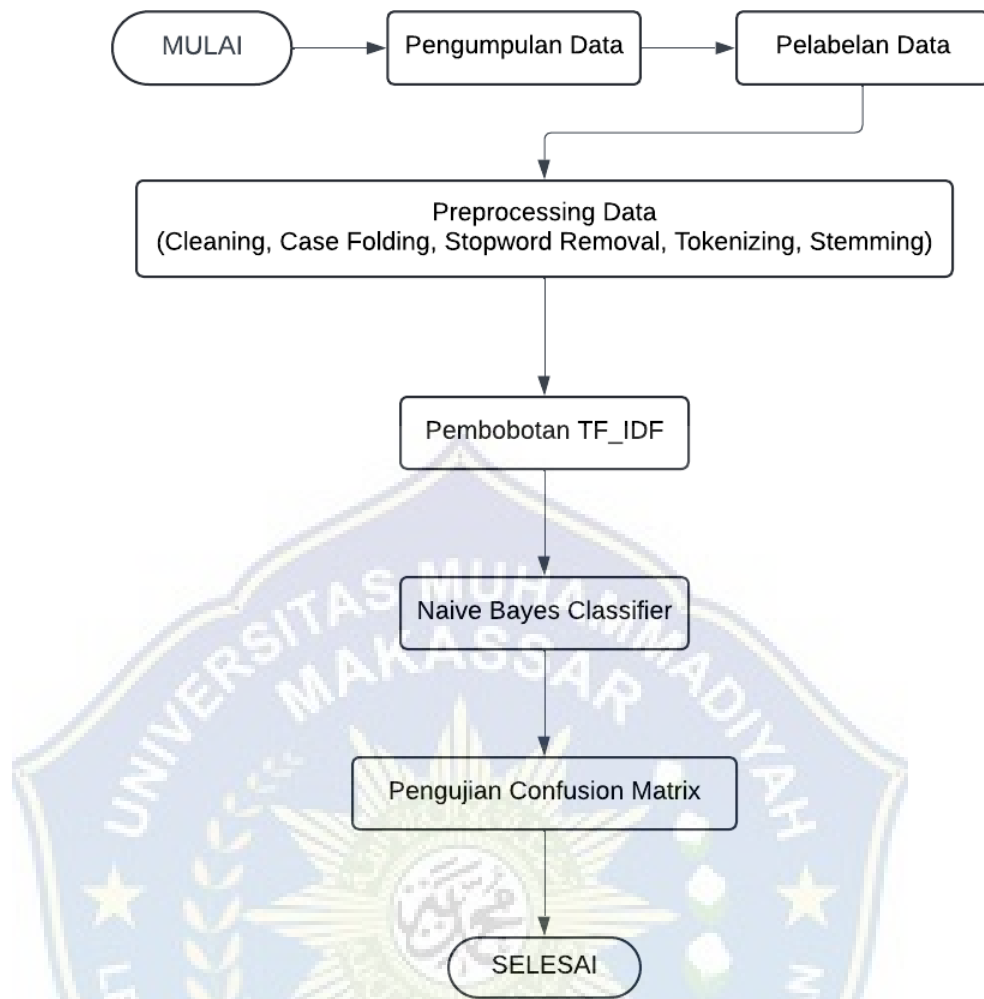
Penelitian ini berfokus pada data yang berasal dari Google Play Store, di mana peneliti mengumpulkan data dari platform tersebut. *WhatsApp* dipilih sebagai objek penelitian karena memiliki jumlah unduhan dan rating bintang yang tinggi. Data yang digunakan dalam penelitian ini berasal dari kolom ulasan aplikasi *WhatsApp* yang tersedia di *Google Play Store*.

B. Alat dan Bahan

Alat yang digunakan dalam penelitian ini meliputi komputer atau laptop yang dilengkapi dengan perangkat keras dan perangkat lunak pendukung. Bahan yang digunakan adalah data ulasan pengguna yang diambil dari *Google Play Store*.

C. Perancangan Sistem

Untuk memudahkan perancangan sistem, peneliti membuat diagram alir (flowchart) sebagai panduan. Proses perancangan sistem meliputi beberapa tahapan:



Gambar 2. *Flowchart System*

Perancangan Sistem yang digunakan dalam penelitian ini adalah :

1. Pengumpulan Data

Pada tahap ini, data ulasan pengguna aplikasi WhatsApp diambil dari Google Play Store menggunakan teknik scraping. Scraping atau web scraping adalah metode untuk mengambil data dari internet, umumnya berupa halaman web, sesuai dengan kebutuhan (Faadilah, 2020). Teknik web scraping ini bisa dilakukan secara manual, dengan cara copy-paste data yang diperlukan, atau secara otomatis dengan

membuat program atau kode yang dapat mengambil data dari halaman web (Baskara & Rahma, 2022).

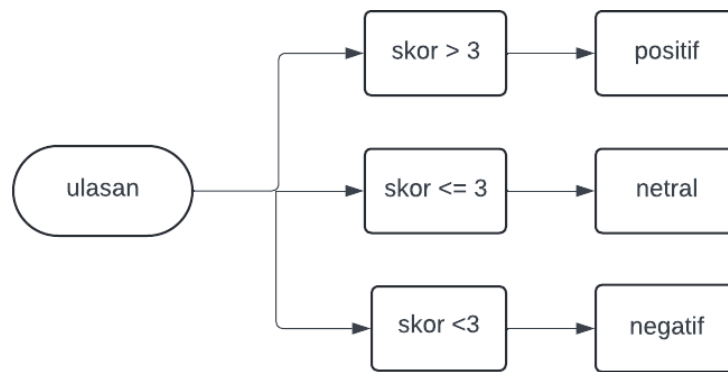
Di Google Play Store, terdapat fitur rating dan ulasan yang mencerminkan kinerja suatu aplikasi. Rating dengan skor 1 bintang menunjukkan kinerja yang sangat buruk, 2 bintang berarti buruk, 3 bintang menunjukkan cukup, 4 bintang menunjukkan baik, dan 5 bintang berarti sangat baik. Dalam penelitian ini, ulasan yang dipilih adalah ulasan yang paling relevan dengan rating berbintang 1, 2, 4, dan 5. Kategori ulasan relevan dipilih karena mencakup ulasan yang paling berkaitan dengan aplikasi, dengan isi ulasan yang jelas mengenai keluhan atau kepuasan pengguna, serta beberapa ulasan yang memberikan saran untuk pengembangan di masa depan.

2. Pelabelan Data

Pada tahap ini, ulasan pengguna diberi label sebagai positif, netral, atau negatif berdasarkan skornya. Ulasan positif adalah komentar yang mendukung dan merekomendasikan aplikasi, sedangkan ulasan negatif adalah komentar yang berisi kritik atau saran terhadap aplikasi ini (Kosasih, 2021). Ulasan dengan skor kurang dari 3 diberi label negatif, skor 4 dan 5 diberi label positif, dan skor 3 dianggap sebagai ulasan netral. Proses pelabelan ini bertujuan untuk mengklasifikasikan ulasan secara otomatis sehingga dapat digunakan dalam proses training dan testing.

Pelabelan manual dilakukan untuk menentukan label sentimen setiap ulasan sebagai positif, negatif, atau netral.

- Preprocessing dilakukan untuk membersihkan data.
- Pembobotan fitur menggunakan TF-IDF.
- Seleksi fitur berdasarkan nilai DF.

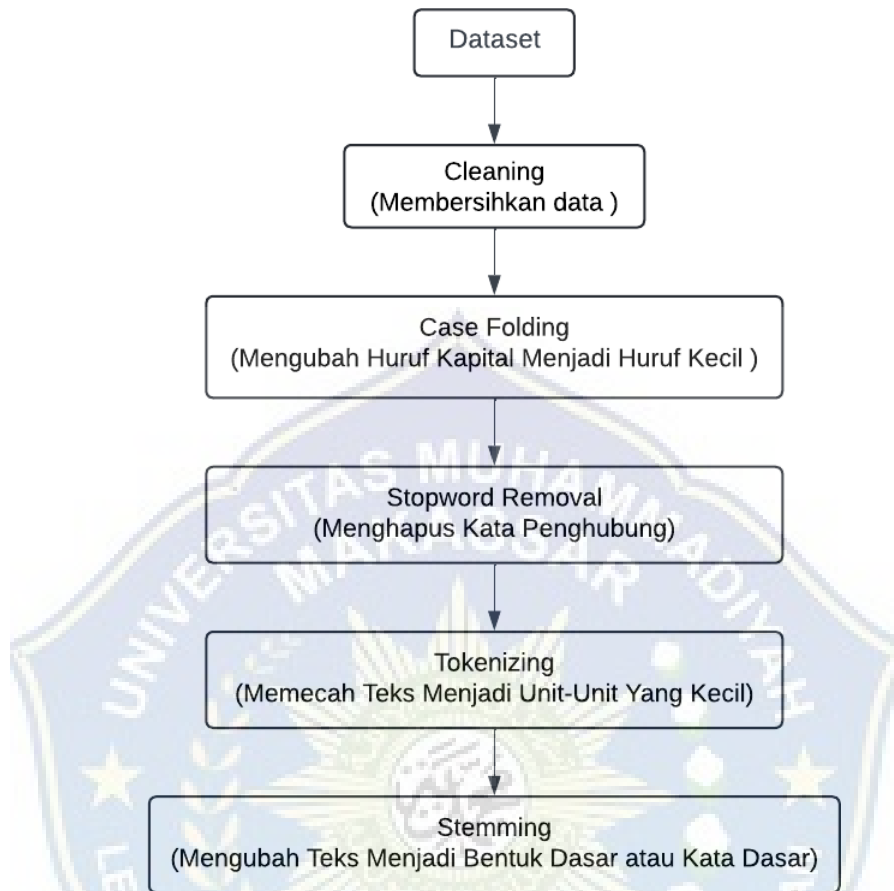


Gambar 3. Pelabelan Kelas

3. Tahap *preprocessing*

Data ulasan yang diperoleh sering kali mengandung kata-kata yang tidak terstruktur, seperti singkatan dan simbol, sehingga perlu dilakukan preprocessing. Tahap preprocessing adalah langkah awal dalam pengolahan data yang bertujuan untuk membersihkan, memformat, dan mempersiapkan data mentah sebelum analisis lebih lanjut. Proses ini membantu mengubah informasi dari ulasan dan menstandarkan kata-kata yang tidak terstruktur (Faadilah, 2020).

Berikut merupakan tahapan-tahapan dari proses *preprocessing*:



Gambar 4. Tahap Proses Processing

4. *Splitting Data*

Splitting data dalam analisis sentimen adalah proses memisahkan data teks menjadi dua bagian: data pelatihan (training data) dan data pengujian (testing data). Tujuannya adalah untuk melatih model menggunakan data pelatihan dan menguji kinerjanya dengan data pengujian guna mengevaluasi seberapa baik model tersebut memprediksi data baru. Setelah preprocessing, data dipilih dan dibagi untuk proses training dan testing algoritma machine learning (Iriananda et al., 2021). Data pelatihan digunakan untuk melatih model dan mengidentifikasi pola sentimen, sementara data pengujian digunakan

untuk mengukur akurasi model dalam memprediksi sentimen pada data baru.

5. Pembobotan TF-IDF

TF-IDF (Term Frequency-Inverse Document Frequency) adalah metode yang digunakan untuk memberikan bobot pada setiap kata dalam dokumen, berdasarkan seberapa sering kata tersebut muncul dan keunikannya dalam keseluruhan dokumen. Term Frequency (TF) mencerminkan seberapa sering kata muncul dalam suatu dokumen, sedangkan Inverse Document Frequency (IDF) mengurangi bobot kata yang sering muncul, untuk menilai pentingnya makna kata tersebut dalam dokumen. Pembobotan TF-IDF ini mengubah teks menjadi vektor yang bisa digunakan dalam proses klasifikasi dengan Naive Bayes dan model machine learning lainnya. TF-IDF memiliki pengaruh signifikan dalam tugas klasifikasi (Iriananda et al., 2021).

6. Pengklasifikasian

Pada tahap ini, metode Naive Bayes Classifier digunakan. Naive Bayes adalah metode klasifikasi yang sangat populer dan efektif dalam machine learning dan data mining. Metode ini didasarkan pada teorema Bayes, yang menyatakan bahwa probabilitas suatu hipotesis (atau kelas) dapat dihitung berdasarkan probabilitas fitur-fitur yang terkait dengannya. Naive Bayes Classifier adalah algoritma yang sederhana dan mudah digunakan, serta mampu memperkirakan kejadian berdasarkan hasil klasifikasi dengan baik (Amaliah & Dwi Nuryana, 2022).

Naive Bayes memodelkan probabilitas kelas sebagai produk dari probabilitas setiap fitur yang terkait dengan kelas tersebut. Dengan menggunakan aturan Bayes, probabilitas posterior untuk setiap kelas dapat dihitung berdasarkan probabilitas prior dan probabilitas kelas yang telah dihitung sebelumnya. Salah satu keunggulan dari algoritma Naive Bayes adalah kebutuhan akan sedikit data pelatihan untuk menentukan estimasi parameter yang diperlukan dalam proses

klasifikasi. Metode ini sering kali bekerja dengan baik dalam situasi kompleks di dunia nyata, lebih baik dari yang diharapkan (Rifai et al., 2019).

D. Teknik Pengujian Sistem

Metode pengujian sistem yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Pengujian *Confusion Matrix*

Confusion Matrix adalah tabel yang digunakan untuk mengevaluasi kinerja model klasifikasi, termasuk dalam konteks analisis sentimen. Metode ini digunakan untuk menghitung akurasi dengan cara membandingkan hasil klasifikasi yang dilakukan oleh model dengan hasil klasifikasi aktual (Amaliah & Dwi Nuryana, 2022). *Confusion Matrix* membantu menilai sejauh mana model dapat melakukan klasifikasi dengan benar, dengan membandingkan hasil prediksi model terhadap label kelas yang sebenarnya. Menurut Ilmawan & Mude dalam Maulana, informasi yang terdapat dalam *Confusion Matrix* sangat penting untuk mengevaluasi kinerja model klasifikasi (Maulana, 2023).

Dalam analisis *sentimen*, *Confusion Matrix* digunakan untuk menghitung akurasi, presisi, *recall*, dan *F1-score* dari model.

- **Akurasi** adalah proporsi dari jumlah data yang diklasifikasikan dengan benar terhadap keseluruhan data. Rumus untuk menghitung akurasi adalah:

$$= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \dots\dots\dots (1)$$

- **Presisi** mengukur persentase keakuratan hasil prediksi yang dibuat oleh model. Rumus untuk menghitung presisi adalah:

$$= \frac{\text{TP}}{\text{TP} + \text{FP}} \dots\dots\dots (2)$$

- **Recall** adalah proporsi dari data positif yang berhasil diklasifikasikan dengan benar dari total data positif yang ada, yang mencerminkan kemampuan model dalam menemukan kembali informasi yang relevan (Anggreany, 2020). Rumus untuk menghitung recall adalah:

$$= \frac{\text{TP}}{\text{TP} + \text{FP}} \dots\dots\dots (3)$$

- **F1-score** adalah gabungan antara presisi dan recall yang memberikan nilai yang lebih baik dalam mengukur kinerja model secara keseluruhan (Amaliah & Dwi Nuryana, 2022). Rumus untuk menghitung F1-score adalah:

$$F1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

E. Teknik Analisis Data

Analisis data adalah proses sistematis untuk mencari dan mengatur data hasil ulasan serta bahan-bahan yang telah dikumpulkan, dengan tujuan memperdalam pemahaman terhadap semua informasi yang terkumpul dan memungkinkan untuk menyajikan temuan dengan lebih jelas. Tahapan analisis data yang digunakan dalam penelitian ini meliputi:

1. Pengumpulan Data

Data yang dikumpulkan dapat berupa teks, dokumen, atau ulasan yang mengandung opini atau sentimen. Sumber data bisa berasal dari berbagai media sosial, platform digital, atau melalui wawancara. Pengumpulan data merupakan langkah yang sangat penting dalam penelitian, bahkan lebih penting daripada menyusun instrumen penelitian itu sendiri (Alhamid & Anufia, 2019).

2. Klasifikasi *Sentimen*

Dalam proses analisis sentimen, teks atau dokumen yang dikumpulkan dikategorikan ke dalam sentimen positif atau negatif.

Proses ini dapat dilakukan dengan menggunakan metode pengolahan bahasa alami (Natural Language Processing) atau dengan bantuan perangkat lunak analisis sentimen.

3. Penghitungan Frekuensi

Setelah teks atau dokumen diklasifikasikan ke dalam kategori sentimen, langkah berikutnya adalah menghitung frekuensi masing-masing kategori. Frekuensi ini dapat dihitung secara absolut (jumlah sentimen positif dan negatif) maupun relatif (persentase sentimen positif dan negatif dari total dokumen yang ada).

4. Interpretasi Hasil

Hasil analisis data kemudian diinterpretasikan untuk memberikan pemahaman tentang sentimen atau opini yang diungkapkan dalam dokumen atau teks. Interpretasi ini dapat memberikan wawasan tentang pandangan masyarakat terhadap suatu topik atau produk tertentu (Faadilah, 2020).

BAB IV HASIL DAN PEMBAHASAN

A. Pengambilan Data

Dengan menggunakan *package google-play-scraper*, data ulasan *WhatsApp* diambil dari *Google Play Store*. Ulasan disusun dalam bahasa Indonesia dan diurutkan menurut relevansi. Jumlah ulasan yang dikumpulkan adalah 2.750. Skrip yang digunakan untuk mengumpulkan data ulasan adalah sebagai berikut:

```
print("Mengambil ulasan dari Google Play Store...") ulasan = reviews_all(
    'com.whatsapp', lang='id',
    country='id',
    sort=Sort.MOST_RELEVANT, count=2750
)
print("Ulasan berhasil diambil.")
```

Gambar 5. Hasil Jumlah Pengumpulan Data

id	nama	jenis	deskripsi	rating	jumlah_ulasan	tanggal_terakhir	tanggal
448701448	WhatsApp	Persepsi	Persepsi	5	231107	2020-02-28	2020-02-28
448701448	WhatsApp	Persepsi	Persepsi	5	231107	2020-02-28	2020-02-28
448701448	WhatsApp	Persepsi	Persepsi	5	231107	2020-02-28	2020-02-28
448701448	WhatsApp	Persepsi	Persepsi	5	231107	2020-02-28	2020-02-28
448701448	WhatsApp	Persepsi	Persepsi	5	231107	2020-02-28	2020-02-28
448701448	WhatsApp	Persepsi	Persepsi	5	231107	2020-02-28	2020-02-28
448701448	WhatsApp	Persepsi	Persepsi	5	231107	2020-02-28	2020-02-28
448701448	WhatsApp	Persepsi	Persepsi	5	231107	2020-02-28	2020-02-28
448701448	WhatsApp	Persepsi	Persepsi	5	231107	2020-02-28	2020-02-28
448701448	WhatsApp	Persepsi	Persepsi	5	231107	2020-02-28	2020-02-28

Gambar 6. Hasil Data Sebelum Di Filter Dan Diubah Menjadi *Data Frame*

username	content	score
Pengguna Google	Inglai kan flu pada status seperti memapung durasi, kualitas video, pemerkasan video pada status, dan audio dari video yang di unggah ke status itu, tetapi video lebih dari 1 menit kurang lebih itu saja yang ingin diperhatikan terutama durasi dan kualitas video karena di pedati percuma banyak durasi atau kualitas	1
Pengguna Google	Aku upload video antara suara dan video tidak sinkron, ada delay. Contoh ketika upload status kemarin gitar atau piano, suara dan video tidak sinkron ada delay	1
Pengguna Google	Manap sekarang bisa update status video dengan durasi sampai 1 menit, saat mungkin bisa diambatkan flu HD saat status video ini bisa ngak pecah atau flu lagi status video yang kita up	5
Pengguna Google	Entah bagaimana caranya WA ini memotong banyak sekali memori internal sampai 4GB? Padahal ini WhatsApp ga internal itu, Ga mengikat memori? Jantung ada yg chat. Kenapa bisa hingga begitu... ML aja ga sampai 4GB di HP ini bagaimana tolong beri penjelasan @GARAN	2
Pengguna Google	Sebenarnya sih bagus tapi entah kenapa di ponselku saya kalo mau masuk ke WhatsApp itu harus masalah nomor dua harus nanti dikirimin kode. Jadi ya ribet banget, jadi tolong buat aplikasi ini semoga bisa berkembang lebih baik lagi	4
Pengguna Google	Kenapa story na sng tangkapnya beda ?? Ga sesuai update waktu nya hanya untuk kontak dan hanya itu saja yg muncul ?	4
Pengguna Google	Tolong admin WhatsApp kelap saat akun saya di blokir padahal saya telah mematuhi persyaratan dan peraturan yang dibuat oleh WhatsApp setiap saat akun saya di blokir lalu dikaba kembali setelah itu di blokir kembali tolong berikan penjelasan yang spesifik mengenai masalah yang saya buat	1
Pengguna Google	Saya sudah 2x daftarin nomor baru saya, tapi kenapa selalu kena spom, dan selalu melakukan peninjauan, dibuat sng ribet malah gabisa masuk akun WA tidak resmi, padahal saya blaster di play store!	1
Pengguna Google	Dari aplikasi ini bisa komunikasi dengan teman samela...sungguh aplikasi yang benar benar sangat bisa membantu...jika memang benar sih banyak aplikasi lain yang bisa membantu tapi aplikasi ini sangat aku apresiasi banget sukses buat WhatsApp semoga makin sukses dan bisa memperbaiki flu flu nya ke yang lebih bagus lagi 🙏	5
Pengguna Google	ApaKasihya bagus, kalau bisa tolong admin flu juga kalau mau tap kali unggah anter! Makasih!	4

Gambar 7. Hasil data sesudah di filter dan diubah menjadi *Data Frame*

B. Pelabelan Data

Setelah data diambil, langkah selanjutnya adalah melabeli ulasan berdasarkan rating yang diberikan oleh pengguna. Rating lebih dari 3 dianggap sebagai *sentimen* positif, rating kurang dari 3 dianggap sebagai *sentimen* negatif, dan rating sama dengan 3 dianggap sebagai *sentimen* netral.



Gambar 8. Proses Pelabelan Data

Setelah penglabelan menggunakan score, penglabelan kedua dilakukan lagi dengan merujuk beberapa kata positif, negatif, serta netral agar dapat mendapatkan hasil yang sesuai dengan ulasan dan skor.

```

kata_positif = [
    'bagus', 'baik', 'mantap', 'hebat', 'luar biasa', 'puas', 'nyaman', 'suka',
    'sempurna', 'menyenangkan', 'keren', 'cepat', 'efisien', 'ramah', 'menyukainya',
    'recommended', 'terbaik', 'memuaskan', 'positif', 'cantik', 'murah', 'fantastis',
    'mudah', 'mengagumkan', 'berguna', 'bermanfaat', 'membantu', 'terjangkau', 'unggul', 'handal'
]

kata_negatif = [
    'jelek', 'buruk', 'tidak baik', 'kecewa', 'lelet', 'lemot', 'parah', 'tidak suka',
    'lambat', 'mahal', 'menyebalkan', 'payah', 'sampah', 'mengecewakan', 'negatif', 'jelekan',
    'tidak puas', 'sulit', 'tidak berguna', 'merugikan', 'mengecewakan', 'mengerikan',
    'buruk sekali', 'tidak menyenangkan', 'tidak membantu', 'gagal', 'cacat', 'rusak'
]

kata_netral = [
    'biasa', 'standard', 'umum', 'normal', 'rata-rata', 'cukup', 'sedang', 'lumayan'
]

```

Gambar 9. Proses Pelabelan Data 2

Dengan menggabungkan kedua proses pelabelan ini, analisis sentimen ulasan menjadi lebih akurat karena mempertimbangkan baik skor yang diberikan pengguna maupun kata-kata yang digunakan dalam ulasan.

```

# Menambahkan field positif, negatif, netral print("Menambahkan field positif,
negatif, netral...") data['positif'] = data['label2'].apply(lambda x: 1 if x ==
'positif' else 0)
data['negatif'] = data['label2'].apply(lambda x: 1 if x == 'negatif' else 0)
data['netral'] = data['label2'].apply(lambda x: 1 if x == 'netral' else 0)

```

Proses penambahan kolom untuk sentimen dilakukan dengan mengaplikasikan fungsi `labelingfromcontent` yang telah dimodifikasi untuk mengidentifikasi sentimen dari teks ulasan. Setelah setiap ulasan diberi label sebagai positif, negatif, atau netral, tiga kolom baru (positif, negatif, netral) ditambahkan ke dalam dataframe. Setiap kolom ini diisi dengan nilai 1 jika label sentimen ulasan sesuai dengan kolom tersebut, dan 0 jika tidak. Misalnya, jika sebuah ulasan diberi label positif, maka

kolom positif akan bernilai 1, sementara kolom negatif dan netral bernilai 0. Proses ini memungkinkan analisis lebih lanjut dengan mudah berdasarkan kategori sentimen yang terpisah.

Proses pelabelan data ulasan aplikasi whatsapp di google play store dapat dilihat pada gambar adalah hasil pelabelan data berikut :

Tabel 1. Hasil Tahap Pelabelan Data

NO	Content	Score	Label	Positif	Negatif	Netral
1	Utk kirim gambar bagus, cepat, telpon juga vagus. Semua bisa terhubung melalui whatsApp. Kirim video juga gambarnya jeenih, suaranya bagus. WhatsApp memang aplikasi yg sangat isrimewa karena bisa terhubung ke semua aplikasi yang tersedia pada saat sekarang. Baik youtube, instagram, twitter maupun yg lainnya. Media whatsApp ini disamping mudah dan banyak penggunaanya juga murah biayanya sehinggs banyak orang yang menyukainya. Kirim dokumen lewat media WhatsApp juga bisa shg bantu kerja	5	Positif, Negatif, Netral	1	0	0
2	Untuk Saat Ini Lumayan.. tapi untuk ke depannya tolong adakan fitur tema pada whattsap nya biar, supaya whattsap nya background nya bisa pakai tema, biar gak bosan liat background nya yang gitu gitu aja yg latar belakang nya putih polos..	3	Positif, Negatif, Netral	0	0	1
3	Mantap sekarang bisa	5	Positif, Negatif, Netral	1	0	0

update status video dengan durasi sampe 1 menit, next mungkin bisa ditambahkan fitur HD buat status video ini biar nggak pecah atau blur lagi status video yang kita up

4	Entah bagaimana caranya WA ini memakan Banyak sekali memori internal Sampai 4Gb+ Padahal isi WhatsApp ga seramai itu, Ga mengikuti saluran ² Jarang ada yg chat, Kenapa bisa bengkak begitu,, ML aja ga smpe 4Gb di HP ini bug/gimana tolong beri penjelasan #SARAN	2	Positif, Negatif, Netral	0	1	0
5	Aplikasi WhatsApp sudah mulai aneh" sekarang. Pengaturan nadanya sudah tidak bisa terpakai saat restart hp ngak tau kenapa tapi setiap kita atur nada ehh saat restart hp atau saat kita pindah mode hening dan di kembalikan ke mode dering nadanya sudah kembali lagi seperti nada dering panggilan dan telepon yang hanya bunyi sekali sja	1	Positif, Negatif, Netral	0	1	0
...				
2750	Selama pakai aplikasi whatsapp sangat keren mudah di gunakan, semakin mudah dan semakin nyaman, selama memakai whatsapp sangat membantu pekerjaan, aplikasi sangat membantu dlm segala aktifitas, sangat keren.	5	Positif, Negatif, Netral	1	0	0

Tabel 2. Hasil Data Positif

NO	Content	Score	Label	Positif	Negatif	Netral
1	Dari aplikasi ini bisa komunikasi dengan teman saudara „sungguh aplikasi yang bener bener sangat bisa membantu ,yaa memang benar sih banyak aplikasi lain yang bisa membantu tapi aplikasi ini sangat aku apresiasi banget sukses buat WhatsApp semoga makin sukses dan bisa memperbaiki fitur fiturnya ke yang lebih bagus lagi	5	Positif, Negatif, Netral	1	0	0
2	Aplikasinya bagus, kalau bisa tolong adain fitur lagu kalau mau tiap kali unggah foto! Makasih!!!	4	Positif, Negatif, Netral	1	0	0
3	Hai semua , aku suka banget pake app ini karena dia sangat mudah diakses dulu hanya dapat video call maksimal 4 orang , sekarang setelah saya perbarui di versi baru saya dapat video call dengan 8 orang guys senangnya , terus whatapp ini juga bisa di mode gelap jadi dijamin nyaman uhuy trimakasih pihak whatapp terus maju lagi ya whatapp Dan semoga whatapp pemakaian nya gratis tidak pake koutađŸ~,đŸ~...	5	Positif, Negatif, Netral	1	0	0
4	Semakin banyak yg memakai apk ini, dan bahkan menjadi pilihan apk terbaik dan utama	4	Positif, Negatif, Netral	1	0	0

dalam berbagi pesan dan moment. Tolong perbaiki kualitas gambar/video di status WhatsApp ini. Dan juga hilangkan fitur lock chat yg berlebihan. Karena bisa memicu ketidakjujuran seseorang (pasangan). Terimakasih

5	Apk bagus untuk komunikasi,,,,tapi mohon di perbaiki bug waktu pembuatan story karena story orang gak ada jam atau menit pembuatannya, kalo mau liat pembuatan story nya harus di klik dulu, itu pun jam nya gak sesuai sama pemostingan asli story nya	4	Positif, Negatif, Netral	1	0	0
...			
2750	Selama pakai aplikasi whatsapp sangat keren mudah di gunakan, semakin mudah dan semakin nyaman,selama memakai whatsapp sangat membantu pekerjaan, aplikasi sangat membantu dlm segala aktifitas, baik menu dan cara aplikasi sangat keren, untuk media foto sangat bagus hasilnya, semoga selalu di depan untuk aplikasi whatsapp semakin maju, semakin keren dan sangat TOP	5	Positif, Negatif, Netral	1	0	0

Tabel 3. Hasil Tabel Data Negatif

NO	Content	Score	Label	Positif	Negatif	Netral
1	gw mau lapor nih, akhir-akhir ini sering banget ada yang kirim virtex atau bug yang bikin WhatsApp gue jadi force close terus. Bener-bener ganggu banget deh. Tiap kali kejadian, gue harus hapus data aplikasi dan login ulang, dan itu ribet banget. Tolong dong cari solusi buat masalah ini. Mungkin bisa tambahin fitur keamanan atau filter biar pesan-pesan aneh kayak gini nggak masuk ke WhatsApp.	1	Positif, Negatif, Netral	0	1	0
2	Untuk pihak whatsapp tolong perbaiki bug yang ada di whatsapp, saya tidak bisa memberi pesan ke orang lain maupun membalas pesan, di karenakan lambang untuk "menetik pesan" Ke pendem di bawah layar yang mengakibatkan keyboard saya tidak bisa muncul karna bug tersebut. Terimakasih	1	Positif, Negatif, Netral	0	1	0
3	Tolong dong ini bug nya, tiap buka status langsung close sendiri aplikasinya, gambarnya makin burem gak karu2an... di update bukannya makin bagus malah makin gak karuan. Baru kali ini kecewa sama wa setelah bertahun2 pakai.	2	Positif, Negatif, Netral	0	1	0

4	Maaf kalau bisa, saran nya ea itu tanpa di buka, apk nya bisa menerima telfon atau panggilan chat jg atau notifikasi . Ini apk nya skrang kurang baik atau kurang lah sebagai pengguna yg menggunakan apk ini di karenakan nanti pd saatdi buka baru ada notifikasi nya atau pemberitahuan. Pd posisi hp atau jaringan nya ada aktif, kalu menurut atau sebagian pengguna wa aga kecewa dgn ada nya pengaturan ini.	2	Positif, Negatif, Netral	0	1	0
5	Aplikasi yg jelek,saya download terus masukan data diri terus ada pemberitahuan anda belum memenuhi syarat,yh jadi buat apakah aplikasinya disimpan,mo di uninstal nggak bisa,sepuluh kali dibikin malah cuman mengarah ke tutup aplikasi tutup aplikasi,padahal sudah tutup akun,sudah keluar juga,jadi saya memuat ulang	1	Positif, Negatif, Netral	0	1	0
...			
354	Saluran ga guna, lalu mode gelap dari mana buat android, pake layar amoled klo gelap ya hitam pekat bukan abu abu.. selebihnya yah biasa aja, apalagi kualitas upload status foto atau video jelek sama ngerekam ga bisa lebih dr 30 detik	1	Positif, Negatif, Netral	0	1	0

Tabel 4. Hasil Data Netral

NO	Content	Score	Label	Positif	Negatif	Netral
1	WhatsApp,,memberitahukan bahwa kesalahan ada dipihak WhatsApp,,dgn alasan tdk sengaja dan mohon maaf..tpi nomor saya verifikasi tetap aja gk bisa..alasan hrs pakai WhatsApp yg resmi..ini maksudnya apa..udah satu Minggu ini saya kehilangan teman group komunitas kami...kemarin ganti nomor udah bisa aktif cuma satu malam , Pagi pagi sy mau buka whatsapp sdh ada peringatan ,,bawa anda tdk bisa menggunakan whatsapp lagi Krn spam. Coba bisa tunjukan spam yg menurut whatsapp berasal dari saya,,	3	Positif, Negatif, Netral	0	0	1
2	Untuk Saat Ini Lumayan.. tapi untuk ke depannya tolong adakan fitur tema pada whatsapp nya biar, supaya whatsapp nya background nya bisa pakai tema, biar gak bosan liat background nya yang gitu gitu aja yg latar belakang nya putih polos..	3	Positif, Negatif, Netral	0	0	1
3	Apk whatsapp saya upgrade otomatis, dan menurut saya ini jelek bgt tampilannya bikin pusing apalagi bagian story tampilannya aduhh, kembaliin lagi tampilan yg sebelumnya	3	Positif, Negatif, Netral	0	0	1

4	Kenapa tiba2 akun whatsapp kena spam. Sedangkan menggunakan wa saya tidak sering. Untuk menormalkannya kembali kenapa dipersulit. Mohon jawabannya. Dan tinjau kembali. Hal2 yg bisa membuat pengguna jadi tidak nyaman	3	Positif, Negatif, Netral	0	0	1
5	Saya terkena pelanggaran spam, padahal saya tidak melakukan pelanggaran apapun, saat saya sudah memverifikasi nomor saya malah diberi tulisan "anda tidak memakai WhatsApp resmi" saat saya klik "unduh" saya pergi ke sebuah web saya mengklik "dapatkan di playstore" saat saya sudah di playstore saya melihat bahwa saya sudah mendownload WhatsApp saya juga mengklik unduh WhatsApp secara langsung ternyata sama saja dan sekarang saya tidak bisa berkomunikasi dengan teman" saya tolong segera diperbaiki.	3	Positif, Negatif, Netral	0	0	1
--	-----	---	----			

Tabel 5. Tabel Data Ulasan Yang Tidak Sesuai Score

N O	Content	Score	Label	Positif	Negatif	Netral
1	tolong diperbaiki agar tidak terblokir, kenapa setiap mengirimkan pesan sekali 1000 kontak selalu terblokir. tolong	5	Positif, Negatif, Netral	1	0	0

	diperbaiki agar bisa mengirimkan lebih dari 1000 kontak seluruh teman. terimakasih. tolong perkuat dan perbaiki					
2	Gimana ya kok jadi error sih, padahal centang dua malah jadi centang dua biru, semua chat jd centang 2 biru padahal belum dibaca, tolong di perbaiki secepatnya, gk nyaman karna seolah olah udh di read padahal belum di read	5	Positif, Negatif, Netral	1	0	0
3	game nya bagus sekali, aku suka saat saya memakai hero zilong. bagus sekali	5	Positif, Negatif, Netral	1	0	0
4	Aplikasi sudah bagus dan cepat, mohon tolong perbaiki bug terhadap kontak saya tiba" hilang	5	Positif, Negatif, Netral	1	0	0
5	Aplikasinya bagus banget sih tapi tolong, tampilan jumlah teks pesan yang terkirim ke setiap kontak di tampilkan lagi kaya dulu terus akhir akhir ini notif suka telat masuk padahal jaringan lancar tolong di perbaiki	2	Positif, Negatif, Netral	0	1	0
6	Saat ini WA aplikasi chat dan tlp	5	Positif, Negatif, Netral	1	0	0

yg paling bagus.. Sedikit saran, sebaiknya di tambahkan ON/OFF button.. Jadi bagi user yg sedang tdk mau di ganggu dia tinggal tekan tombol ON/OFF..bkn HP nya yg dimatikan.. Mudah2an ke depannya ada tombol ON/OFF ini... Terima Kasih...

7	Apk ini bagus banget...tapi mohon maaf kepada developer WhatsApp karena kualitas untuk video call dan kirim video / gambar ke orang lain melalui WhatsApp grafiknya kurang bagus...Dari saya sih semoga ditingkatkan lagi untuk bisa video dan gambar makin diperbaiki biar bisa lebih bagus..â~i,•	2	Positif, Negatif, Netral	0	1	0
--	-----	---	----			

Ulasan positif, skor rendah: Ada pengguna yang memberikan ulasan positif tentang aplikasi, misalnya memuji fungsionalitas atau fitur, tetapi memberikan skor rendah (1 atau 2 bintang). Ini bisa terjadi karena kesalahan atau ketidaksesuaian antara penilaian pengguna dan sistem pemberian skor.

Ulasan negatif, skor tinggi: Sebaliknya, ada pengguna yang memberikan ulasan negatif, tetapi secara tidak sengaja memberi skor yang tinggi (4 atau 5 bintang). Ini juga bisa disebabkan oleh kesalahan pengguna atau ketidaktahuan tentang bagaimana sistem penilaian bekerja.

C. Tahap Preprocessing

Cleaning, case folding, stopword removal, dan tokenizing adalah semua proses yang dilakukan pada tahap *preprocessing* ini. Proses ini digunakan untuk membersihkan data dari tanda baca dan simbol yang tidak dibutuhkan, seperti tanda titik, tanda koma, tanda tanya, dan tanda seru, serta yang tidak dibutuhkan. Pada titik ini, Anda juga dapat memotong sebuah kalimat menjadi potongan kata, menghilangkan *stopword*, mengubah semua huruf dalam data menjadi huruf kecil, dan mengubah kata berimbuhan menjadi kata dasar.

1. *Cleaning*

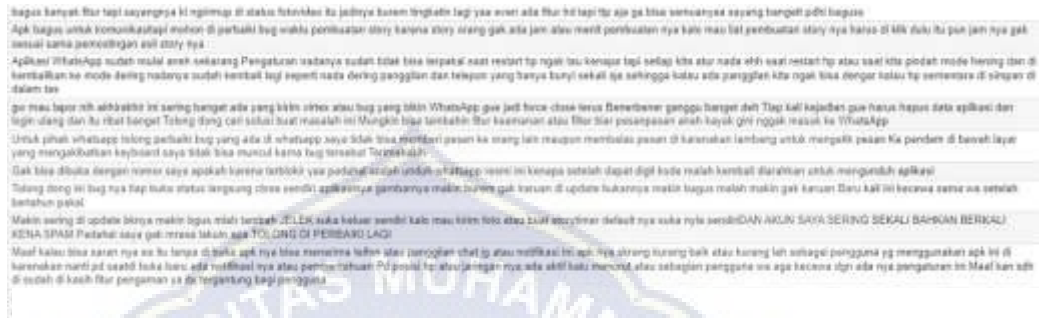
Pada tahap pembersihan data yang sudah dilabeli, karakter, seperti tanda baca, dan simbol yang tidak relevan, akan dihapus. Karena penelitian ini hanya membutuhkan ulasan negatif dan positif, ulasan berlabel netral akan dihapus. Gambar berikut menunjukkan hasil dari



Gambar 10. Data Ulasan WhatsApp Setelah Dilakukan Cleaning tahap membersihkan:

2. Case folding

Dalam tahap *case folding* ini, huruf dalam data ulasan akan diubah menjadi huruf kecil atau *lowercase* supaya komputer lebih mudah membacanya. Gambar berikut menunjukkan hasil dari tahap persiapan



Gambar 11. Hasil Data Ulasan WhatsApp Sebelum Proses Case Folding

case:



Gambar 12. Hasil Data Ulasan WhatsApp Setelah Proses Case Folding

3. Stopword Removal

Pada titik ini, kata-kata yang biasa dan sering digunakan dalam teks seperti kata depan, kata ganti, dan kata penghubung dihilangkan. Tujuannya adalah untuk membuat analisis *sentimen* lebih akurat. Hasil dari tahap penghapusan *stopword* ditunjukkan dalam gambar berikut:

Tabel 6. Hasil Data Ulasan WhatsApp Setelah Tahap Stopword Removal

Sebelum	Sesudah
<p>keluhan saat videocall gambar terjeda tidak bisa dikembalikan seperti semula menyalakan kamera lagi harus mematikan sambungan telepon terlebih dahulu share lokasi tidak update lokasi terbaru padahal sinyal pengirim bagus harus mengirim berulangung untuk mengetahui lokasi terkini kalau membuat story kualitas gambar dan video jelek sekali bahkan untuk gambar tulisan juga kadang tidak terbaca bagaimana ini wa sudah lama pakai aplikasi ini perbaiki lagi hal kecil seperti ini</p>	<p>keluhan videocall gambar terjeda dikembalikan menyalakan kamera mematikan sambungan telepon share lokasi update lokasi terbaru sinyal pengirim bagus mengirim berulangung lokasi terkini story kualitas gambar video jelek gambar tulisan kadang terbaca wa pakai aplikasi perbaiki</p>

4. *Tokenizing*

Dalam penelitian ini, *library* NLTK digunakan untuk *tokenize* ulasan pada tahap *tokenizing*. Dengan menggunakan teknik *tokenizing*, kita dapat memotong satu kalimat dari satu ulasan menjadi potongan-potongan yang terdiri dari potongan kata. Gambar berikut menunjukkan hasil sebelum dan sesudah *tokenizing*:

Tabel 7. Hasil Data Ulasan WhatsApp Setelah Tahap Tokenizing

Sebelum	Sesudah
---------	---------

keluhan videocall gambar terjeda dikembalikan menyalakan kamera mematikan sambungan telepon share lokasi update lokasi terbaru sinyal pengirim bagus mengirim berulangulang lokasi terkini story kualitas gambar video jelek gambar tulisan kadang terbaca wa pakai aplikasi perbaiki	['keluhan', 'videocall', 'gambar', 'terjeda', 'dikembalikan', 'menyalakan', 'kamera', 'mematikan', 'sambungan', 'telepon', 'share', 'lokasi', 'update', 'lokasi', 'terbaru', 'sinyal', 'pengirim', 'bagus', 'mengirim', 'berulangulang', 'lokasi', 'terkini', 'story', 'kualitas', 'gambar', 'video', 'jelek', 'gambar', 'tulisan', 'kadang', 'terbaca', 'wa', 'pakai', 'aplikasi', 'perbaiki']
--	---

5. Steaming

Stemming adalah proses dalam text preprocessing untuk mengubah katakata menjadi bentuk dasar atau kata dasar (root word) dengan cara memotong akhiran kata. Tujuan stemming adalah untuk mengurangi variasi kata yang muncul dalam teks dan mempertahankan makna kata meskipun berbeda dalam bentuk. Contohnya, kata "membaca", "membacaan", dan "membacakan" semuanya memiliki akar kata yang sama yaitu "baca". Dengan menggunakan algoritma stemming, ketiga kata tersebut akan diubah menjadi "baca". Untuk melakukan stemming bahasa indonesia kita dapat menggunakan library python sastrawi. berikut adalah tabel hasil data ulasan aplikasi *WhatsApp* setelah tahap stemming.

Tabel 8. Hasil Data Ulasan WhatsApp Setelah Tahap Stemming

Sebelum	Sesudah
keluhan videocall gambar terjeda dikembalikan menyalakan kamera mematikan sambungan telepon share lokasi update lokasi terbaru sinyal pengirim bagus mengirim berulangulang lokasi terkini story kualitas gambar	keluh videocall gambar jeda kembali nyala kamera mati sambung telepon share lokasi update lokasi baru sinyal kirim bagus kirim berulangulang lokasi kini story kualitas gambar video jelek gambar tulisan kadang baca wa pakai aplikasi baik

video jelek gambar tulisan
kadang terbaca wa pakai
aplikasi perbaiki

D. Spiliting Data

Data dibagi menjadi set pelatihan dan set pengujian dengan rasio 80:20. Dari total 2750 data, sebanyak 2185 data digunakan untuk set pelatihan, dan 565 data digunakan untuk set pengujian. Set pelatihan digunakan untuk melatih model, sementara set pengujian digunakan untuk mengevaluasi kinerja model. Dengan pembagian ini, kami memastikan bahwa model memiliki cukup data untuk belajar dan dapat diuji secara akurat pada data yang tidak dilihat sebelumnya.

Tabel 9. Jumlah Data

Data Pelatihan	Data Pengujian
2185	565

E. Pembobotan TF-IDF

TF-IDF (Term Frequency-Inverse Document Frequency) digunakan untuk mengubah teks ulasan menjadi vektor numerik yang dapat diproses oleh model pembelajaran mesin. TF-IDF memberikan bobot yang lebih tinggi pada kata-kata yang sering muncul dalam dokumen tertentu tetapi jarang muncul dalam dokumen lain. Ini membantu model untuk lebih fokus pada kata-kata yang lebih informatif dan mengurangi dampak dari kata-kata umum yang kurang berarti.

```

▶ # Pembobotan data, TF-IDF
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(data['stemmed_text'])
y = data['label']

```

Gambar 13. Hasil Tahap Pembobotan TF-IDF

F. Klasifikasi

Model *Naive bayes* digunakan untuk mengklasifikasikan ulasan sebagai *sentimen* positif atau negatif. Model dilatih menggunakan set pelatihan dan dievaluasi menggunakan set pengujian. Hasil evaluasi mencakup akurasi, matriks konfusi, dan laporan klasifikasi yang memberikan informasi detail tentang kinerja model. Setelah proses pelatihan, model dievaluasi dengan menggunakan metrik berikut:

1. Akurasi (Accuracy): Persentase prediksi yang benar dari total prediksi yang dibuat oleh model.
2. Matriks Konfusi (Confusion Matrix): Tabel yang menunjukkan jumlah prediksi benar dan salah yang dibuat oleh model untuk setiap kelas. Matriks ini membantu dalam memahami kesalahan yang dibuat oleh model dalam klasifikasi.
3. Laporan Klasifikasi (Classification Report): Laporan yang mencakup metrik seperti *precision*, *recall*, dan F1-score untuk setiap kelas. *Precision* mengukur seberapa tepat prediksi positif, *recall* mengukur seberapa baik model dapat mendeteksi semua instance positif, dan F1-score adalah harmonisasi antara *precision* dan *recall*.

```

accuracy = accuracy_score(y_test, predicted)
precision = precision_score(y_test, predicted, average='weighted')

```



```

recall = recall_score(y_test, predicted, average='weighted') f1 = f1_score(y_test,
predicted, average='weighted') confusion = confusion_matrix(y_test, predicted)
class_report = classification_report(y_test, predicted, zero_division=0)

print("Akurasi MultinomialNB: ", accuracy) print("Presisi
MultinomialNB: ", precision) print("Recall MultinomialNB: ",
recall) print("F1 Score MultinomialNB: ", f1) print(f'Matriks
Confusion:\n {confusion}')

print("=====")
print(class_report)

```

Gambar 14. Hasil Proses Tahap Klasifikasi Naïve Bayes

G. Hasil Pengujian

Berdasarkan hasil proses tahap preprocessing didapatkan jumlah data sebanyak 2750 data ulasan aplikasi dari dataset. Kemudian pada proses splitting data, data tersebut dibagi menjadi dua yaitu data latih dan data uji dengan skenario pembagian data 80:20, yaitu dengan 80% data latih dan 20% data uji. Jadi, data latih sebanyak 2185 data dan data uji sebanyak 565 data. Berdasarkan hasil analisis ulasan aplikasi menggunakan metode Naive Bayes, didapatkan nilai sebagai berikut:


```

Akurasi MultinomialNB: 0.8550561797752809
Presisi MultinomialNB: 0.8702801594636412
Recall MultinomialNB: 0.8550561797752809
F1 Score MultinomialNB: 0.8517522451036028
Matriks Confusion:
[[295  2  0]
 [ 11 269 15]
 [  6  95 197]]
=====

```

	precision	recall	f1-score	support
negatif	0.95	0.99	0.97	297
netral	0.73	0.91	0.81	295
positif	0.93	0.66	0.77	298
accuracy			0.86	890
macro avg	0.87	0.86	0.85	890
weighted avg	0.87	0.86	0.85	890

Gambar 15. Hasil Proses Tahap Klasifikasi Naïve Bayes

Hasil diatas menunjukkan bahwa kombinasi antara pendekatan Natural Language Processing (NLP) dan metode Naive Bayes Classifier memberikan hasil pengelompokan sentimen yang lebih akurat dan relevan. Efektivitas pendekatan ini terlihat dari peningkatan akurasi dan relevansi hasil klasifikasi. NLP memungkinkan pemrosesan teks ulasan yang lebih mendalam dengan mempertimbangkan konteks dan makna dari setiap ulasan, sementara Naive Bayes Classifier memberikan metode yang sederhana namun efektif dalam mengelompokkan data ke dalam kategori sentimen. Dengan demikian, pendekatan ini terbukti meningkatkan kinerja klasifikasi dibandingkan metode yang hanya menggunakan satu teknik.

Dalam hitungan manual ini, kita menggunakan Teorema Bayes untuk menghitung probabilitas posterior dari setiap label (Negatif, Positif, Netral) berdasarkan data yang ada.

Konsep dasar dari Teorema Bayes adalah:

$$P(i) = \frac{P(i) \cdot O}{O} \dots \dots \dots (1)$$

Dimana:

- $P(Y|X)$ adalah probabilitas posterior dari Y mengingat X.
- $P(X|Y)$ adalah likelihood, yaitu probabilitas dari X mengingat Y.
- $P(Y)$ adalah prior, yaitu probabilitas awal dari Y.
- $P(X)$ adalah evidence, yaitu probabilitas dari X.

Selanjutnya, untuk menghitung akurasi keseluruhan, kita menjumlahkan total prediksi yang benar (True Negatives + True Positives + True Neutrals) dan membaginya dengan total data, kemudian dikali 100 untuk mendapatkan persentase dan disesuaikan dengan hitungan datanya maka

- Total prediksi yang benar (True Negatives + True Positives + True Neutrals) = $295 + 269 + 197 = 761$
- Total data = 890
- Akurasi Keseluruhan = $(\text{Total Prediksi Benar} / \text{Total Data}) \times 100$
- Akurasi Keseluruhan = $(761 / 890) \times 100 = 85.51\%$

$$h(\text{ }) = \left(\frac{\text{Total Prediksi Benar}}{\text{Total Data}} \right) \times 100$$

Dengan memasukkan nilai:

$$h(\text{ }) = \left(\frac{295 + 269 + 197}{761} \right) \times 100$$

$$= 85.51\%$$

Hasil Perhitungan:

- $(\text{ }) = 97.02\%$
- $(\text{ }) = 92.87\%$
- $(|) = 72.29\%$

- Akurasi keseluruhan (manual): 85.51%



H. Confusion Matrix

Pada tahap pengujian menggunakan metode confusion matrix, dimana peneliti akan membandingkan hasil klasifikasi yang dilakukan secara manual dengan hasil klasifikasi yang telah dihasilkan oleh model dengan metode naive bayes. Dalam membangun model ini, peneliti menggunakan sebuah ukuran atau biasa disebut metric untuk mengetahui seberapa baik kinerja model yang dibuat. Precision, recall, fi-score, macro avg dan weighted avg digunakan untuk evaluasi metric dalam model klasifikasi ini.

Gambar berikut adalah Hasil Confusion matrix naïve bayes untuk data ulasan aplikasi WhatsApp :

```

Akurasi MultinomialNB: 0.8550561797752809
Presisi MultinomialNB: 0.8702801594636412
Recall MultinomialNB: 0.8550561797752809
F1 Score MultinomialNB: 0.8517522451036028
Matriks Confusion:
[[295  2  0]
 [ 11 269 15]
 [  6  95 197]]
    
```

Gambar 16. Confusion Matrix Data ulasan WhatsApp

Setelah melakukan pengujian data , maka didapat hasil data yang meliputi seperti konten ,:

Tabel 10. Hasil Tahap Akhir Pelabelan Klasifikasi Data True Positif

No	Content	Data bersih	Label	True Label
1.	Utk kirim gambar bagus, cepat, telpon juga bagus. Semua bisa terhubung melalui whatsApp. Kirim video juga gambarnya jeenih, suaranya bagus. WhatsApp memang aplikasi yg sangat isrimewa karena bisa terhubung ke semua aplikasi	utk kirim gambar bagus cepat telpon vague hubung whatsapp kirim video gambar jeenih suara bagus whatsapp aplikasi yg isrimewa hubung aplikasi sedia youtube instagram	Positif	Positif

	yang tersedia pada saat sekarang. Baik youtube, instagram, twitter maupun yg lainnya. Media whatsApp ini disamping mudah dan banyak penggunaanya juga murah biayanya sehinggs banyak orang yang menyukainya. Kirim dokumen lewat media WhatsApp juga bisa shg bantu kerja	twitter yg media whatsapp samping mudah guna murah biaya sehinggs orang suka kirim dokumen media whatsapp shg bantu kerja		
2	Apk-nya bagus! Ada fitur-fitur baru yang bermanfaat dan mempermudah dalam pemakaian apk-nya.cuman saya kasih bintang 4 karena ada bug, yaitu saat post video/foto HD di status WhatsApp resolusinya jadi jelek, jadi kaya pixel-pixel itu. Mohon perbaiki bug nya	apknya bagus fiturfitur manfaat muda pakai apknya cuman kasih bintang bug post videofoto hd status whatsapp resolusi jelek kaya pixelpixel mohon baik bug nya	Positif	Positif
3	Tolong dong untuk whatsapp inu lagi eror ya dibagian kameranya, kalo mau ambil gambar baik ngirim foto atau buat story, suka zoom out sendiri, tapi hasilnya normal. Tolong segera diperbaiki ya. Jadi susah buat kerja, soalnya sering ambil foto buat kirim dokumen.	tolong whatsapp inu eror ya bagi fitur kamera kalo ambil gambar ngirim foto story suka zoom out hasil normal tolong baik ya susah kerja ambil foto kirim dokumen	Positif	Positif
3	Sangat membantu dalam berkomunikasi dengan saudara juga teman.serta sangat membantu juga dalam pekerjaan ,makasih saya sangat puas dengan layanan ni	bantu komunikasi saudara temanserta bantu kerja makasih puas layanan ni	Positif	Positif

4	mantap suka aplikasi nya sangat membantu	mantap suka aplikasi nya bantu	Positif	Positif
5	Gimana ya kok jadi error sih, padahal centang dua malah jadi centang dua biru, semua chat jd centang 2 biru padahal belum kebaca, tolong di perbaiki secepatnya, gk nyaman karna seolah olah udh di read padahal belum di read	gimana ya error sih centang centang biru chat jd centang biru baca tolong baik cepat gk nyaman karna olah udh read read	Positif	Positif

Tabel 11. Hasil Tahap Akhir Pelabelan Klasifikasi Data True Negatif

No	Content	Data Bersih	Label	True Label
1	kenapa harus mengaktifkan fitur info terakhir dilihat atau online diri sendiri hanya untuk melihat info terakhir dilihat atau online nya orang lain ah mengecewakan sekali lah aaaaaaarrrrrggghhhhhhhh hhh	aktif fitur info online info online nya orang ah kecewa aaaaaaarrrrrggghhhhhhhh	Negatif	Negatif
2	Saya kasih bintang 1 karena ini WhatsApp sangat merugikan saya .. dan saya selalu ke blokir tanpa ada masalah udh lebih dari 30x saya terbelokir dan bisa LG di gunakan setelah itu terbelokir LG mengecewakan dan harus hati2 dengan WhatsApp	kasih bintang whatsapp rugi blokir udh x terbelokir lg terbelokir lg kecewa hati whatsapp	Negatif	Negatif
3	Saya kecewe dengan keamanan whatsapp	kecewe aman whatsapp beli sim card hp satunyya	Negatif	Negatif

yang sekarang, saya baru beli SIM CARD baru untuk hp satunya, saya login dengan SIM CARD baru terus saya misscall nomor lama saya pake nomor baru, beberapa saat tidak berdering di nomor yang lamaa, pas di cek dinomor baru Whatsappnya malah dinonaktifkann. Disuruh aju banding, yang perihal masalahnya sama sekali saya tidak tau, sangat buruk performa yang sekarang

login sim card misscall nomor pake nomor dering nomor lamaa pas cek nomor whatsappnya dinonaktifkann suruh aju banding perihal tau buruk performa

4	Setelah ada update aplikasi WA ini semua semua payah, ada chat wa masuk maupun telpon masuk nanti gak nampak. Baru nampak setelah kita buka aplikasi WA. Ini aplikasi WA tolong diperbaharui sistemnya. Sangat lelet dan bobrok kali pekerjaan yang punya aplikasi ini.	update aplikasi wa payah chat wa masuk telpon masuk gak nampak nampak buka aplikasi wa aplikasi wa tolong baharu sistem lelet bobrok kali kerja aplikasi	Negatif	Negatif
5	Tidak pernah membuat pelanggaran tapi nomer saya di blokir oleh pihak WhatsApp, Tolong lebih di perhatikan saya sudah menggunakan selama bertahun tahun tapi kali ini saya di buat kecewa oleh WhatsApp.	langgar nomer blokir whatsapp tolong perhati tahun kali kecewa whatsapp	Negatif	Negatif
---	---	--		

Tabel 12. Hasil Tahap Akhir Pelabelan Klasifikasi Data True Netral

No	Content	Data Bersih	Label	True Label
1	sdh mengalami keuntungan dan kelebihan menggunakan Whaatsapp sudah dapat mengerti manfaat serta proteksi dari penggunaan aplikasi ini. Sangat membantu. dalam menyajikan dokumen. Terima kasih atas pelayanannya. Sangat bermanfaat. hebat. Terasa sangat bermanfaat dengan Fasilitas WA ini terbantu dalam berbagai keperluan dan kepentingan komunikasi.	sdh alami untung lebih whaatsapp erti manfaat proteksi guna aplikasi inisangat bantu saji dokumen terima kasih layan bermanfaatbatterasa manfaat fasilitas wa bantu perlu kepenting komunikasi	Netral	Netral
2	Jangan diperbarui terus wa nya, kasian yang MB nya dikit, belum punya uang buat beli hp yang MB nya lebih tinggi/kuat, ga bisa buat punya wa, padahal wa penting buat hubungi keluarga/bisnis lainnya, kakak saling membantu/mempermudah ah lebih baik	baru wa nyakasian mb nya dikit uang beli hp mb nya tinggikuatga wa wa hubung keluargabisnis kakak membantumempermudah	Netral	Netral
3	ya coba tp lumayanlah tuk hiburan.	ya coba tp lumayan tuk hibur	Netral	Netral
4	Tolong wharsapp kenapa setelah di update malah makin jelek saat kirim foto dan vidio gambarnya pecah bahkan saat	tolong wharsapp update jelek kirim foto vidio gambar pecah update story pecah tolong kualitas nya bagusin lebih nya bagus kualitas update story kirim	Netral	Netral

update story pun masih pecah tolong kualitas nya di bagusin lagi selain itu selebih nya bagus hanya kurang di kualitas update story dan dan kirim foto video saja yang jelek ., Tolong dengan sangat whatsapp perbaiki lagi thanks .,untuk sementara bintang 3 dlu

foto video jelek tolong whatsapp baik thanks bintang dlu

5	<p>Pilih ditengah . Antara suka dan kecewa Suka karna aplikasi ini sangat membantu saya dalam bersosialisasi dengan orang lain ... baik teman, teman dekat , keluarga, dan juga gebetan .aplikasi ini juga membantu saya untuk lebih dekat dengannya .hua2 bucin Kecewanya karna ... aplikasi kadang berhenti tanpa sebab, kadang juga eror ..mau upload foto atau video ke sw ga bisa karna wa berhenti ... padahal udah ngedit alakadarnya . Mohon wa nya suruh intropesi diri . Jangan bikin emosi.trimakasih</p>	<p>pilih tengah suka kecewa suka karna aplikasi bantu sosial orang teman teman keluarga gebetan aplikasi bantu dengan hua bucin kecewa karna aplikasi kadang henti kadang eror upload foto video sw ga karna wa henti udah ngedit alakadarnya mohon wa nya suruh intropesi bikin emositrimakasih</p>	Netral	Netral
--	---	--	---	

BAB V

PENUTUP

A. Kesimpulan

Berdasarkan penelitian analisis sentimen yang telah dilakukan, beberapa kesimpulan dapat diambil:

1. Penelitian ini berhasil menganalisis sentimen ulasan pengguna terhadap aplikasi WhatsApp di Google Play Store menggunakan metode Naive Bayes Classifier. Metode ini efektif dalam mengkategorikan sentimen ulasan pengguna ke dalam tiga kategori utama: positif, negatif, dan netral. Proses analisis ini menunjukkan bahwa Naive Bayes Classifier mampu menghasilkan akurasi yang cukup tinggi, yaitu sebesar 85,51%, dalam klasifikasi sentimen ulasan.
2. Faktor-faktor yang memengaruhi sentimen pengguna dalam ulasan mereka terhadap aplikasi WhatsApp meliputi kualitas layanan, fitur aplikasi, Pembaruan dan Perubahan, Keamanan dan Privasi, Dukungan Pengguna dan pengalaman pengguna secara keseluruhan. Analisis terhadap ulasan pengguna menunjukkan bahwa sentimen positif umumnya dipengaruhi oleh kepuasan terhadap fitur-fitur yang ditawarkan oleh WhatsApp, sementara sentimen negatif sering kali muncul karena masalah teknis atau kekurangan tertentu dalam aplikasi.
3. Klasifikasi sentimen ulasan pengguna WhatsApp ke dalam kategori positif, negatif, atau netral dapat dilakukan secara efektif dengan mengintegrasikan pendekatan Natural Language Processing (NLP) dengan metode Naive Bayes Classifier. Kombinasi ini memungkinkan pemrosesan teks ulasan yang lebih mendalam, sehingga hasil klasifikasi menjadi lebih akurat dan relevan. Hal ini menunjukkan bahwa pendekatan NLP yang diterapkan dalam penelitian ini telah berhasil meningkatkan kinerja Naive Bayes Classifier dalam analisis sentimen ulasan pengguna di platform Google Play Store.

B. Saran

Pemantauan sentimen secara berkala akan membantu dalam mengidentifikasi tren perubahan sentimen dan memungkinkan implementasi solusi yang lebih tepat. Untuk mempercepat proses analisis data dalam skala besar, dan juga menerapkan teknik optimasi seperti pemrosesan paralel atau menggunakan perangkat keras yang lebih kuat.



DAFTAR PUSTAKA

- Agustina, N., Citra, D. H., Purnama, W., Nisa, C., & Kurnia, A. R. (2022). Implementasi Algoritma *Naive bayes* untuk Analisis *Sentimen* Ulasan Shopee pada *Google Play Store*. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 2(1), 47–54. <https://doi.org/10.57152/malcom.v2i1.195>
- Amaliah, F., & Dwi Nuryana, I. K. (2022). Perbandingan Akurasi Metode Lexicon Based Dan *Naive bayes Classifier* Pada Analisis *Sentimen* Pendapat Masyarakat Terhadap Aplikasi Investasi Pada Media Twitter. *Journal of Informatics and Computer Science (JINACS)*, 3(03), 384–393. <https://doi.org/10.26740/jinacs.v3n03.p384-393>
- Apriyanto Alhamad, Azminuddin I. S. Azis, Budy Santoso, & Sunarto Taliki. (2019). Prediksi Penyakit Jantung Menggunakan Metode-Metode Machine Learning Berbasis Ensemble – Weighted Vote. *Jurnal Edukasi Dan Penelitian Informatika*, 5(3), 352.
- Arlovin, T., Kusriani, & Kusnawi. (2024). Analisis *Sentimen* Review Pengguna Aplikasi Fizzo Novel Di *Google Play* Menggunakan Algoritma Naive Bayes. *Jurnal Informatika Teknologi Dan Sains (Jinteks)*, 6(1), 65–70. <https://doi.org/10.51401/jinteks.v6i1.3909>
- Azriansyah, N., Indra, E., & Azriansyah, N. (2023). Penerapan *Natural language processing* Untuk Analisis. *Jurnal Ilmiah BETRIK (Besemah Teknologi Informasi Dan Komputer)*, 14(2), 273–282. <https://ejournal.pppmitpa.or.id/index.php/betrik/article/view/96>
- Baskara, R., & Rahma, F. (2022). Implementasi Web Scraping Pada Media Sosial Instagram. *Automata*, 3, 1–3.
- Faadilah, A. (2020). Analisis *Sentimen* Pada Ulasan Aplikasi Tokopedia di *Google Play Store* Menggunakan Metode Long Short Term Memory. 1–46.

- Furqan, M., Sriani, S., & Shidqi, M. N. (2023). Chatbot Telegram Menggunakan *Natural language processing*. *Walisongo Journal of Information Technology*, 5(1), 15–26. <https://doi.org/10.21580/wjit.2023.5.1.14793>
- Gumilar, T. S., Astuti, R., & Wijaya, Y. A. (2024). Analisis *Sentimen* Ulasan Aplikasi Lita Di *Play Store* Menggunakan Algoritma Naive Bayes. *JATI (Jurnal Mahasiswa Teknik ...)*, 8(1), 543–550. <https://ejournal.itn.ac.id/index.php/jati/article/view/8778%0Ahttps://ejournal.itn.ac.id/index.php/jati/article/download/8778/5018>
- Informatika, P., Teknik, F., & Nusantara, U. M. (2023). *JTS : Jurnal Teknik P-ISSN: 2302-8734 E-ISSN: 2581-0006*. 12(02), 128–137.
- Kosasih, R. (2021). PENGGUNAAN METODE LINEAR DISCRIMINANT ANALYSIS UNTUK PENGENALAN WAJAH DENGAN MEMBANDINGKAN BANYAKNYA DATA LATIH. *Jurnal Ilmiah Teknologi Dan Rekayasa*, 26(1), 25–34. <https://doi.org/10.35760/tr.2021.v26i1.3520>
- Maulana, Y. (2023). Optimalisasi Support Vector Machine (SVM) Menggunakan Pelabelan VADER pada Analisis *Sentimen* Ulasan *Google Classroom*. *Doctoral Dissertation, Universitas Siliwangi*, 2023.
- Muhamad Anton Permana1), S. W. S. S. 3). (2023). Analisis *Sentimen* Penggunaan Aplikasi VideoConference Pada Ulasan *Google Play Store* Menggunakan Metode Nbc (*Naive bayes Classifier*). *Jurnal Riset Sistem Informasi Dan Teknologi Informasi (JURSISTEKNI)*, 5(1), 178–191. <https://jursistekni.nusaputra.ac.id/article/view/178>
- Nurian, A., Ma`arif, M. S., Amalia, I. N., & Rozikin, C. (2024). Analisis *Sentimen* Pengguna Aplikasi Shopee Pada Situs *Google Play* Menggunakan *Naive bayes Classifier*. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(1), 704–713. <https://doi.org/10.23960/jitet.v12i1.3631>
- Prasetya, M., Wulandari, M., & Nikmah, S. A. (2024). Implementasi NLP

Suarna, N., Prihartono, W., Informatika, T., Akuntansi, K., & Store, P. (2024). *PENERAPAN NLP (NATURAL LANGUAGE PROCESSING) DALAM ANALISIS*. 8(2), 1841–1846.

Surya Sayogo, D., Irawan, B., & Bahtiar, A. (2024). Analisis *Sentimen* Ulasan Instagram Di *Google Play Store* Menggunakan Algoritma Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3314–3319. <https://doi.org/10.36040/jati.v7i6.8178>

Yolanda, K., Yusra, Y., & Fikry, M. (2023). Klasifikasi *Sentimen* Ulasan Aplikasi *WhatsApp* di *Play Store* Menggunakan *Naive bayes Classifier*. *J-Intech*, 11(1), 1–9. <https://doi.org/10.32664/j-intech.v11i1.867>



LAMPIRAN

Lampiran 3. Hasil Data Sebelum Di Filter Dan Diubah Menjadi Data Frame

id	nama	username	password	score	score	score	score
1	123456789	123456789	123456789	1	1	1	1
2	123456789	123456789	123456789	2	2	2	2
3	123456789	123456789	123456789	3	3	3	3
4	123456789	123456789	123456789	4	4	4	4
5	123456789	123456789	123456789	5	5	5	5
6	123456789	123456789	123456789	6	6	6	6
7	123456789	123456789	123456789	7	7	7	7
8	123456789	123456789	123456789	8	8	8	8
9	123456789	123456789	123456789	9	9	9	9
10	123456789	123456789	123456789	10	10	10	10

Lampiran 2. Hasil data sesudah di filter dan diubah menjadi Data Frame

username	content	score
Pengguna Google	Enggak lah itu aku zina baperi mendingan bawak. Anaknya video, pemangsaan video pada status, dan audio dari video yang di unggah ke status lalu berlagu video lebih dari 1 menit...	1
Pengguna Google	Aku apasih aku nnta suara dan video total 10000. ada bany. Content kaba apasih status format gbr atau photo, suara dan video total skrin aja mlay.	1
Pengguna Google	Mantap sei zany bisa update status video dengan durasi sampai 1 menit. bawak mending bisa diantarkan. Itu HD buat status video di bisa nggak pecan abis tra lagi status video yang kita ap...	5
Pengguna Google	Trenk bagaimana caranya WA ini merakun. Bawak lokal merakun interval sampai 4500. Pabrik di WhatsApp ya ini merakun. Ga mending sakun? Jepang ada yg chat, Kanada ada bawak...	2
Pengguna Google	Sabarannya sih bawak lagi entah kenapa di parangin. aya bawak mau masuk ka WhatsApp itu harus masukin nomor tuh bawak nanti di kirim kode. Jadinya chat bawak jadi bawak buat aplikasi...	4
Pengguna Google	Kenapa bawak wa stng tangkalinnya bawak?? Ga bawak update video ya bawak ngak korting dan bawak ka bawak ag mauca?!	4
Pengguna Google	Tolong admin WhatsApp bawak video saat aku saya di bawak pethel saya bawak mendingan panyaman dan panyaman yang dibuat oleh WhatsApp selap saat aku saya di bawak lalu bawak...	0
Pengguna Google	Saya sudah 2x daftarin nomor baru saya, tapi kenapa selalu kena spant. Bawak status mendingan panyaman, status wa bawak malah paku mauk akaw WA bawak ream, padahal saya bawak...	1
Pengguna Google	Dari aplikasi ini bisa komunikasi dengan teman sekiranya...juga aplikasi yang bawak bawak yang bawak bawak yang bawak bawak yang bawak bawak yang bawak bawak yang bawak bawak...	0
Pengguna Google	Apa kabarnya bawak. kaku itu bawak...	4

Lampiran 1. Proses Pelabelan Data 1

```

# Pelabelan data berdasarkan score
def label_data(score):
    if score > 3: # rating lebih dari tiga akan di klasifikasikan sebagai "positif"
        return 'positif'
    elif score < 3: # rating kurang dari tiga akan di klasifikasikan sebagai "negatif"
        return 'negatif'
    else:
        return 'netral'

data['label'] = data['score'].apply(label_data)
    
```


Lampiran 7. Hasil Data Ulasan WhatsApp Setelah Proses Case Folding

bagus kayak itu tapi sayangnya di whatsapp di status foto video itu jadinya buram bngt krn lag ya even ada foto itu tapi tp aja ga bisa semuanya kayak banget pdfi bagus
 apk bagus untuk komunikasi tp mshn di perbaiki bug waktu pembuatan story karena story orang gam ada jern atau mshn pembuatan nya kalo mau luf pembuatan story nya harus di klik dulu lu pan jern nya gak
 sesuai sama pembuatan ai story nya
 aplikasi whatsapp sudah mulai aneh sekarang pengaturan nadanya sudah tidak bisa lepakal saat restart hp ngak tau kenapa lagi selap kta stur mada ehh saat restart hp atau saat kta pindah mode hening dan di
 kembalikan ke mode biasa nadanya sudah kembali lagi seperti mode hening panggilan dan keypad yang hanya bunyi sekali aja sehingga kalau ada panggilan kta ngak bisa denger kalau hp sementara di simpan di
 dalam tas
 hp mau later nih akhirnya di sering banget ada yang bikin virus atau bug yang bikin whatsapp gam jadi bisa biasa baru banget banget ganggu banget deh lag full keadilan gam harus hapus data aplikasi dan login
 ulang dan itu dhar banget hlong dng cai sialat buat masalah ini mungkin bisa ditambah fitur keamanan atau fitur biar pntanmen aneh kayak gini ngak masuk ke whatsapp
 untuk pihak whatsapp tolong perbaiki bug yang ada di whatsapp saya bilang bisa membuat pesan ke orang lain maupun membaca pesan di kemudian lanjutkan untuk mengirim pesan ke pemberi di bawah layar
 yang mengakibatkan keyboard saya tidak bisa muncul karena lag tersebut lohokalah
 gam bisa dibuka dengan nomor saya apakah karena bndidur ya padahal sudah untk whatsapp resmi ini kenapa setelah dapat dng kode malah kembali diarahkan untuk menginstal aplikasi
 hlong dng ini bug nya lag buka status langsung chta sendiri aplikasinya gambarnya malah buram gam karun di update bukannya malah bagus malah makin gam karun baru kali ini kecewa sama wa setelah
 setahun pakai
 malah sering di update diknya malah bguu malah tambah jelek wala keluar sendiri kalo mau bikin foto atau buat story/mem default nya sudah nya sendiri dan akun saya sering sekali buatkan kembali karena apert padahal
 saya gam mshn selain apa hlong di perbaiki lag
 malah kalau bisa saran nya ke lu tempo di buka apk nya bisa memntra helen atau ganggung chat yg atau notifikasi ini apk nya kadang kadang baik atau kadang lain sebagai pengguna yg menggunakan apk ini di
 simpanan nam di rusak buka bany ada notifikasi nya atau pembuatanan pd pesan hp atau jerning nya ada atof kta memuat atau sebagai pengguna wa aga kecewa tpe ada nya pengaturan ini malah gam edh
 di sudah di kasih itu pengaman ya itu tergantung lag pengunat

Lampiran 8. Hasil Tahap Pembobotan TF-IDF

```
# Pembobotan data, TF-IDF
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(data['stemmed_text'])
y = data['label']
```

Lampiran 9. Hasil Proses Tahap Klasifikasi Naïve Bayes

```
Akurasi MultinomialNB: 0.8550561797752809
Presisi MultinomialNB: 0.8702801594636412
Recall MultinomialNB: 0.8450561797752809
F1 Score MultinomialNB: 0.8517522451036028
Matriks Confusion:
[[295  2  0]
 [ 11 269 15]
 [  6  95 197]]
```

	precision	recall	f1-score	support
negatif	0.95	0.99	0.97	297
netral	0.73	0.91	0.81	295
positif	0.93	0.66	0.77	298
accuracy			0.86	890
macro avg	0.87	0.86	0.85	890
weighted avg	0.87	0.86	0.85	890

Lampiran 10. Confusion Matrix Data ulasan WhatsApp

```
Akurasi MultinomialNB: 0.8550561797752809
Presisi MultinomialNB: 0.8702801594636412
Recall MultinomialNB: 0.8550561797752809
F1 Score MultinomialNB: 0.8517522451036028
Matriks Confusion:
[[295  2  0]
 [ 11 269 15]
 [  6  95 197]]
```

Lampiran 11. Source Code

```
import pandas as pd

import nltk

from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split

from sklearn.naive_bayes import MultinomialNB

from sklearn.metrics import accuracy_score,
precision_score, recall_score, f1_score, confusion_matrix,
classification_report

from imblearn.over_sampling import SMOTE

# Download necessary NLTK data
nltk.download('punkt')
nltk.download('stopwords')

# Setup Sastrawi Stemmer
factory = StemmerFactory()
stemmer = factory.create_stemmer()

# Load the validated data
```

```

print("Membaca data dari validasi.csv...")

data = pd.read_csv('validasi.csv')

# Clean the content text

def clean_text(text):

    import re

    from nltk.tokenize import word_tokenize

    from nltk.corpus import stopwords

    text = re.sub(r'\d+', '', text) # Menghapus angka

    text = re.sub(r'[^\\w\\s]', '', text) # Menghapus tanda
    baca

    text = re.sub(r"(@[A-Za-z0-9]+)|([^0-9A-Za-z
    \t])|(\w+:\/\/\S+)|^rt|http.+?", "", text) # Menghapus
    mention dan URL

    text = text.lower() # Mengubah huruf menjadi kecil

    words = word_tokenize(text) # Tokenisasi

    words = [word for word in words if word not in
    stopwords.words('indonesian')] # Menghapus stopwords

    words = [stemmer.stem(word) for word in words] #
    Stemming

    return ' '.join(words)

print("Membersihkan teks konten...")

data['cleaned_text'] = data['content'].apply(clean_text)

print("Teks konten telah dibersihkan.")

# Add fields for positive, negative, and neutral

print("Menambahkan field positif, negatif, netral...")

```



```

data['positif'] = data['label'].apply(lambda x: 1 if x ==
'positif' else 0)

data['negatif'] = data['label'].apply(lambda x: 1 if x ==
'negatif' else 0)

data['netral'] = data['label'].apply(lambda x: 1 if x ==
'netral' else 0)

# Naive Bayes Model for Sentiment Analysis

print("Melatih Model Naive Bayes untuk Analisis
Sentimen...")

X = data['cleaned_text']
y = data['label']

vectorizer = TfidfVectorizer(ngram_range=(1,2)) #
Menggunakan bigrams

X_tfidf = vectorizer.fit_transform(X)

# Balancing the data using SMOTE

smote = SMOTE()

X_resampled, y_resampled = smote.fit_resample(X_tfidf, y)

X_train, X_test, y_train, y_test =
train_test_split(X_resampled, y_resampled, test_size=0.2,
random_state=40)

clf = MultinomialNB()

clf.fit(X_train, y_train)

predicted = clf.predict(X_test)

print("Model berhasil dilatih.")

# Model evaluation

```



```

print("Mengevaluasi model...")

print("Akurasi MultinomialNB: ", accuracy_score(y_test,
predicted))

print("Presisi MultinomialNB: ", precision_score(y_test,
predicted, average='weighted'))

print("Recall MultinomialNB: ", recall_score(y_test,
predicted, average='weighted'))

print("F1 Score MultinomialNB: ", f1_score(y_test,
predicted, average='weighted'))

print(f'Matriks Kebingungan:\n {confusion_matrix(y_test,
predicted)}')

print("=====")

print(classification_report(y_test, predicted,
zero_division=0))

# Save validated data

print("Menyimpan data yang telah divalidasi ke
validasi_updated.csv...")

data.to_csv('validasi_updated.csv', index=False)

print("Data yang telah divalidasi disimpan ke
validasi_updated.csv.")

# Display the first 10 rows of the updated data

print("Menampilkan 10 data pertama setelah ditambahkan
field:")

print(data[['reviewId', 'positif', 'negatif',
'netral']].head(10))

print("Distribusi label setelah validasi:")

print(data['label'].value_counts())

print("Jumlah total data:", len(data))

```



**MAJELIS PENDIDIKAN TINGGI PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH MAKASSAR
UPT PERPUSTAKAAN DAN PENERBITAN**

Alamat kantor: Jl.Sultan Alauddin NO.259 Makassar 90221 Tlp.(0411) 866972,881593, Fax.(0411) 865588

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

SURAT KETERANGAN BEBAS PLAGIAT

**UPT Perpustakaan dan Penerbitan Universitas Muhammadiyah Makassar,
Menerangkan bahwa mahasiswa yang tersebut namanya di bawah ini:**

Nama : Syamrilla Dewi

Nim : 105841102720

Program Studi : Teknik Informatika

Dengan nilai:

No	Bab	Nilai	Ambang Batas
1	Bab 1	10 %	10 %
2	Bab 2	25 %	25 %
3	Bab 3	9 %	10 %
4	Bab 4	7 %	10 %
5	Bab 5	5 %	5 %

Dinyatakan telah lulus cek plagiat yang diadakan oleh UPT- Perpustakaan dan Penerbitan Universitas Muhammadiyah Makassar Menggunakan Aplikasi Turnitin.

Demikian surat keterangan ini diberikan kepada yang bersangkutan untuk dipergunakan seperlunya.

Makassar, 14 Agustus 2024

Mengetahui,

Kepala UPT- Perpustakaan dan Penerbitan,





Bab I SYAMRILLA DEWI 105841102720

by Tahap Tutup

Submission date: 13-Aug-2024 09:47AM (UTC+0700)

Submission ID: 2431332663

File name: BAB_I_Skripsi-Dewi_1_1.docx (24.98K)

Word count: 772

Character count: 5244

ab I SYAMRILLA DEWI 105841102720

ORIGINALITY REPORT

10%

SIMILARITY INDEX

7%

INTERNET SOURCES



6%

PUBLICATIONS

3%

STUDENT PAPERS

PRIMARY SOURCES

1

Romario Onsu, Daniel Febrian Sengkey, Feisy Diane Kambey. "Implementasi Bi-LSTM dengan Ekstraksi Fitur Word2Vec untuk Pengembangan Analisis Sentimen Aplikasi Identitas Kependudukan Digital", Jurnal Teknologi Terpadu, 2024

Publication

3%

2

Submitted to University of Wollongong

Student Paper

2%

3

www.ranktracker.com

Internet Source

2%

4

jurnal.untan.ac.id

Internet Source

2%

5

scholar.unand.ac.id

Internet Source

2%

Exclude quotes Off

Exclude matches < 2%

Exclude bibliography On

Bab II SYAMRILLA DEWI 105841102720

by Tahap Tutup



Submission date: 12-Aug-2024 02:22PM (UTC+0700)

Submission ID: 2430907273

File name: BAB_II_Skripsi-Dewi_1.docx (57.98K)

Word count: 680

Character count: 4420

ab II SYAMRILLA DEWI 105841102720

ORIGINALITY REPORT

25%

SIMILARITY INDEX

19%

INTERNET SOURCES

19%

PUBLICATIONS

7%

STUDENT PAPERS

PRIMARY SOURCES



1	www.whatsapp.com Internet Source	5%
2	digilib.uinsby.ac.id Internet Source	5%
3	publikasi.dinus.ac.id Internet Source	3%
4	repository.uin-suska.ac.id Internet Source	3%
5	www.beritabaik.id Internet Source	2%
6	Dwi Rahma Firmansyah, Endang Lestariningsih. "Analisis Sentimen Ulasan Aplikasi Smart Campus Unisbank di Google Playstore Menggunakan Algoritma Naive Bayes", Jurnal JTIC (Jurnal Teknologi Informasi dan Komunikasi), 2024 Publication	2%
7	Muhammad Guruh Andriawan, Tati Ernawati. "PENGGUNAAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE UNTUK	2%

ANALISIS SENTIMEN KONFLIK PALESTINA
DAN ISRAEL PADA PLATFORM X", Jurnal
Informatika dan Teknik Elektro Terapan, 2024
Publication

8 Linggar Nursingah, Ruuhwan Ruuhwan,
Teuku Mufizar. "ANALISIS SENTIMEN
PENGGUNA APLIKASI X TERHADAP PROGRAM
MAKAN SIANG GRATIS DENGAN METODE
NAÏVE BAYES CLASSIFIER", Jurnal Informatika
dan Teknik Elektro Terapan, 2024
Publication

9 www.detik60.com 2%
Internet Source

Exclude quotes On
Exclude bibliography On

Exclude matches

Bab III SYAMRILLA DEWI 105841102720

by Tahap Tutup



Submission date: 14-Aug-2024 08:49AM (UTC+0700)

Submission ID: 2431759241

File name: BAB_III_skripsi-Dewi_1_1_1.docx (66.39K)

Word count: 1253

Character count: 8387

III SYAMRILLA DEWI 105841102720

PLAGIARISM REPORT

9%

SIMILARITY INDEX

7%

INTERNET SOURCES

2%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

1

rifqimulyawan.com

Internet Source

2%

2

ejournal.itn.ac.id

Internet Source

2%

3

Submitted to Konsorsium Perguruan Tinggi Swasta Indonesia II

Student Paper

2%

4

docplayer.info

Internet Source

2%

Exclude quotes

On

Exclude matches

< 2 >

Exclude bibliography

On





Bab IV SYAMRILLA DEWI
105841102720
by Tahap Tutup

Submission date: 12-Aug-2024 02:22PM (UTC+0700)

Submission ID: 2430907521

File name: BAB_IV_skripsi-Dewi-1.docx (661.68K)

Word count: 6105

Character count: 34861

Lab IV SYAMRILLA DEWI 105841102720

ORIGINALITY REPORT

7%

SIMILARITY INDEX

6%

INTERNET SOURCES

3%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES



1	ejournal.unesa.ac.id Internet Source	1%
2	digilibadmin.unismuh.ac.id Internet Source	1%
3	Submitted to Universitas Negeri Semarang - iTh Student Paper	1%
4	ejournal.akakom.ac.id Internet Source	1%
5	jurnal.undhirabali.ac.id Internet Source	1%
6	dqlab.id Internet Source	<1%
7	jres1.ejournal.unsri.ac.id Internet Source	<1%
8	Leni Kusneti Sukiman, Anggitta Ratu Dolok Saribu. "ANALISIS SENTIMEN ULASAN APLIKASI LINKEDIN DALAM GOOGLE PLAY	<1%

STORE DENGAN MODEL NAÏVE BAYES",
Djtechno: Jurnal Teknologi Informasi, 2023
Publication

9	repository.uin-suska.ac.id Internet Source	<1 %
10	Submitted to Padjadjaran University Student Paper	<1 %
11	Submitted to Sriwijaya University Student Paper	<1 %
12	Submitted to UIN Maulana Malik Ibrahim Malang Student Paper	<1 %
13	ksnugroho.medium.com Internet Source	<1 %
14	Muhammad Dzikry Afandi, Ahmad Homaidi, Abd Ghofur, Ach Zubairi. "Penerapan Information Retrieval dalam Sistem Analisis Kemiripan Proposal Skripsi menggunakan Cosine Similarity", Swabumi, 2024 Publication	<1 %
15	ejournal.itn.ac.id Internet Source	<1 %
16	journal.eng.unila.ac.id Internet Source	<1 %
17	Chris Moulana Bachri, Wawan Gunawan. "Deteksi Email Spam menggunakan Algoritma	<1 %

Convolutional Neural Network (CNN)", Jurnal
Edukasi dan Penelitian Informatika (JEPIN),
2024
Publication

18	Submitted to Loughborough University Student Paper	<1 %
19	play.google.com Internet Source	<1 %

Exclude quotes Off
Exclude bibliography Off

Exclude matches Off



Bab V SYAMRILLA DEWI 105841102720

by Tahap Tutup



Submission date: 13-Aug-2024 09:47AM (UTC+0700)

Submission ID: 2431332874

File name: BAB_V_skripsi-Dewi_1_1_-1.docx (26.59K)

Word count: 1015

Character count: 7032

Ab V SYAMRILLA DEWI 105841102720

ORIGINALITY REPORT

5%

SIMILARITY INDEX

5%

INTERNET SOURCES

6%

PUBLICATIONS

2%

STUDENT PAPERS

PRIMARY SOURCES

1

stmik-budidarma.ac.id
Internet Source



5%



Exclude quotes Off
Exclude bibliography Off

Exclude matches < 2

