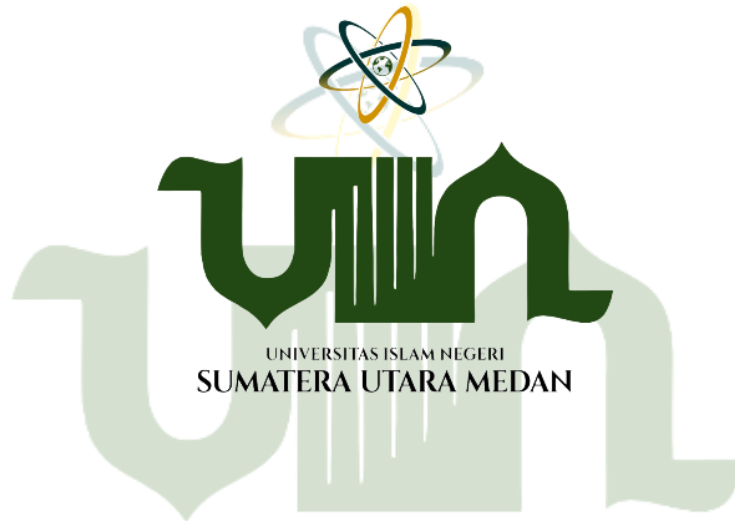


**ANALISIS SENTIMEN TERHADAP TINDAKAN  
PEMERINTAH INDONESIA UNTUK MENAMPUNG  
SEMENTARA PENGUNGSY ETNIS ROHINGYA  
MENGUNAKAN *NAÏVE BAYES CLASIFIER***

**TUGAS AKHIR**



UNIVERSITAS ISLAM NEGERI  
SUMATERA UTARA MEDAN  
**IRWAN GUNAWAN**

**0701201002**

**PRODI ILMU KOMPUTER  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA  
MEDAN**

**2024**

**ANALISIS SENTIMEN TERHADAP TINDAKAN  
PEMERINTAH INDONESIA UNTUK MENAMPUNG  
SEMENTARA PENGUNGSIS ETNIS ROHINGYA  
MENGUNAKAN *NAÏVE BAYES CLASIFIER***

**TUGAS AKHIR**

Diajukan untuk Memenuhi Tugas Akhir dan Melengkapi Syarat Guna  
Memperoleh Gelar Sarjana Strata Satu (S-1) Ilmu Komputer



**IRWAN GUNAWAN**

**0701201002**

**PRODI ILMU KOMPUTER  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA  
MEDAN**

**2024**

## LEMBAR PENGESAHAN

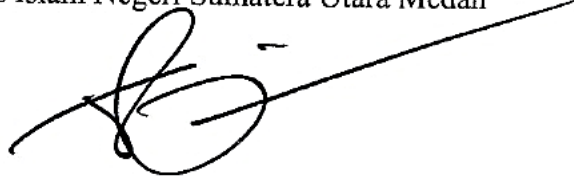
Judul : Analisis Sentimen Terhadap Tindakan Pemerintah  
Indonesia Untuk Menampung Sementara Pengungsi Etnis  
Rohingya Menggunakan *Naïve Bayes Clasifier*  
Penyusun : Irwan Gunawan  
NIM : 0701201002  
Pembimbing : Dr. Mhd Furqan, S.Si, M.Comp, Sc  
Tanggal Seminar : 27 Maret 2024

Disetujui Oleh  
Pembimbing



Dr. Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui,  
Ketua Program Studi Ilmu Komputer  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Sumatera Utara Medan



Ilka Zufria, M.Kom  
NIP. 198506042015031006

## PERSETUJUAN TUGAS AKHIR

Hal : Surat Persetujuan Tugas Akhir

Lamp : -

Kepada Yth.

Dekan Fakultas Sains dan Teknologi

Univeritas Islam Negeri Sumatera Utara Medan

*Assalmu'alaikum Wr. Wb.*

Setelah membaca, meneliti, memberikan beberapa petunjuk dan mengoreksi serta mengadakan perbaikan, saya selaku pembimbing berpendapat Tugas Akhir saudara.

Nama : Irwan Gunawan

Nim : 0701201002

Program Studi : Ilmu Komputer

Judul : Analisis Sentimen Terhadap Tindakan Pemerintah  
Indonesia Untuk Menampung Sementara Pengungsi Etnis  
Rohingya Menggunakan Naïve Bayes Clasifier

dapat disetujui untuk segera *dikolokumkan*. Atas perhatiannya terima kasih.

*Wassalamu'alaikum Wr. Wb*

Medan, 18 Juli 2024 M

11 Muharram 1446 H

Pembimbing



Dr, Mhd Furqan, S.Si, M.Comp, Sc

NIP. 198008062006041003



**KEMENTERIAN AGAMA REPUBLIK INDONESIA  
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA MEDAN  
FAKULTAS SAINS DAN TEKNOLOGI**

Jl. Lapangan Golf, Desa Durian Jangak, Kec. Pancur Batu  
Kab. Deli Serdang, Prov. Sumatera Utara, Kode Pos 20353  
Url: <http://saintek.uinsu.ac.id> , E-mail: [saintek@uinsu.ac.id](mailto:saintek@uinsu.ac.id)

**PENGESAHAN TUGAS AKHIR**

Nomor: B.526/ST/ST.V.2/PP.01.1/08/2024

Judul : Analisis Sentimen Terhadap Tindakan Pemerintah  
Indonesia Untuk Menampung Sementara Pengungsi Etnis  
Rohingya Menggunakan Naïve Bayes Clasifier  
Nama : Irwan Gunawan  
NIM : 0701201002  
Program Studi : Ilmu Komputer  
Fakultas : Sains dan Teknologi

Telah dipertahankan dihadapan Dewan Penguji Tugas Akhir Program Studi Ilmu  
Komputer Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera Medan  
dan dinyatakan **LULUS**.

Pada Hari/Tanggal : Senin / 12 Agustus 2024

Tempat : Ruang Sidang Fakultas Sains dan Teknologi

Tim Ujian Kolokium,  
Ketua,

Ilka Zufria, M.Kom  
NIP. 198506042015031006

Dewan Penguji,

Penguji

Ilka Zufria, M.Kom  
NIP. 198506042015031006

Pembimbing

Dr. Mhd Furqan, S.Si, S.H, M.Comp, Sc  
NIP. 198008062006041003

Mengesahkan,  
Dekan Fakultas Sains dan Teknologi  
UIN Sumatera Utara Medan



Edhum S.H.I., M.Hum  
NIP. 197703012009011008

## SURAT PERNYATAAN KEASLIAN TUGAS AKHIR

Saya yang bertanda tandan dibawah ini

Nama : Irwan Gunawan  
Nomor Induk Mahasiswa : 0701201002  
Program Studi : Ilmu Komputer  
Judul : Analisis Sentimen Terhadap Tindakan  
Pemerintah Indonesia Untuk Menampung  
Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Clasifier

Menyatakan bahwa tugas akhir jurnal ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutka sumbernya. Apabila di kemudia hari ditemukan plagiat dalam tugas akhir jurnal ini maka saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi lainnya sesuai dengan peraturan yang berlaku.

Medan, 05 Juli 2024

  
REPUBLIK INDONESIA  
10000  
METERAI  
TEMPEL  
4243ALX249818012

Irwan Gunawan

NIM. 0701201002

## ABSTRAK

Etnis rohingya merupakan penduduk asli di negara myanmar yang sebagaian besar mayoritasnya beragama muslim. Konflik yang terjadi pada etnis tersebut dimulai sejak ditetapkannya kebijakan Burma Citizen Law oleh pemerintah myanmar. kebijakan ini berisi terkait penolakan pemerintah myanmar terhadap etnis Rohingya sebagai etnis resmi dan memutuskan jika etnis tersebut tidak termasuk dari negara Myanmar. Indonesia merupakan salah satu negara di ASEAN yang masih menampung sementara pengungsi rohingya, tindakan ini dilakukan berdasarkan konsep Human Security dan mengacu pada Peraturan Presiden Republik Indonesia Nomor 125 Tahun 2016 Tentang Penanganan Pengungsi Dari Luar Negeri Pasal 4 Ayat 2 mengenai koordinasi penanganan pengungsi yang meliputi Penemuan, Penampungan, Pengamanan dan Pengawasan. Akibatnya, terjadinya cemburu sosial yang berdampak pada keberagamannya opini masyarakat dan menjadi isu yang sering dibicarakan. Penelitian ini bertujuan untuk mengetahui kecenderungan opini berdasarkan klasifikasi sentimen yang diperoleh melalui video youtube. Manfaat dari penelitian ini adalah agar pemerintah indonesia dapat mengetahui tindakan tersebut cenderung positif atau negatif. Dalam penelitian ini menerapkan algoritma Naïve Bayes Clasifier dengan dataset berjumlah 7547 yang dibagi menjadi 6037 data latih dan 1510 data uji. Hasil Confussion Matrix pada penelitian ini menunjukkan tingkat akurasi yang dihasilkan mencapai 93%

**Kata Kunci:** Analisis Sentimen, Menampung, Pemerintah, Rohingya, Penambangan Teks, Youtube

## ABSTRACT

The Rohingya ethnic group is a native population in the country of Myanmar, where the majority are Muslim. The conflict that occurred in this ethnic group began when the Burma Citizen Law policy was implemented by the Myanmar government. This policy contains the Myanmar government's rejection of the Rohingya ethnicity as an official ethnicity and deciding that this ethnicity is not included in the state of Myanmar. Indonesia is one of the countries in ASEAN which is still temporarily hosting Rohingya refugees. This action is carried out based on the concept of Human Security and refers to Presidential Regulation of the Republic of Indonesia Number 125 of 2016 concerning Handling of Refugees from Abroad Article 4 Paragraph 2 concerning coordination of refugee handling which includes Discovery, Shelter, Security and Supervision. As a result, social jealousy occurs which has an impact on the diversity of public opinion and becomes an issue that is often discussed. To find out this, this research aims to find out opinion trends based on sentiment testimonials obtained through YouTube videos. The benefit of this research is that the Indonesian government can find out whether these actions tend to be positive or negative. In this research, the Naïve Bayes Classifier algorithm was applied with a dataset of 7547 which was divided into 6037 training data and 1510 test data. The results of the Confusion Matrix in this research show that the resulting accuracy level reached 93%

**Keywords:** Sentimen Analysis, Accomodate, Government's, Rohingya, Text Mining, Youtube



UNIVERSITAS ISLAM NEGERI  
SUMATERA UTARA MEDAN

## KATA PENGANTAR

Syukur alhamdulillah kepada Allah Swt atas rahmat-Nya sehingga penulis dapat menyelesaikan Proposal Penelitian Tugas Akhir bentuk Jurnal yang berjudul “Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia Untuk Menampung Sementara Pengungsi Etnis Rohingya Menggunakan *Naïve Bayes Clasifier*”.

Dalam penulisan Proposal penelitian ini, penulis banyak menerima bantuan, bimbingan dukungan dan nasihat dari berbagai pihak. Oleh sebab itu pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada :

1. Ibu Prof. Dr. Nurhayati, M.Ag selaku Rektor UIN Sumatera Utara.
2. Bapak Dr. Zulham, S.H.I, M.Hum selaku Dekan Fakultas Sains dan Teknologi UIN Sumatera Utara.
3. Bapak Ilka Zufria, M.Kom selaku Ketua Prodi Ilmu Komputer Fakultas sains dan Teknologi UIN Sumatera Utara.
4. Bapak Dr. Mhd Furqon, S.Si, M.Comp, Sc selaku Dosen Pembimbing Tugar Akhir Jurnal dan juga sebagai Dosen Penasehat Akademik.
5. Kedua Orang Tua saya, Ayahanda Sriadi,SH dan Ibunda Sriani tercinta yang tiada henti dalam mendukung serta mensupport sepanjang waktu dalam mendoakan disetiap sholatnya.
6. Raisa Azhari dan group ALWI SDN3 selaku sahabat seperjuangan saya selama menempuh perkuliahan yang telah mensupport dan membantu.

Penulis berharap semoga Proposal penelitian ini dapat bermanfaat dan menambah wawasan keilmuan. Kritik dan saran yang sifatnya membangun sangat penulis harapkan untuk perbaikan dimasa yang akan datang.

Medan, 16 Februari 2024

Penyusun

Irwan Gunawan

## DAFTAR ISI

<b>ABSTRAK .....</b>	<b>i</b>
<b>ABSTRACT .....</b>	<b>ii</b>
<b>KATA PENGANTAR.....</b>	<b>iii</b>
<b>DAFTAR ISI.....</b>	<b>iv</b>
<b>DAFTAR GAMBAR.....</b>	<b>vi</b>
<b>DAFTAR TABEL .....</b>	<b>vii</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>viii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian.....	4
1.4 Kontribusi Penelitian .....	4
1.5 Kajian Terdahulu .....	5
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>8</b>
2.1 Etnis Rohingya .....	8
2.2 Analisis Sentimen.....	8
2.3 Klasifikasi.....	9
2.4 Text Mining .....	9
2.5 Text Preprocessing .....	9
2.5.1 Case Folding .....	10
2.5.2 Tokenizing.....	10
2.5.3 Word Normalization.....	10
2.5.4 Stopword Removal.....	10
2.5.5 Stemming .....	11
2.6 Naïve Bayes Clasifier.....	11
2.7 Confussion Matrix.....	11
2.8 Term Frequency-Inverse Document Frequency (TF-IDF).....	12
2.9 Youtube .....	14
2.10Python.....	14

2.11 Google Collaboratory .....	14
2.12 Wordcloud .....	14
2.13 Flowchart.....	15
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>17</b>
3.1 Kerangka Penelitian .....	17
3.2 Rencana Pembahasan .....	19
3.3 Lokasi Penelitian .....	20
3.4 Waktu Penelitian .....	20
3.5 Rencana Penerbitan .....	21
<b>BAB IV HASIL DAN PEMBAHASAN.....</b>	<b>22</b>
4.1 Pembahasan.....	22
4.1.1 Analisis Data .....	22
4.1.2 Representasi Data.....	26
4.1.2.1 Pemberian Label .....	26
4.1.2.2 Preprocessing Data .....	27
4.1.3 Hasil Analisis Data.....	32
4.1.3.1 Pembobotan TF-IDF .....	32
4.1.3.2 Seleksi Kata (Chi Square).....	36
4.1.3.3 Klasifikasi Naïve Bayes Clasifier.....	37
4.2 Hasil .....	39
4.2.1 Pengujian.....	39
4.2.2 Confussion Matrix.....	40
4.2.3 Evaluasi Hasil.....	41
4.2.4 Visualisasi .....	42
4.2.5 Penerapan .....	43
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>44</b>
5.1 Kesimpulan.....	44
5.2 Saran.....	44
<b>DAFTAR PUSTAKA.....</b>	<b>45</b>

## DAFTAR GAMBAR

Gambar	Nama Gambar	Halaman
3.1	Kerangka Penelitian .....	17
3.2	Flowchart Sistem Klasifikasi Analisis Sentimen .....	19
4.1	Konten Video Youtube .....	22
4.2	Google Collaboratory .....	22
4.3	Instalasi Library Googleapiclient .....	23
4.4	Import Library Googleapiclient dan Pandas .....	23
4.5	Secret Key Google API .....	23
4.6	Youtube Data API V3 .....	24
4.7	Dataset yang Berhasil di Crawling .....	24
4.8	Menyimpan Dataset ke Excel .....	25
4.9	<i>Dataset</i> Komentar Youtube .....	25
4.10	Dataset Sudah diberikan Label .....	26
4.11	Persentase Sentimen Positif dan Negatif .....	26
4.12	Hasil Konversi Label .....	27
4.13	Hasil Pembobotan TFIDF .....	36
4.14	Perbandingan Kata yang akan diseleksi .....	36
4.15	1000 Kata yang sudah diseleksi .....	36
4.16	Jumlah Data yang akan di Uji .....	40
4.17	Hasil Confussion Matrix .....	40
4.18	Visualisasi Confussion Matrix .....	41
4.19	<i>Wordcoud</i> pada Sentimen Negatif .....	43
4.20	<i>Wordcoud</i> pada Sentimen Positif .....	43
4.21	<i>Wordcoud</i> pada Sentimen <i>Positif dan Negatif</i> .....	43

## DAFTAR TABEL

<b>Tabel</b>	<b>Nama Tabel</b>	<b>Halaman</b>
2.1	Confussion Matrix.....	12
2.2	Contoh Pembobotan TF-IDF.....	13
2.3	Simbol-Simbol Flowchart .....	15
3.1	Kebutuhan Perangkat .....	18
3.2	Waktu Pelaksanaan Penelitian.....	20
3.3	Rencana Penerbitan Jurnal .....	21
4.1	Proses <i>Case Folding</i> .....	28
4.2	Proses <i>Cleansing</i> .....	28
4.3	Proses <i>Tokenizing</i> .....	29
4.4	Proses World Normalization .....	30
4.5	Proses Stopwords Removal.....	31
4.6	Proses <i>Stemming</i> .....	32
4.7	Contoh Menghitung <i>Term Frequency</i> .....	33
4.8	Contoh Menghitung <i>Document Frequency</i> .....	34
4.9	Contoh Menghitung <i>Inverse Document Frequency</i> .....	34
4.10	Contoh Menghitung TFIDF .....	35
4.11	Contoh Menghitung Naive Bayes Clasifier .....	37
4.12	Hasil <i>Confussion Matrix</i> .....	41

## DAFTAR LAMPIRAN

Lampiran	Nama Lampiran
1.	Logbook
2.	Letter of Accepted (LoA)
3.	Jurnal
4.	Kartu Bimbingan
5.	Daftar Riwayat Hidup



UNIVERSITAS ISLAM NEGERI  
SUMATERA UTARA MEDAN

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Etnis rohingya merupakan penduduk asli di negara Myanmar yang sebgaiian besar mayoritasnya beragama Muslim, konflik yang terjadi pada etnis tersebut dimulai sejak ditetapkan kebijakan *Burma Citizen Law* oleh pemerintah Myanmar pada tahun 1982 (Pramita & Zahidi, 2023), kebijakan tersebut membahas terkait penolakan dari pemerintah Myanmar terhadap keberadaan Rohingya sebagai etnis resmi dan mengklaim bahwasannya etnis tersebut bukan bagian dari negara Myanmar.

Indonesia termasuk salah satu negara di ASEAN (*Asean South East Nations*) yang menampung pengungsi rohingya. Penyebaran etnis tersebut mencapai kurang lebih 1000 orang di indonesia (Rameliah & Sagena, 2023). Selain itu, berdasarkan data dari lembaga UNHCR (*United Nations High Commissioner for Refugees*) bulan maret tahun 2021 memperkirakan total pengungsi rohingya di indonesia mencapai 5% dari 13.497 keseluruhan pengungsi dari berbagai negara yang datang ke indonesia (Adha, 2023).

Penyebab pemerintah indonesia menampung sementara pengungsi rohingya dikarenakan kemanusiaan. Menurut tirza listiarani, Indonesia memiliki kebijakan luar negeri yang terbuka terhadap pengungsi rohingya. dalam perencanaannya dilakukan dengan pendekatan *Human Security* bukan dengan *National Security* (Listiarani, 2020). Selain itu, berdasarkan Peraturan Presiden Republik Indonesia Nomor 125 Tahun 2016 Tentang Penanganan Pengungsi Dari Luar Negeri yang terdapat pada Pasal 4 Ayat 2 mengenai koordinasi penanganan pengungsi yang meliputi Penemuan, Penampungan, Pengamanan dan Pengawasan.

Pada pasal 26 Ayat 5, pemerintah juga memberikan bantuan kepada pengungsi rohingya mencakup penyediaan air bersih, kebutuhan makanan, pakaian, pelayanan kesehatan dan kebersihan serta fasilitas ibadah. Membantu terhadap sesama manusia yang membutuhkan telah dijelaskan di dalam Al-Quran pada Surah Ar-Rum Ayat 38 sebagai berikut :

قَاتِ ذَا الْقُرْبَىٰ حَقَّهُ وَالْمِسْكِينَ وَابْنَ السَّبِيلِ ذَلِكَ خَيْرٌ لِلَّذِينَ يُرِيدُونَ وَجْهَ اللَّهِ وَأُولَٰئِكَ هُمُ الْمُفْلِحُونَ

Artinya “Oleh karena itu, beri kerabat dekat haknya, juga orang miskin, dan orang yang dalam perjalanan. Itulah yang lebih baik bagi orang-orang yang mencari keridaan Allah. Mereka itulah orang-orang yang beruntung” (Al-Qur’an Surah Ar-Rum Ayat 38).

Ayat tersebut menjelaskan mengenai urgensinya memberikan hak kepada kerabat dekat, orang miskin dan seseorang yang sedang dalam perjalanan. Untuk itu, perlu dikaji bahwasannya yang seharusnya diprioritaskan adalah bangsa sendiri. Akibatnya, terjadi cemburu sosial yang berdampak pada beragamnya pandangan dan opini masyarakat yang menjadi isu yang cukup sering di bicarakan di media sosial, Media sosial telah menjadi sarana komunikasi dan interaksi dalam kehidupan masyarakat dunia (Meilani et al., 2024).

*Youtube* menjadi salah satu media yang menyebarkan berita melalui konten video terkait tindakan pemerintah indonesia yang menampung pengungsi rohingya. Menurut Hermawan, sejak tahun 2022 diperkirakan 139 juta masyarakat indonesia menggunakan *youtube* (Hermawan, 2022). Maka dari itu, *youtube* menjadi media yang sering digunakan oleh masyarakat indonesia, serta menghasilkan data yang cukup banyak dan komentar pro dan kontra yang dapat di amati untuk di analisis.

Untuk mengetahui pandangan masyarakat terhadap tindakan pemerintah indonesia yang menampung sementara pengungsi rohingya, diperlukan analisis sentimen yang bertujuan untuk mengetahui kecenderungan opini berdasarkan klasifikasi sentimen melalui komentar video *youtube*. Analisis sentimen adalah pengolahan bahasa natural yang berperan menganalisis pendapat (Furqan et al., 2022), dalam penerapannya analisis sentimen mengacu pada kemampuan perangkat lunak untuk memahami sikap, pendapat, atau pendirian pembicara atau penulis (Furqan & Nasir, 2024). Analisis sentimen juga digunakan untuk menentukan sentimen tentang isu, peristiwa, kepuasan layanan, harga sahan dan analisis persaingan berdasarkan data teks (Setiawan & Zufria, 2023).

Analisis Sentimen yang dimaksud dalam penelitian ini menerapkan algoritma *Naïve Bayes Clasifier* dalam mengklasifikasikan komentar. *Naïve Bayes Clasifier* merupakan hasil probabilitas masa depan yang di ambil dari pengalaman

masa sebelumnya (Adelia et al., 2023). Dan *Term Frequency-Inverse Document Frequency* untuk pembobotan kata (*term*). Menurut Mulyani *Term Frequency-Inverse Document Frequency* termasuk salah satu pembobotan kata terbaik dibandingkan dengan pembobotan kata lainnya (Mulyani & Novita, 2022).

Kajian terkait analisis sentimen menggunakan algoritma *Naïve Bayes Classifier* sebelumnya telah dilakukan oleh beberapa penelitian terdahulu. Seperti pada penelitian berjudul “Analisis Sentimen Tanggapan *Public* Mengenai E-Tilang Melalui Media Sosial *Youtube* Menggunakan Algoritma *Naïve Bayes*” yang dilakukan oleh (Aulia et al., 2023). Berdasarkan hasil penelitiannya opini yang dikumpulkan diambil dari media *Youtube*, karena media tersebut menjadi salah satu objek untuk menyampaikan opini karena bersifat *user friendly* dan terbuka dalam mengakses komentar. Namun, data yang digunakan selama penelitian berlangsung hanya berjumlah 500 komentar dan menghasilkan akurasi sebesar 79.44%. Selanjutnya, penelitian yang berjudul “Analisis Sentimen Terhadap Opini Feminisme Menggunakan Metode *Naïve Bayes*” yang dilakukan oleh (Wahyuni, 2022), berdasarkan hasil penelitiannya digunakan teknik pembobotan kata menggunakan *Term Frequency-Inverse Document Frequency* dengan pola data 80:20 atau 80% data latih dan 20% data uji yang menghasilkan akurasi sebesar 88%. Selanjutnya, penelitian yang berjudul “Analisis Sentimen Masyarakat Terhadap Tindakan Vaksinasi Dalam Upaya Mengatasi Pandemi Covid-19” yang dilakukan oleh (Laurensz & Sedyono, 2021). Berdasarkan hasil penelitiannya pemberian label dilakukan secara manual dan menghasilkan akurasi sebesar 85.59%.

Nantinya, penelitian ini menggunakan data berjumlah 7.547 komentar yang diambil dari *platform youtube* dan diberi label secara manual, sentimen yang akan digunakan hanya positif dan negatif, serta pembobotan kata yang akan diterapkan menggunakan teknik *Term Frequency-Inverse Document Frequency* dengan 70% data *training* dan 30% data *testing*.

Berdasarkan uraian diatas, maka penulis berkeinginan untuk mengangkat topik penelitian yang berjudul “ANALISIS SENTIMEN TERHADAP TINDAKAN PEMERINTAH INDONESIA UNTUK MENAMPUNG SEMENTARA PENGUNGSI ETNIS ROHINGYA MENGGUNAKAN *NAÏVE BAYES CLASIFIER*”

## 1.2 Rumusan Masalah

Berdasarkan penyampaian sebelumnya, maka permasalahan yang akan dibahas pada penelitian ini meliputi:

1. Bagaimana menerapkan atau mengimplementasikan ilmu *machine learning* dengan algoritma klasifikasi *Naïve Bayes Classifier* dalam mengklasifikasikan sentimen terhadap tindakan pemerintah indonesia untuk menampung sementara pengungsi etnis rohingya di media sosial *youtube*.
2. Bagaimana hasil akurasi yang dihasilkan dari model klasifikasi sentimen terhadap tindakan pemerintah indonesia untuk menampung sementara pengungsi etnis rohingya menggunakan metode *naïve bayes clasifier* di media sosial *youtube*.

## 1.3 Tujuan Penelitian

Adapun beberapa tujuan penelitian yang ingin dicapai dari dilaksanakannya penelitian ini adalah untuk :

1. Menerapkan atau mengimplementasikan keilmuan bidang *Machine Learning* dan *Text Mining* menggunakan algoritma klasifikasi *Naive Bayes Classifier* dalam mengklasifikasikan sentimen masyarakat terhadap tindakan pemerintah indonesia untuk menampung sementara pengungsi etnis rohingya di *Youtube*.
2. Menerapan atau mengimplementasikan teknik pembobotan kata *Term Frequency-Inverse Document Frequency (TF-IDF)* dalam klasifikasi sentimen masyarakat terhadap tindakan pemerintah indonesia untuk menampung sementara pengungsi etnis rohingya di *Youtube*.
3. Mengetahui Tingkat akurasi yang dihasilkan dari model klasifikasi sentimen masyarakat terhadap tindakan pemerintah indonesia untuk menampung sementara pengungsi etnis rohingya menggunakan algoritma klasifikasi *Naïve Bayes Classifier*.

## 1.4 Kontribusi Penelitian

Manfaat yang diperoleh dari dilaksanakannya penelitian ini bagi peneliti yaitu Mengimplementasikan ilmu yang sudah didapatkan selama perkuliahan ataupun diluar perkuliahan, mengembangkan ilmu dan kemampuan dalam bidang

penelitian khususnya dalam penelitian terkait analisis sentimen dan *text mining*. Selain itu, Hasil penelitian ini dapat dijadikan pedoman bagi Pemerintah Indonesia untuk mengevaluasi dan mempertimbangkan lebih lanjut terkait tindakan pemerintah untuk menampung sementara pengungsi etnis rohingya, serta dapat memberikan informasi tambahan kepada masyarakat berdasarkan sentimen yang di klasifikasikan dan *wordcloud* yang disajikan dalam penelitian ini.

### 1.5 Kajian Terdahulu

Berikut beberapa penelitian terdahulu yang digunakan pada penelitian ini untuk dapat memecahkan permasalahan yang ada menggunakan metode *Naïve Bayes Clasifier*, diantaranya adalah :

1. Penelitian yang dilakukan oleh (Setiawan & Zufria, 2023) yang berjudul “Analisis Sentimen Pembatalan Indonesia Sebagai Tuan Ruma Piala Dunia FIFA U-20 Menggunakan Naïve Bayes Clasifier”, berdasarkan hasil penelitiannya yang menerapkan algoritma *Naïve Bayes Clasifier* menyimpulkan bahwasannya skor akurasi yang didapatkan sebesar 84% dengan data sebanyak 946 komentar *twitter*.
2. Penelitian yang dilakukan oleh (Mulyani & Novita, 2022) yang berjudul “Implementasi Algoritma *Naïve Bayes Clasifier* Untuk Klasifikasi Sentimen Masyarakat Tentang Depresi Pada *Youtube*”. Berdasarkan hasil penelitiannya menyimpulkan bahwasanya algoritma *Naïve Bayes Clasifier* digunakan untuk mengklasifikasikan sentimen masyarakat pada media sosial *youtube* dengan dataset sebanyak 3784 data yang di *crawling* menggunakan *python* dan dipisah menjadi 3826 data latih dan 957 data uji. Pada proses pembobotan kata dilakukan dengan teknik *Term Frequency-Inverse Document Frequency* (TF-IDF) dan dilakukan pemberian label secara manual. Adapun hasil penelitiannya menghasilkan tingkat akurasi sebesar 84.11%
3. Penelitian yang dilakukan oleh (Naraswati et al., 2021) yang berjudul “Analisis Sentimen Publik Dari *Twitter* Tentang Kebijakan Penanganan Covid-19 Di Indonesia Dengan *Naive Bayes Clasification*”. Berdasarkan hasil penelitiannya menyimpulkan bahwasannya algoritma *Naïve Bayes Clasifier* digunakan untuk mengklasifikasikan sentimen masyarakat pada media sosial *twitter* dengan

- dataset sebanyak 4292 data dan dipisah menjadi 1864 data latih dan 466 data uji. Adapun hasil penelitiannya menghasilkan tingkat akurasi sebesar 87,34%
4. Penelitian yang dilakukan oleh (Santoso et al., 2022) yang berjudul “Analisis Sentimen Mahasiswa Terkait Pembelajaran Tatap Muka Menggunakan Metode *Naïve Bayes Clasifier*”. Berdasarkan hasil penelitiannya dilakukan klasifikasi sentimen dengan algoritma *Naïve Bayes Clasifier* terhadap 5350 data sentimen yang didapat berdasarkan kuisisioner (angket) dari 1070 responden dari berbagai fakultas di UIN-SU Medan. Pemberian label dilakukan dengan cara manual karena menghasilkan nilai presisi yang cukup tinggi, selain itu pembobotan kata yang digunakan menerapkan teknik *Term Frequency-Inverser Document Frequency* dan menghasilkan akurasi rata-rata sebesar 83%.
  5. Penelitian yang dilakukan oleh (Hertati et al., 2023) yang berjudul “Analisis Sentimen Masyarakat Terhadap Kenaikan Biaya Haji Tahun 2023 Menggunakan Metode *Naïve Bayes Clasifier*”. Berdasarkan hasil penelitiannya dilakukan klasifikasi sentimen masyarakat terhadap 3000 data komentar pengguna *twitter*, adapun pelabelan sentimen dilakukan secara manual oleh Guru Bahasa Indonesia, dan untuk Pembobotan Kata digunakan teknik *Term Frequency-Inverser Document Frequency* (TF-IDF), adapun data yang akan di modelkan dibagi menjadi beberapa percobaan pembagian data yaitu dengan pola 70:30, 80:20, dan 90:10 dan dengan percobaan 70:30 mendapatkan akurasi sebesar 81,46%.
  6. Penelitian yang dilakukan oleh (Ghozali et al., 2023) yang berjudul “Anslisis Sentimen Pinjam *Online* Di Media Sosial *Twitter* Menggunakan Metode *Naïve Bayes*”. Berdasar kan hasil penelitiannya yang menerapkan algoritma *Naïve Bayes Clasfieir* pada 500 data yang telah di *crawling* yang terbagi menjadi 400 data latih dan 100 data uji, adapun pembobotan dilakukan menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) dan mendapatkan hasil akurasi sebesar 80%.
  7. Penelitan yang dilakukan oleh (Furqan et al., 2022) yang berjudul “Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia”. Berdasarkan penelitian yang telah dilakukan dalam

proses pembobotan kata (*term*) menggunakan teknik *Term Frequency-Inverse Document Frequency* pada 100 komentar dan menghasilkan akurasi 100%.

8. Penelitian yang dilakukan oleh (Kurniasih & Suseno, 2022) yang berjudul “Analisis Sentimen Terhadap Bantuan Subsidi Upah (BSU) pada Kenaikan Harga Bahan Bakar BBM”. Berdasarkan hasil penelitiannya dalam menerapkan *Term Frequency-Inverse Document Frequency* (TF-IDF) mendapatkan tingkat akurasi sebesar 92,89%.
9. Penelitian yang dilakukan oleh (Firsttama et al., 204 C.E.) yang berjudul “Analisis Sentimen Komentar *Youtube* Konferesi Tingkat Tinggi G20 Menggunakan Metode *Naïve Bayes*”. Berdasarkan penelitian yang telah dilaksanakan. Algoritma *Naïve Bayes Clasifier* diterapkan untuk mengklasifikasikan sentimen berdasarkan 19.215 data komentar yang diumpulkan dari platform *Youtube*, berdasarkan sentimennya 51.5% komentar berlabel positif dan 48.5% berlabel negatif, dengan pola pembagian data 90%:10% dan meghasilkan akurasi 76%.
10. Penelitian yang dilakukan oleh (Adelia et al., 2023) yang berjudul “Analisis Sentimen Belajar Programming Pada Media Sosial *Youtube* Menggunakan Algoritma Klasifikasi *Naïve Bayes*”. Berdasarkan penelitian yang telah dilaksanakan, data di peroleh dengan teknik *crawling* pada media sosial *Youtube*. Karena youtube menjadi salah satu jaringan sosial yang sering digunakan untuk berbagai kegiatan. Selain itu, setiap kata dilakukan pembobotan dengan teknik *Term Frequency-Inverse Document Frequency* dan menghasilkan akurasi sebesar 100%.
11. Penelitian yang dilakukan oleh (Yasir & Suraji, 2023) yang berjudul “Perbandingan Metode Klasifikasi *Naïve Bayes Clasifier*, *Decision Tree*, *Random Forest* Terhadap Analisis Sentimen Kenaikan Biaya Haji 2023 Pada Media Sosial *Youtube*”. Berdasarkan hasil penelitiannya yang membandingkan metode *Naïve Bayes Clasifier* dengan *Decision Tree* dan *Random Forset* pada 1.013 teks komentar. Adapun dari ketiga perbandingan metode klasifiaksi tersebut menunjukkan bahwasannya akurasi yang tinggi dihasilkan oleh algoritma *Naïve Bayes Clasifier* sebesar 90%.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Etnis Rohingya**

Etnis rohingya merupakan salah satu etnis dari sejumlah etnis yang berada di negara Myanmar, penduduk Myanmar yang beragama islam sebagian besar berasal dari etnis rohingya dan etnis rohingya termasuk etnis yang menganut agama islam dinegara tersebut (Indradipradana & Haridha, 2023). Keberadaan etnis rohingya di negara Myanmar merupakan minoritas islam dan minoritas yang terisolasi di Myanmar, rohingya adalah bangsa yang terbentuk selama peristiwa yang panjang. Mereka ada di wilayah Rakhine sebelum inggris berkuasa di Rakhine dan Burma antara tahun 1824 dan 1948 (Hamdi et al., 2023). Pelanggaran dan tindakan kejahatan kemanusiaan yang dilakukan oleh pemerintah Myanmar terhadap etnis rohingya berawal dari di teapkannya Undang-undag status kewarganegaraan tahun 1962. Peraturan tersebut menghapus status kewarganegaraan kelompok rohingya sehingga sebagian dari mereka kehilangan hak-hak dan perlindungannya dari negara tersebut (Mahadevi et al., 2022). Selain itu penindasan dan pembunuhan yang diterima oleh masyarakat dari etnis rohingya ini membuat etnis tersebut mendapatkan dua pilihan yaitu bertahan di Myanmar atau mengungsi untuk menjadi pengungsi di negara lain (Prabowo et al., 2022). Etnis rohingya di pandang sebagai ilegal bengali karena etnis ini dinilai sebagai imigran ilegal yang memiliki karakteristik yang sama seperti budaya dan fisik seperti orang bangladesh (Indradipradana & Haridha, 2023)

#### **2.2 Analisis Sentimen**

Analisis sentimen merupakan penambangan teks yang bertujuan untuk menganalisis dan mengekstraksi suatu informasi yang bersifat subjektif yang dapat membantu pihak yang membutuhkannya (Normawati & Prayogi, 2021). Analisis sentimen juga digunakan untuk mengekstrak data opini, memahami serta mengolah tekstual data secara otomatis untuk melihat sentimen yang terkandung di sebuah opini. Sentimen yang di kategorikan terbagi menjadi sentimen positif, netral dan

negatif (Sari & Wibowo, 2019). Umumnya, analisis sentimen diterapkan untuk melakukan proses identifikasi terhadap suatu opini masyarakat atau pendapat perorangan terhadap suatu objek yang diungkapkan atau diutarakan melalui bentuk teks. Dalam analisis sentimen, data diambil dari komentar, ulasan atau testimoni dari produk yang akan di teliti dengan tujuan untuk mengetahui secara umum opini masyarakat terhadap objek yang di analisis (Normawati & Prayogi, 2021). Selain itu tujuan dari dilakukannya analisis sentimen adalah untuk mengidentifikasi kecenderungan opini pada suatu masalah, yang menghasilkan opini tersebut cenderung positif atau negatif (Sumanjaya et al., 2022)

### **2.3 Klasifikasi**

Klasifikasi merupakan suatu cara untuk mengelompokkan suatu objek berdasarkan ciri-ciri yang dimiliki. Dalam prosesnya, klasifikasi dapat dilakukan dengan berbagai cara baik itu manual ataupun dibantu dengan penggunaan teknologi. Klasifikasi yang dilakukan oleh manusia dilakukan tanpa adanya bantuan teknologi sedangkan yang dilakukan oleh komputer dengan menggunakan bantuan algoritma cerdas (Supiyandi et al., 2023). Dalam klasifikasi terdiri dari dua tahap yaitu pembelajaran dan tahap pengklasifikasian, pada tahap pembelajaran sebuah algoritma klasifikasi akan membangun sebuah model klasifikasi dengan cara menganalisis data latih (Susana et al., 2022).

### **2.4 Text Mining**

*Text Mining* merupakan proses penambangan yang dilakukan oleh komputer untuk mendapatkan sesuatu yang baru, dan tidak diketahui sebelumnya atau menemukan kembali informasi-informasi yang tersirat secara implisit, temuan tersebut berasal dari informasi yang di ekstrak secara otomatis dari sumber-sumber data teks yang berbeda (Sari & Wibowo, 2019)

### **2.5 Text Preprocessing**

*Text Preprocessing* merupakan langkah atau tahapan awal pada bidang *text mining* yang bertujuan untuk mengolah data dari data yang sebelumnya tidak terstruktur menjadi lebih terstruktur sesuai dengan kebutuhan, hal ini dilakukan untuk mempersiapkan data yang bersih sebelum digunakan pada proses

selanjutnya (Sumanjaya et al., 2022). *Text preprocessing* merupakan bagian dari *text mining* yang dilakukan untuk menghapus *noise* pada kalimat, selain itu bertujuan untuk menghindari data yang kurang sempurna, gangguan pada data dan data yang tidak konsisten (Sari & Wibowo, 2019). Tahapan yang dilakukan dalam *text preprocessing* meliputi *Case Folding*, *Tokenizing*, *Normalization*, *Filtering* dan *Stemming* (Fahlevvi, 2022).

### **2.5.1 Case Folding**

*Case Folding* merupakan tahapan normalisasi menjadi huruf kecil (Adelia et al., 2023). Dalam tahapan ini akan diubah semua huruf dalam dokumen menjadi bentuk *lowercase* atau huruf kecil, adapun contoh dari diterapkannya tahap ini adalah kalimat "Kemarin" diubah menjadi "kemarin" (Sumanjaya et al., 2022).

### **2.5.2 Tokenizing**

*Tokenizing* merupakan tahapan pemisahan setiap kata dalam sebuah kalimat menjadi satu kesatuan kata (Adelia et al., 2023). Tujuannya adalah untuk memisahkan kata-kata menjadi bentuk yang lebih kecil, contohnya terhadap kalimat "kita juga pisah" akan di *tokenizing* menjadi satu kesatuan seperti "kita" "juga" "pisah" (Sumanjaya et al., 2022).

### **2.5.3 Word Normalization**

*Normalize* merupakan tahapan yang memiliki fungsi untuk mengurangi kata atau mengubah kata yang tidak baku menjadi baku (Adelia et al., 2023). Tujuannya adalah untuk mengubah kata gaul menjadi kata yang baku sesuai dengan Kamus Besar Bahasa Indonesia (KBBI). Contohnya adalah kalimat "dgn" yang diubah menjadi "dengan" (Sumanjaya et al., 2022).

### **2.5.4 Stopword Removal**

*Stopword Removal* merupakan tahapan yang memiliki tujuan untuk menghilangkan atau menghapus kata yang tidak memiliki arti pada suatu kalimat (Santoso et al., 2022). Tujuan dari dilakukannya tahap ini adalah untuk mengurangi waktu pemrosesan, ukuran indeks serta mengurangi *level noise*.

Seperti penggunaan kata hubug seperti “di”, “ke”, “dan”, “dari” dan lainnya (Sumanjaya et al., 2022).

### 2.5.5 Stemming

*Stemming* merupakan tahapan yang digunakan untuk merubah kata yang memiliki imbuhan menjadi kata dasar (Adelia et al., 2023). Tujuan dari dilakukannya tahap ini adalah untuk mengubah kata kedalam bentuk dasar yang di khususkan untuk pemrosesan teks bahasa indonesia, contohnya adalah kalimat “makanan” menjadi “makan” (Sumanjaya et al., 2022).

### 2.6 Naïve Bayes Clasifier

*Naïve Bayes Clasifier* merupakan sebuah metode pengklasifikasian dengan menggunakan probabilitas sederhana yang berakar pada *Teorema Bayes* dan memiliki asumsi ketidaktergantungan (*independent*) yang tinggi dari masing-masing kondisi atau kejadian (Sari & Wibowo, 2019). metode *Naïve Bayes Clasifier* tergolong kedalam *Supervised Learning*, karena membutuhkan data latih untuk mengelompokkan data kedalam sebuah kelas (Sumanjaya et al., 2022). Selain itu metode ini digunakan untuk memprediksi probabilitas keanggotaan suatu kelas (Supiyandi et al., 2023).

Rumus perhitungan algoritma *Naïve Bayes Clasifier* dalam (Naraswati et al., 2021)dijabarkan dalam persamaan (1) :

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Keterangan persamaan :

X = Data dengan kelas yang belum diketahui

H = Hipotesis data yang merupakan suatu kelas spesifik

$P(H|X)$  = Probabilitas hipotesis H berdasarkan kondisi X (probabilitas posterior)

$P(H)$  = Probabilitas hipotesis H (probabilitas prior)

$P(X|H)$  = Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$  = Probabilitas

### 2.7 Confussion Matrix

*Confussion Matrix* merupakan tabel yang digunakan untuk menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah (Normawati

& Prayogi, 2021). Adapun tabel *confussion matrix* yang dimaksud dalam (Normawati & Prayogi, 2021) disajikan pada tabel 2.1.

**Tabel 2.1** *Confussion Matrix*

		Kelas Prediksi	
		1	0
Kelas sebenarnya	1	TP	FN
	0	FP	TN

Keterangan tabel

TP (*True Positif*) = Jumlah dokumen dari kelas 1 yang benar diklasifikasikan sebagai kelas 1

TN (*True Negatif*) = jumlah dokumen dari kelas 0 yang benar diklasifikasikan sebagai kelas 0

FP (*False Positif*) = jumlah dokumen dari kelas 0 yang salah diklasifikasikan sebagai kelas 1

FN (*False Negatif*) = Jumlah dokumen dari kelas 1 yang salah diklasifikasikan sebagai kelas 0

Selanjutnya, berikut rumus *confussion matrix* untuk mendapatkan nilai *accuracy*, *precision*, *f1 score* dan *recall* dalam (Hertati et al., 2023) disajikan pada persamaan (2) (3) (4) dan (5):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (4)$$

$$F1\ Score = 2 \times \frac{precision \times recall}{precision+recall} \quad (5)$$

## 2.8 Term Frequency-Inverse Document Frequency (TF-IDF)

*Term Frequency-Inverse Document Frequency* merupakan salah satu metode untuk melakukan pembobotan kata dari proses ekstraksi kata dengan menerapkan perhitungan kata umum di *information retrieval*, metode pembobotan ini merupakan gabungan antara *term frequency* dan *inverse document frequency*, yang mana *term frequency* merupakan jumlah kemunculan dari sebuah kata (*term*)

pada sebuah dokumen. Sedangkan *inverse document frequency* adalah proses untuk mengukur seberapa penting kata dalam suatu dokumen (Fahlevvi, 2022).

Rumus perhitungan pembobotan *Term Frequency-Inverse Document Frequency* dalam (Santoso et al., 2022) disajikan dalam persamaan (6) dan (7):

Persamaan untuk menghitung TF (*Term Frequency*) :

$$TF = \frac{f_{t,d}}{\sum_{t \in d} f_{t,d}} \quad (6)$$

Keterangan persamaan :

TF = *Term Frequency*

$d$  = dokumen

$t$  = *term* (kata/token)

$f_{t,d}$  = jumlah *term* (kata/token) pada setiap dokumen

$\sum_{t \in d} f_{t,d}$  = jumlah dokumen yang memuat  $t$

Persamaan untuk menghitung IDF (*Inverse Document Frequency*)

$$idf_t = \log \frac{N}{dft} \quad (7)$$

Keterangan persamaan :

$idf_t$  = Nilai *idf* dari *term* (token)  $t$

$N$  = Banyaknya dokumen yang tersedia

$dft$  = Intensitas kemunculan kata pada dokumen

Berikut contoh penerapan *Term Frequency-Inverse Document Frequency* pada kalimat-kalimat berikut ini :

1. Rohingya harus dibantu
2. Indonesia bebas imigran
3. Utamakan hak asasi manusia

**Tabel 2.2** Contoh Pembobotan TF-IDF

Kata ( <i>Term</i> )	TF		
	D1	D2	D3
Rohingya	1	0	0
Harus	1	0	0
Bantu	1	0	0
Indonesia	0	1	0

Bebas	0	1	0
Imigran	0	1	0
utamakan	0	0	1
hak	0	0	1
Asasi	0	0	1
Manusia	0	0	1

## 2.9 Youtube

*Youtube* merupakan media sosial yang digunakan untuk mengunggah video, menonton berbagai video dan dapat berbagi video dimanapun dan kepada siapapun oleh semua orang (Tinambunan & Siahaan, 2022). Media sosial merupakan salah satu media *online* yang mendukung interaksi sosial manusia. Sosial media menggunakan teknologi berbasis web yang mengubah komunikasi menjadi dialog interaktif (Hendra & Nurdin, 2020). semua orang dipermudah dalam pencarian berita karena tidak perlu lagi keluar rumah membeli koran, Selain itu *youtube* juga dikenal sebagai hiburan karena ada banyak jenis konten video yang menghibur (Tinambunan & Siahaan, 2022).

## 2.10 Python

*Python* merupakan suatu bahasa pemrograman tingkat tinggi yang dapat melakukan eksekusi beberapa perintah atau intruksi multi guna secara langsung atau interpretatif (Oktavia et al., 2023).

## 2.11 Google Collaboratory

*Google Colab* merupakan sebuah IDE pemrograman *Python* dimana pemrosesan dilakukan di server *Google* yang memiliki perangkat keras performa tinggi, dan menyediakan pustaka (*library*) yang dibutuhkan (Guntara, 2023).




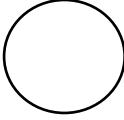
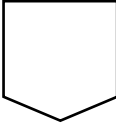
## 2.12 Wordcloud

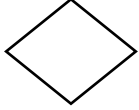



*Word Cloud* merupakan visualisasi data teks yang akan menampilkan kata-kata dari kumpulan data teks di suatu model yang sudah dikembangkan. Visualisasi tersebut menjelaskan bahwasannya semakin besar ukuran suatu kata, maka kata-kata tersebut semakin sering kemunculannya dan lebih banyak muncul didalam suatu data teks (Yasir & Suraji, 2023)

### 2.13 Flowchart

*Flowchart* atau bagan alir merupakan sebuah gambaran dari sistem yang akan dibuat dengan memperlihatkan urutan dengan hubungan proses beserta intruksinya, tujuannya adalah agar urutan proses lebih jelas (Nugraha et al., 2022). *flowchart* memiliki bagan-bagam yang mempunyai arus yang menggambarkan langkah-langkah dalam penyelesaian suatu masalah (Putra & Faiza, 2022). Penggunaan *flowchart* dapat merepresentasikan grafis yang menggambarkan aliran data dan pertukaran informasi yang diterapkan sebagai data yang mengalir dari *input* ke *output* (Semendawai et al., n.d.). *Flowchart* yang dimaksud disajikan pada Tabel 2.1 sebagai berikut.

**Tabel 2.3** Simbol-Simbol *Flowchart*

No	Simbol	Nama	Keterangan
1		Input atau Otput	Digunakan untuk mewakili data masukan dan keluaran
2		Proses	Digunakan untuk mewakili suatu proses
3		Garis Alir	Digunakan untuk menunjukkan arus atau arah dari sebuah proses
4		Penghubung	Digunakan untuk menunjukkan sambungan dari bagian alir yang terputus di halaman lain
5		Penghubung	Digunakan untuk menunjukkan sambungan dari bagian alir yang terputus di halaman lainnya

6		Keputusan	Digunakan untuk penyeleksian kondisi di dalam program
7		Proses Terdefinisi	Digunakan untuk menunjukkan suatu operasi yang rincihannya ditunjukkan di tempat lain
8		Persiapan	Digunakan untuk memberi nilai awal suatu besaran
9		Titik Terminal	Digunakan untuk menunjukkan awal dan akhir dari suatu proses



UNIVERSITAS ISLAM NEGERI  
SUMATERA UTARA MEDAN



b. Studi Lapangan (*Field Research*)

Berbeda dengan proses pengumpulan data dilapangan pada umumnya, dalam penelitian analisis sentimen ini, peneliti mengumpulkan data komentar di sosial media *youtube*. teknik *crawling* dilakukan dalam pengambilan data dan peneliti juga menggunakan layanan *Google API Youtube V3* untuk dapat mengakses ke *database youtube*.

### 3. Analisis Kebutuhan

Analisis kebutuhan data bertujuan untuk mengetahui kategori data yang akan digunakan dalam proses analisis. Data yang diperlukan dalam penelitian ini dibagi menjadi 2 kelompok sentimen yaitu Positif dan Negatif. Pada penelitian ini peneliti menggunakan kombinasi pembagian data dengan pola 70% data *training* dan 30% data *testing* dari 7547 total data sentimen yang dimiliki. Pada data *training* diberi label secara manual sesuai dengan sentimen yang diperoleh. selanjutnya analisis kebutuhan perangkat yaitu hal-hal yang diperlukan untuk merancang sistem berupa perangkat keras dan perangkat lunak yang disajikan dalam tabel 3.1.

**Tabel 3.1** Kebutuhan Perangkat

Perangkat Keras	Perangkat Lunak
Laptop Toshiba Sattelite L745	Windows 7 Ultimate 64 Bit
Intel Core I3-2350M 2.30 Ghz	Visual Studio Code versi 1.70.3
SSD 120 GB	Python versi 3.8.10 dan Jupyter versi 1.0.0
Memory 6 GB	Google Chrome versi 109.0.5414.120

### 4. Perancangan Sistem

Pada tahap ini peneliti akan memberikan gambaran mengenai sistem yang akan dibuat. Perancangan sistem dalam penelitian ini menggunakan diagram alir (*flowchart*) yang bertujuan untuk menggambarkan secara akurat semua langkah yang terlibat dalam proses pembuatan sistem klasifikasi sentimen.

### 5. Implementasi *Naïve Bayes Classifier*

Pada tahap ini peneliti melaksanakan tahapan perancangan yaitu membuat model klasifikasi dengan menerapkan algoritma *Naïve Bayes Classifier*, tahap ini menghasilkan model yang diperoleh berdasarkan 70% data latih.

## 6. Pengujian Sistem

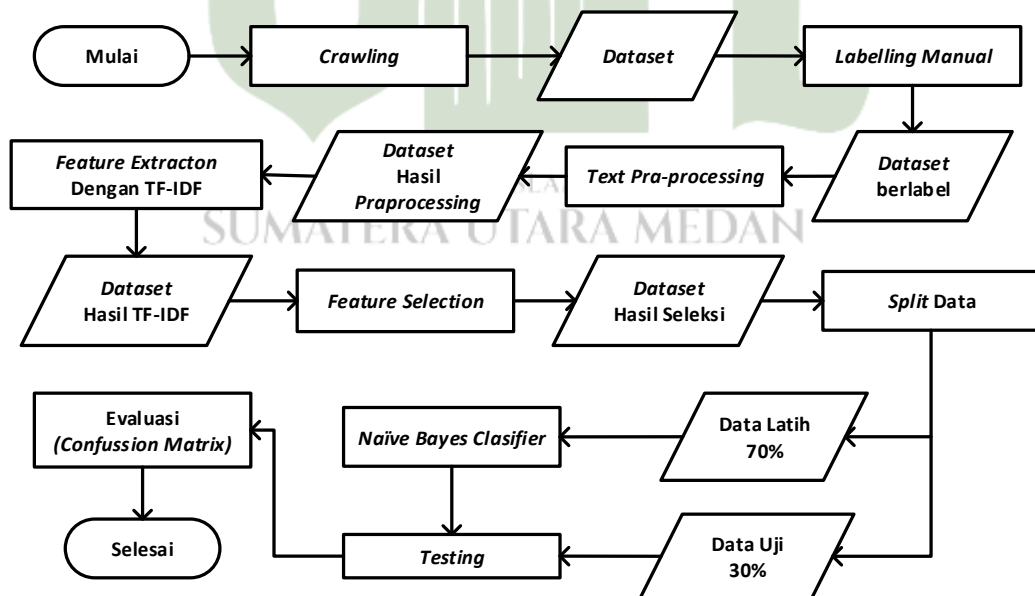
Pada tahap ini peneliti memastikan bahwa sistem klasifikasi sentimen telah berjalan sebagaimana mestinya. Proses ini dilakukan dengan menggunakan data uji (*testing*) yang sebelumnya telah melewati *preprocessing*, *feature extraction* dan *feature selection*. lalu diuji dengan mengklasifikasikannya ke model yang telah dikembangkan pada tahapan implementasi. Hasil pengujiannya akan menghasilkan nilai *accuracy*, *precision*, *recall*, *f1* dan *score* dalam *confusion matrix*.

## 7. Evaluasi Sistem

Pada tahap ini peneliti akan mengevaluasi sistem klasifikasi sentimen yang telah diuji berdasarkan hasil *confusion matrix* yang dihasilkan, peneliti melakukan perhitungan manual *Confussion Matrix* untuk memperoleh evaluasi yang akurat.

### 3.2 Rencana Pembahasan

Dalam penelitian ini akan merancang sistem analisis sentimen yang dapat melakukan klasifikasi sentimen yang dimodelkan menggunakan algoritma *Naïve Bayes Clasifier*. adapun sistem analisis sentimen yang dimaksud disajikan dalam *flowchart* sistem pada Gambar 3.2.



**Gambar 3.2** Flowchart Sistem Klasifikasi Analisis Sentimen

Proses diawali dengan memasukan *dataset* yang telah crawling dan diberikan label secara manual, selanjutnya dilakukan pembersihan *teks* pada tahap

*text preprocessing*. lalu akan di ekstrak dengan dilakukan pembobotan kata menggunakan teknik *Term Frequency-Inverse Document Frequency (TF-IDF)*, hasil dari pembobotan kata menggunakan TF-IDF akan di seleksi pada tahap *feature selection*. Lalu akan di modellkan menggunakan algoritma *Naïve Bayes Clasifier* yang sebelumnya telah dilakukan pemisahan data latih sebesar 70% dan data uji sebesar 30%, dan model yang telah dihasilkan akan diuji menggunakan data uji dan diukur akurasi klasifikasinya menggunakan *Confusion Matrix*.

### 3.3 Lokasi Penelitian

Peneliti menempatkan lokasi penelitian di Laboratorium Ilmu Komputer Fakultas Sains dan Teknnologi UIN Sumatera Utara yang berlokasi di Jalan Lapangan Golf, Desa Durin Jangak, Kecamatan Pancur Batu, Kota Medan.

### 3.4 Waktu Penelitian

Waktu pelaksanaan penelitian pada penelitian ini disajikan pada Tabel 3.2.

**Tabel 3.2** Waktu Pelaksanaan Penelitian

No	Kegiatan	Tahun 2024																			
		Januari				Februari				Maret				April				Mei			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Identifikasi Masalah	■	■																		
2	Studi Pustaka			■	■																
3	Pengumpulan Data					■	■														
4	Analisis Kebutuhan							■	■												
5	Perancangan									■	■										
6	Implementasi											■	■								
7	Pengujian dan Evaluasi													■	■						
8	Submit Jurnal ke Penerbit															■	■				
9	Review dan Revisi Jurnal																	■	■	■	■
10	Publish Jurnal																			■	■

### 3.5 Rencana Penerbitan

Adapun rencana penerbitan jurnal yang akan ditargetkan oleh peneliti dalam melaksanakan penelitian ini disajikan pada Tabel 3.3.

**Tabel 3.3** Rencana Penerbitan Jurnal

<b>Penerbit 1</b>	Nama Penerbit	Jurnal Ilmu Komputer dan Agri-Informatika (JIKA)
	Akreditasi Penerbit	Sinta 3
	Website Penerbit	<a href="https://journal.ipb.ac.id/index.php/jika/index">https://journal.ipb.ac.id/index.php/jika/index</a>
	Frekuensi Terbitan	Biannual (Mei dan November)
	Biaya	Rp. 750.000
<b>Penerbit 2</b>	Nama Penerbit	Jurnal Informatika : Jurnal Pengembangan IT
	Akreditasi Penerbit	Sinta 3
	Website Penerbit	<a href="https://ejournal.uin-suska.ac.id/index.php/coreit/index">https://ejournal.uin-suska.ac.id/index.php/coreit/index</a>
	Frekuensi Terbitan	Januari, Mei, September
	Biaya	Rp. 500, 000
<b>Penerbit 3</b>	Nama Penerbit	Jurnal CoreIT
	Akreditasi Penerbit	Sinta 3
	Website Penerbit	<a href="https://ejournal.uin-suska.ac.id/index.php/coreit/index">https://ejournal.uin-suska.ac.id/index.php/coreit/index</a>
	Frekuensi Terbitan	Juni dan Desember
	Biaya	Gratis
<b>Penerbit 4</b>	Nama Penerbit	Jurnal Teknologi dan Manajemen Informatika
	Akreditasi Penerbit	Sint 3
	Website Penerbit	<a href="https://jurnal.unmer.ac.id/index.php/jtmi/index">https://jurnal.unmer.ac.id/index.php/jtmi/index</a>
	Frekuensi Terbitan	Juni dan Desember
	Biaya	Rp. 1.000.000

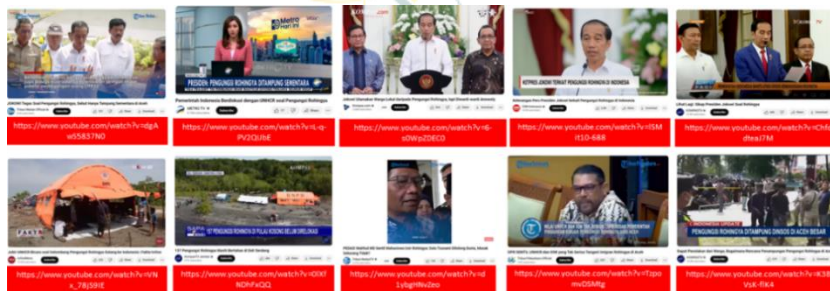
## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1 Pembahasan

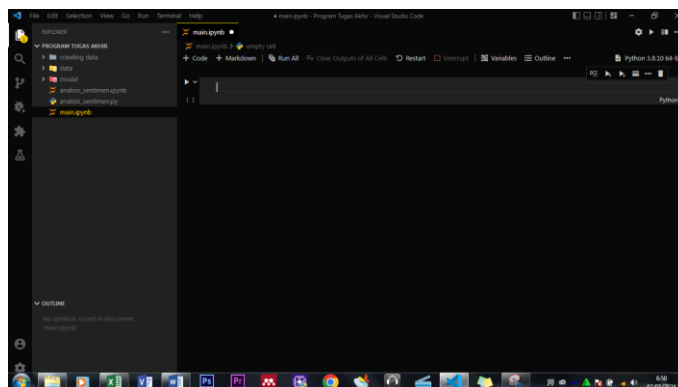
##### 4.1.1 Analisis Data

Dalam proses pengumpulan data, komentar yang digunakan dalam penelitian ini di ambil dengan teknik *crawling* pada *platform youtube*, peneliti mulai mencari beberapa video *youtube* terkait topik penelitian ini, dan peneliti mengumpulkan kurang lebih 10 video yang akan diambil komentarnya untuk di analisis. Masing-masing video memiliki *video\_id* yang tersedia didalam *link* video tersebut. Nantinya, atribut tersebut akan digunakan sebagai parameter saat proses *crawling*, 10 video yang dimaksud dapat dilihat pada Gambar 4.1



Gambar 4.1 Konten Video Youtube

Peneliti menggunakan *google collaboratory* yang di *install* di dalam *Visual Studio Code* untuk menjalankan program *python*, *tools* ini efektif karena tidak memakan *memory* dan tidak perlu koneksi jaringan internet. *Google collaboratory* yang di *install* didalam *Visual Studio Code* dapat dilihat pada Gambar 4.2.



Gambar 4.2 Google Collaboratory

Peneliti juga menggunakan bahasa pemrograman *python*, bahasa ini sangat baik digunakan dalam penambangan data. Karena bahasa *python* menyediakan beberapa *library* yang cukup banyak, serta dalam penulisan kode programnya *python* sangat sederhana dalam penggunaannya. Selain itu, peneliti menggunakan *Library pandas* dan *googleapiclient*, dimana *library googleapiclient* digunakan untuk menghubungkan program *python* yang telah dibuat agar terhubung ke *server youtube*, dan *library pandas* digunakan untuk menyimpan data yang telah di *crawling* ke format *excel*. Sebelum menggunakan *library* tersebut harus dilakukan instalasi dengan perintah PIP, instalasi dapat dilihat pada Gambar 4.3 dan Gambar 4.4.

```
pip install googleapiclient
pip install pandas
```

Python

**Gambar 4.3** Instalasi *Library Googleapiclient*

```
import googleapiclient.discovery
import pandas as pd
```

Python

**Gambar 4.4** *Import Library Googleapiclient dan Pandas*

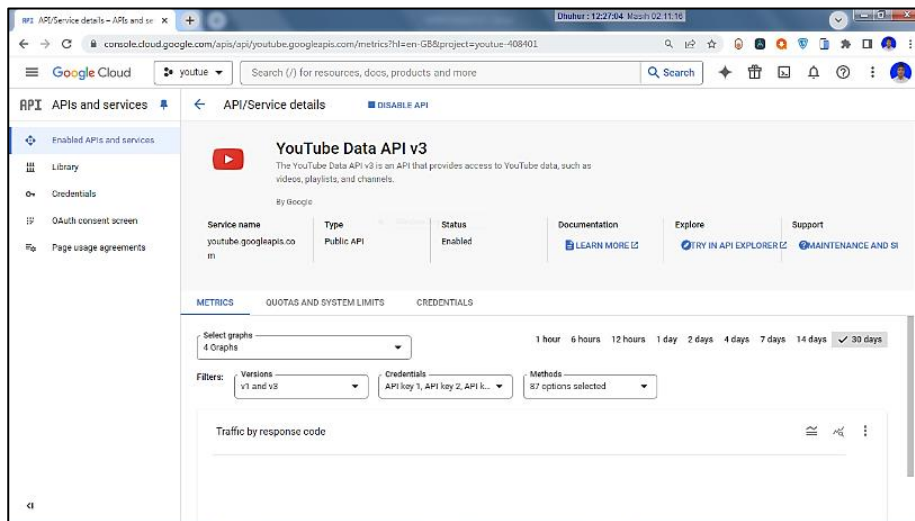
Selain itu, peneliti juga menggunakan *secret\_key* atau kunci rahasia yang diambil dari layanan *Youtube data API V3*, kunci ini digunakan sebagai identitas dari akun *google* yang peneliti gunakan agar dapat mengambil komentar-komentar video yang ada di *platform youtube*, *secret\_key* tersebut dapat dilihat pada Gambar 4.5.

```
api_service_name = "youtube"
api_version = "v3"
DEVELOPER_KEY = "AIzaSyCC1VyJaZrwTRQc5hgeCTT1un67dpJ6VdY"
```

Python

**Gambar 4.5** *Secret Key Google API*

Layanan *youtube data API V3* merupakan sebuah fasilitas yang disediakan oleh *google* yang bertujuan untuk dapat terhubung ke beberapa produk *google* seperti *Gmail*, *Youtube*, *Google Drive* dan lainnya. Dengan memanfaatkan fasilitas tersebut peneliti mengambil *secret\_key* yang akan digunakan untuk dimasukkan kedalam kode program *python* yang peneliti buat agar dapat terhubung dan memperoleh data yang di inginkan. *Youtube data API V3* dapat dilihat pada Gambar 4.6.



**Gambar 4.6** Youtube Data API V3

Atribut `video_id` yang telah disimpan sebelumnya, dimasukkan kedalam kode program untuk dilakukan proses *crawling*, perulangan (*looping*) digunakan bertujuan untuk mengulangi pengambilan komentar dari masing-masing video dan akan disimpan didalam *DataFrame pandas* untuk ditampilkan. Tahap *crawling* ini membutuhkan waktu beberapa menit tergantung dari seberapa banyak komentar yang akan diambil.

Berdasarkan hasil *crawling* yang telah dilakukan diperoleh kurang lebih 7.547 data komentar yang diambil berdasarkan 10 konten video *youtube*. Pada data tersebut mencakup informasi seperti *author*, *updated\_at*, *like\_count*, *text*, *video\_id*, dan *public* yang sebelumnya telah ditambahkan sebagai informasi yang akan diambil. Hasil *crawling* tersebut dapat dilihat pada Gambar 4.7.

```
print('Total data yang didapatkan : ',len(df))
df
```

Total data yang didapatkan : 7547

	author	updated_at	like_count	text	video_id	public
0	@yanienaro8606	2024-02-17T07:12:58Z	0	Mungkin masih Mengkunsuldasi dgn pemerintah Ban...	dgAwS5837N0	True
1	@putukarme5228	2024-01-25T15:26:43Z	0	Kebalikan kemanusiaan bisa ditandang Rohingya ...	dgAwS5837N0	True
2	@GarenaAov-bk2gr	2024-01-21T09:45:12Z	0	Iya pemerintah hanya menampung sementara tapi...	dgAwS5837N0	True
3	@phonet8325	2024-01-20T18:55:26Z	0	Pak Jokowi minta tolongRuhingya Banglades h...	dgAwS5837N0	True
4	@elinurlina4061	2024-01-14T08:31:09Z	0	Kami bukan warga Aceh juga menolak keras penda...	dgAwS5837N0	True
...	...	...	...	...	...	...
928	@ryana79ryana51	2023-12-11T09:44:05Z	0	org asing angp org Indonesia terlalu mudah dan...	K3BVsk-flk4	True
929	@ridhoet1112	2023-12-11T09:43:35Z	0	Imigran gelap ya bukan pengungsi.	K3BVsk-flk4	True
930	@ryana79ryana51	2023-12-11T09:42:42Z	2	pulgnk ke negara asal	K3BVsk-flk4	True
931	@user-tn7tn7wi1i	2023-12-11T09:41:04Z	0	Bangsa mn	K3BVsk-flk4	True
932	@witoz5573	2023-12-11T09:39:09Z	0	Di kembalikan di laut	K3BVsk-flk4	True

7547 rows x 6 columns

**Gambar 4.7** Dataset yang Berhasil di *Crawling*

Selanjutnya, *dataframe* yang menyimpan seluruh data hasil *crawling* disimpan kedalam *file microsoft excel* dengan format “*xlsx*” menggunakan *library pandas* dan disimpan dengan nama “*dataset\_youtube\_irwan.xlsx*”. proses ini dapat dilihat pada Gambar 4.8.

```
df.to_excel('dataset_youtube_irwan.xlsx')
```

Python

**Gambar 4.8** Menyimpan *Dataset* ke *Excel*

Setelah data hasil *crawling* berhasil disimpan kedalam *file microsoft excel*, selanjutnya akan digunakan untuk proses analisis sentimen pada penelitian ini. Data tersebut terdiri dari 6 kolom yaitu :

1. *author* (penulis)
2. *updated\_at* (tanggal dibuat)
3. *like\_count* (jumlah *like*)
4. *text* (teks)
5. *video\_id* (atribut id didala video)
6. *public* (jenis penayangan)

Dalam penggunaannya nanti hanya kolom *text* yang akan digunakan dan kolom yang lainnya tidak digunakan atau di *drop*, hal ini dikarenakan hanya kolom teks yang memberikan informasi seputar opini pengguna yang dapat di analisis sentimennya. Tampilan data didalam *excel* tersebut dapat dilihat pada Gambar 4.9.

	A	B	C	D	E	F	G
	author	updated_at	like_count	text	video_id	public	
1	0	@yanienar0806	2024-02-17T07:12:5	0 mungkin masih mengkonsultasi dengan pemerintah Bangladesh akan dikembalikan ke coxs bazar di sana sudah dibangun	dgAw5S837N0	TRUE	
2	1	@putukarme5228	2024-01-25T15:26:4	0 Kebalikan kemanusiaan bisa ditandang Rohingnya minta ini minta tanah katanya	dgAw5S837N0	TRUE	
3	2	@GarenaAov-bk2gr	2024-01-21T09:45:1	0 Iya pemerintah hanya menampung sementara tapi niat mereka pergi dari negara nya itu ingin menetap di indo pak	dgAw5S837N0	TRUE	
4	3	@phonene8325	2024-01-20T18:55:2	0 Pak Jokowi minta tolongRuhingya Banglades harusSebab mau pemilu rakyat Indonesia. Orang asing ruhingya banglades	dgAw5S837N0	TRUE	
5	4	@elinturina061	2024-01-14T08:31:0	0 Kami bukan warga Aceh juga menolak keras pendatang gelap Rohingya demi NKRI.	dgAw5S837N0	TRUE	
6	5	@user-ih9fx7pu5v	2024-01-12T09:16:2	0 Jangan percaya Jokowi , dia biasa ngblu , yg sementara itu Jokowi krn kurang dr 1 th dia sdh kehilangan kekuasaannya. 1	dgAw5S837N0	TRUE	
7	6	@sitirhmaniyah3247	2024-01-08T11:42:2	0 Tampung sementara, lalu negara mn selanjudnya yg mau menerima? Sementara itu lama lo g cukup 1, 2 th bs pulu	dgAw5S837N0	TRUE	
8	7	@user-uy1ss7bd8c	2024-01-07T07:50:0	0 Andai Jokowi suharto... Urusan pengungsi sudah selesai.	dgAw5S837N0	TRUE	
9	8	@Sriharti107	2024-01-04T02:50:4	0 Bpk Jokowi yang terhormat , silahkan bantu etnis Rohingnya asal jangan sampai tinggal di Indonesia karena Tanah air ind	dgAw5S837N0	TRUE	
10	9	@sumiyatiyati4531	2024-01-03T17:08:0	0 Sementara bgm yg sdh 9 th rohingnya di Indonesia	dgAw5S837N0	TRUE	
11	10	@abizag7443	2024-01-03T02:53:4	0 Tegas..?ditampungnya berapa lama pak..?Krna faktanya, banyak yg di tampung sampai bertahun tahun disini.Indonesia	dgAw5S837N0	TRUE	
12	11	@maswaraj7629	2024-01-01T15:40:3	0 UNHCR itu penjajah..pk de Jokowi harus tegas..usir Rohingnya.	dgAw5S837N0	TRUE	
13	12	@arifiharun8186	2023-12-31T18:21:2	0 Mana bukti ucapan le joko buktikan buktinya makin nambah	dgAw5S837N0	TRUE	
14	13	@user-oz4ob7uc2l	2023-12-31T10:43:0	0 Antar keselo	dgAw5S837N0	TRUE	
15	14	@livayansugiana3843	2023-12-30T12:48:3	0 kenapa ditampung....thailand sj mampu mengusir Indonesia LEMBEX	dgAw5S837N0	TRUE	
16	15	@didikhario8155	2023-12-28T00:14:0	0 Mnaaa tindakan tegasnya kok blm diangut2 ke negaranya	dgAw5S837N0	TRUE	
17	16	@arwanabur4322	2023-12-27T11:34:3	0 Pak Jokowi kami meminta bubarkan organisasi UNHCR	dgAw5S837N0	TRUE	
18	17	@Wahyusutrisno-hm3l	2023-12-27T01:19:5	0 Yth pak Jokowi usir pengungsi rohingnya dari negara kita Indonesia	dgAw5S837N0	TRUE	
19	18	@user-ek8ui8rf6w	2023-12-26T15:53:1	0 Sementara gimana? Udah kelamaan dan kebanyakan	dgAw5S837N0	TRUE	
20	19	@tubagurizki7470	2023-12-25T23:59:4	0 Sementara nya sampai kapan?1thn/5thn/10thn/50thn....lebih cepat lebih baik usir imigran ilegal	dgAw5S837N0	TRUE	
21	20	@relokasi7417	2023-12-25T08:15:0	0 D TAMPUNG SEMENTARA HINGGA BERANAK PINAK DAN MNJADI NEGARA KESATUAN ROHINGI ROHINHI	dgAw5S837N0	TRUE	
22	21	@rezekoryuh42271	2023-12-24T05:47:5	0 Tegas pak pulangi aja..	dgAw5S837N0	TRUE	
23	22	@putrabanjara4474	2023-12-23T10:44:4	0 Kirim Balik ke Bangladesh. 🇮🇩	dgAw5S837N0	TRUE	
24	23	@user-vf2grjnlj	2023-12-21T14:34:0	0 Secepatnya pak, di pulangkan karna warga Indonesia sudah muak dengan orang ruhingya.	dgAw5S837N0	TRUE	
25	24	@yai-tc2cn	2023-12-20T19:34:1	0 Di tampung sementara,, kira " 10 th, seperti pengungsi yg lbh dulu di Indonesia, jancok..Jancok	dgAw5S837N0	TRUE	
26	25	@user-ge8vdf01t	2023-12-19T22:52:5	0 Belajar dari sejarah, awalnya penjajahan.Adalah ketulusan, keluguan kesantunan bangsa kita menerima dan menghorm	dgAw5S837N0	TRUE	
27	26	@rayanaaceh1210	2023-12-19T21:10:2	0 Secepat nya di pulangkan sebelum Indonesia hancur ,semoga Allah tunjukan jalan kebenaran ,kasihan rakyat sendiri ser	dgAw5S837N0	TRUE	
28	27	@user-zb1xx4sg1n	2023-12-19T17:40:3	0 Sementara nya sampai kapan ? Ingat bapak tll lama lagi sdh purna bakti jadi presiden. Pulangkan secepatnya lebih bagu	dgAw5S837N0	TRUE	

**Gambar 4.9** *Dataset* Komentar Youtube

## 4.1.2 Representasi Data

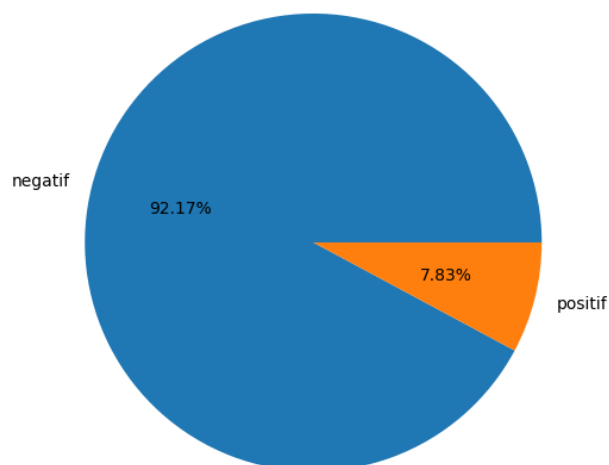
### 4.1.2.1 Pemberian Label

Pemberian label dilakukan dengan cara Manual, peneliti mulai membaca seluruh teks komentar satu per satu. Dalam sehari peneliti dapat memberikan label sebanyak 1000 data selama 7 hari peneliti menghasilkan 6956 sentimen Negatif dan 591 sentimen Positif. Hal ini dilakukan karena sebelumnya peneliti sudah mencoba memberikan label menggunakan bantuan *Lexicon Based* secara otomatis. Namun, cara tersebut tidak efektif dikarenakan sentimen yang diberikan tidak sesuai dengan isi dari komentarnya. Berikut hasil pemberian label pada gambar 4.10 dan 4.11.

	B	C	D	E	F	G	H
	author	updated_at	like_count	text	video_id	publi	label
2	@yanienan8606	2024-02-17T07:12:5		0 mungkin masih mengkonsultasi dengan pemerintah Bangladesh akan dikembalikan ke coxs bazar di sana sudah dibangun dgAw5837NO		TRUE	positif
3	@putukame5228	2024-01-25T15:26:4		0 Kebalikan kemanusiaan bisa ditandang Rohingya minta ini minta tanah katanya dgAw5837NO		TRUE	negatif
4	@GarenaAov-5k2gr	2024-01-21T09:45:1		0 Iya pemerintah hanya menampung sementara tapi niat mereka pergi dari negara nya itu ingin menetap di indo pak dgAw5837NO		TRUE	positif
5	@phonene8325	2024-01-20T18:55:2		0 Pak Jokowi minta tolongRohingya Banglades harussebab mau pemilu rakyat Indonesia. Orang asing ruhingya banglades dgAw5837NO		TRUE	positif
6	@elinlirln4061	2024-01-14T08:31:0		0 Kami bukan warga Aceh juga menolak keras pendatang gelap Rohingya.demi NKRI. dgAw5837NO		TRUE	negatif
7	@user-ih9fx7p3uv	2024-01-12T09:16:2		0 Jangan percaya Jokowi , dia biasa ngibul , yg sementara itu Jokowi krn kurang dr 1 th dia sdh kehilangan kekuasaannya. dgAw5837NO		TRUE	negatif
8	@setirohmaniyah3247	2024-01-08T11:42:2		0 Tampung sementara, lalu negara mn selanjudnya yg mau menerima nya? Sementara itu lama lo g cukup 1, 2 th bs pulu dgAw5837NO		TRUE	negatif
9	@user-uy1ss7b08c	2024-01-07T07:50:0		0 Andai Jokowi suharto... Urusan pengungsi sudah selesai. dgAw5837NO		TRUE	negatif
10	@srHartati107	2024-01-04T02:50:4		0 Bpk Jokowi yang terhormat , silahkan bantu etnis Rohingnya asal jangan sampai tinggal di Indonesia karena Tanah air Ind dgAw5837NO		TRUE	positif
11	@sumiyatijay4531	2024-01-03T17:08:0		0 Sementara bgn yg sdh 9 th rohingya di indonesia dgAw5837NO		TRUE	negatif
12	@abizaa7443	2024-01-03T02:53:4		0 Tegas...ditampungnya berapa lama pak. ?Krn faktanya, banyak yg di tampung sampai bertahun tahun disini.Indonesia dgAw5837NO		TRUE	negatif
13	@maswarija7629	2024-01-01T15:40:3		0 UNHCR itu penjajah..pk de Jokowi harus tegas...usir Rohingya. dgAw5837NO		TRUE	negatif
14	@arifinharun8186	2023-12-31T18:21:2		0 Mana bukti ucapan le joko buktikan buktinya makin nambah. mending orang aceh reverendum klu ga di den dgAw5837NO		TRUE	negatif
15	@user-oz4ob7u2cl	2023-12-31T10:43:0		0 Antar keselo dgAw5837NO		TRUE	positif
16	@wayansugama3943	2023-12-30T12:48:3		0 kenapa ditampung.... thailand sj mampu mengusir Indonesia LEMBEK dgAw5837NO		TRUE	negatif
17	@didiharonon8155	2023-12-28T00:14:0		0 Mnaas tindakan tegasnya kok blm diangut2 ke negaranya dgAw5837NO		TRUE	negatif
18	@arwanabiru4322	2023-12-27T11:34:3		0 Pak jokowi kami meminta buatkan organisasi UNHCR dgAw5837NO		TRUE	negatif
19	@WahyuSutrisno-hm3l	2023-12-27T01:19:5		0 Yth pak Jokowi usir pengungsi rohingya dari negara kita indonesia dgAw5837NO		TRUE	negatif
20	@user-ekku8ir8fw	2023-12-26T15:53:1		0 Sementara gimana? Udah kelamaan dan kebanyakan dgAw5837NO		TRUE	negatif
21	@tubagusizik7170	2023-12-25T23:59:4		0 Sementara nya sampai kapan?11thn/5thn/10thn/50thn...lebih cepat lebih baik usir imigran ilegal dgAw5837NO		TRUE	negatif
22	@relokak7417	2023-12-25T09:15:0		0 D TAMPUNG SEMENTARA HINGGA BERAKAN PINAK DAN MNIADI NEGARA KESATUAN ROHINGI 我我我我我我 dgAw5837NO		TRUE	negatif
23	@reizenkonyu42271	2023-12-24T05:47:5		0 Tegas pak pulangin aja. dgAw5837NO		TRUE	negatif
24	@putrabanjard474	2023-12-23T10:44:4		0 Kirim Balik ke Bangladesh. dgAw5837NO		TRUE	negatif
25	@user-vfzrj6n1j	2023-12-23T14:34:0		0 Secepatnya pak, di pulangkan karna warga indonesia sudah muak dengan orang ruhingya. dgAw5837NO		TRUE	negatif
26	@wai-tc2cn	2023-12-20T19:34:1		0 Di tampung sementara, kira " 10 th, seperti pengungsi yg lbh dulu di indonesia, jancok.. jancok dgAw5837NO		TRUE	negatif
27	@user-ge8d4f01t	2023-12-19T22:52:5		0 Belajar dari sejarah, awalnya penjajahan.Adalah ketulusan, keluguan kesantunan bangsa kita menerima dan menghorm dgAw5837NO		TRUE	negatif
28	@rayanaoeh1210	2023-12-19T21:10:2		0 Secepat nya di pulangkan sebelum Indonesia hancur ,semoga Allah tunjukan jalan kebenaran ,kasihan rakyat sendiri ser dgAw5837NO		TRUE	negatif
29	@user-zblx44gIn	2023-12-19T17:40:3		0 Sementara nya sampai kapan ? ingat bapak tdi lama lagi sdh puma bakti jadi presiden. Pulangkan secepatnya lebih bagus dgAw5837NO		TRUE	negatif

Gambar 4.10 Dataset Sudah diberikan Label

Sentimen Pada Data Tindakan Pemerintah Indonesia Untuk Menampung Etnis Rohingya



Gambar 4.11 Persentase Sentimen Positif dan Negatif

Ketika label telah diberikan, selanjutnya akan dilakukan tahapan konversi label kedalam angka 1 dan 0 yang mendeskripsikan 1 sebagai Positif dan 0 sebagai Negatif. Hal ini dilakukan bertujuan agar komputer dapat mendeteksi label yang diberikan untuk di proses klasifikasi. Karenanya, komputer hanya mengetahui angka dalam prosesnya. Pada tahap ini digunakan *IF-ELSE* sebagai operator logika untuk menentukan konversi label. Hasil konversi label dapat dilihat pada Gambar 4.12.

	text	label	sentimen
0	mungkin masih mengkonsultasi dengan pemerintah...	positif	1
1	Kebalikan kemanusiaan bisa ditandang Rohingya ...	negatif	0
2	Iya pemerintah hanya menampung sementara tapi...	positif	1
3	Pak Jokowi minta tolong\nRohingya Banglades h...	positif	1
4	Kami bukan warga Aceh juga menolak keras penda...	negatif	0
...	...	...	...
7542	org asing angp org Indonesia terlalu mudah dan...	negatif	0
7543	Imigran gelap ya bukan pengungsi.	negatif	0
7544	pulangkan ke negara asal	negatif	0
7545	Bangsa kurang ajar	negatif	0
7546	Di kembalikan di laut	negatif	0

7547 rows x 3 columns

**Gambar 4.12** Hasil Konversi Label

#### 4.1.2.2 Preprocessing Data

Data komentar yang telah dikumpulkan dan telah diberikan label secara manual, selanjutnya akan dilakukan pembersihan teks. Proses ini terdiri dari tahapan *Case Folding, Cleansing, Tokenizing, Word Normalization, Stopwords Removal, dan Stemming*. Tahap ini bertujuan agar data dapat digunakan tanpa adanya gangguan (*noise*) yang tidak di inginkan. Proses ini dapat mempengaruhi kinerja dari model klasifikasi yang akan dibuat. Untuk mempermudah pemahaman pembaca terkait proses ini, peneliti mengambil 5 data acak pada dataset sebagai contoh untuk di *preprocessing* secara manual. Tahap tersebut dijabarkan sebagai berikut :

##### a. Case Folding

Pada tahap ini seluruh komentar akan diubah menjadi huruf kecil (*lowercase*) dari yang sebelumnya berbentuk huruf kapital (*uppercase*). Tahap ini bertujuan untuk menyeragamkan teks agar tidak ada yang kapital ketika di proses. tahap ini dapat dilihat pada tabel 4.1.

Tabel 4.1 Proses *Case Folding*







Text	Casefolding
Sementara??? 😊 yg kemaren 9 thn sdh di indonesia	sementara??? 😊 yg kemaren 9 thn sdh di indonesia
SEMENTARA ITU BERAPA LAMA YANG JELAS DAN PASTI DONG PAK, ITU YANG DI SIDOHARJO SUDAH 9THN GK PULANG2. KEENAKAN ITU IMIGRAN GELAP RUMAH GRATIS MAKAN GRATIS UANG BULANAN JUGA DAPAT... KERJANYA CUMA BERKEMBANG BIAK BIKIN ANAK.	sementara itu berapa lama yang jelas dan pasti dong pak, itu yang di sidoharjo sudah 9thn gk pulang2. keenakan itu imigran gelap rumah gratis makan gratis uang bulanan juga dapat... kerjanya cuma berkembang biak bikin anak.
👤 sementara kalo terus datang apa gk bahaya kalo sudah terlalu banyak Pak, apa sanggup.., kecuali ketika ada kapal yg datang harus juga ada kapal yg pergi pulangin pengungsi ke negara asalnya..	👤 sementara kalo terus datang apa gk bahaya kalo sudah terlalu banyak pak, apa sanggup.., kecuali ketika ada kapal yg datang harus juga ada kapal yg pergi pulangin pengungsi ke negara asalnya..
😞😞😞 sedih karena rohingya cuman minta tolong saja, mari pak dukung please	😞😞😞 sedih karena rohingya cuman minta tolong saja, mari pak dukung please
👍👍 good pak mari bantu rohingya demi saudara seiman kita kasihan mereka	👍👍 good pak mari bantu rohingya demi saudara seiman kita kasihan mereka

b. *Cleansing*

Pada tahap ini komentar akan dicek apakah terdapat simbol seperti  $(:^\wedge)[!?\()(*\&^\%\$#\@!)$ , angka 1234567890, *link url* seperti *httpss* atau *www*, karakter khusus dan *mention* (@). Jika ada maka akan dihapus dan hanya menyisakan teks. Proses ini dilakukan agar teks yang dihasilkan bersih dari gangguan (*noise*), tahap ini dapat dilihat pada tabel 4.2.

Tabel 4.2 Proses *Cleansing*

Casefolding	Cleansing
sementara??? 😊 yg kemaren 9 thn sdh di indonesia	sementara yg kemaren thn sdh di indonesia
sementara itu berapa lama yang jelas dan pasti dong pak, itu yang	sementara itu berapa lama yang jelas dan pasti dong pak itu yang di

di sidoharjo sudah 9thn gk pulang2. keenakan itu imigran gelap rumah gratis makan gratis uang bulanan juga dapat... kerjanya cuma berkembang biak bikin anak.	sidoharjo sudah thn gk pulang keenakan itu imigran gelap rumah gratis makan gratis uang bulanan juga dapat kerjanya cuma berkembang biak bikin anak
 sementara kalo terus datang apa gk bahaya kalo sudah terlalu banyak pak, apa sanggup..., kecuali ketika ada kapal yg datang harus juga ada kapal yg pergi pulangin pengungsi ke negara asalnya..	sementara kalo terus datang apa gk bahaya kalo sudah terlalu banyak pak apa sanggup kecuali ketika ada kapal yg datang harus juga ada kapal yg pergi pulangin pengungsi ke negara asalnya
   sedih karena rohingya cuman minta tolong saja, mari pak dukung please	sedih karena rohingya cuman minta tolong saja mari pak dukung please
  good pak mari bantu rohingya demi saudara seiman kita kasihan mereka	good pak mari bantu rohingya demi saudara seiman kita kasihan mereka

### c. Tokenizing

Pada tahap ini komentar yang masih berbentuk kalimat, akan dipisahkan menjadi satuan kata, hal ini bertujuan untuk memperkecil pola pencarian kata pada tahap *stopwords* dan *stemming*. Tahap ini dapat dilihat pada tabel 4.3.

**Tabel 4.3** Proses *Tokenizing*

Cleansing	Tokenizing
sementara yg kemaren thn sdh di indonesia	['sementara', 'yg', 'kemaren', 'thn', 'sdh', 'di', 'indonesia']
sementara itu berapa lama yang jelas dan pasti dong pak itu yang di sidoharjo sudah thn gk pulang keenakan itu imigran gelap rumah gratis makan gratis uang bulanan juga dapat kerjanya cuma berkembang biak bikin anak	['sementara', 'itu', 'berapa', 'lama', 'yang', 'jelas', 'dan', 'pasti', 'dong', 'pak', 'itu', 'yang', 'di', 'sidoharjo', 'sudah', 'thn', 'gk', 'pulang', 'keenakan', 'itu', 'imigran', 'gelap', 'rumah', 'gratis', 'makan', 'gratis', 'uang', 'bulanan', 'juga', 'dapat', 'kerjanya', 'cuma', 'berkembang', 'biak', 'bikin', 'anak']
sementara kalo terus datang apa gk bahaya kalo sudah terlalu banyak pak apa sanggup kecuali ketika ada kapal yg datang harus juga ada kapal yg pergi pulangin pengungsi ke negara asalnya	['sementara', 'kalo', 'terus', 'datang', 'apa', 'gk', 'bahaya', 'kalo', 'sudah', 'terlalu', 'banyak', 'pak', 'apa', 'sanggup', 'kecuali', 'ketika', 'ada', 'kapal', 'yg', 'datang', 'harus', 'juga', 'ada', 'kapal', 'yg', 'pergi', 'pulangin', 'pengungsi', 'ke', 'negara', 'asalnya']

	'pulangin', 'pengungsi', 'ke', 'negara', 'asalnya']
sedih karena rohingya cuman minta tolong saja mari pak dukung please	['sedih', 'karena', 'rohingya', 'cuman', 'minta', 'tolong', 'saja', 'mari', 'pak', 'dukung', 'please']
good pak mari bantu rohingya demi saudara seiman kita kasihan mereka	['good', 'pak', 'mari', 'bantu', 'rohingya', 'demi', 'saudara', 'seiman', 'kita', 'kasihan', 'mereka']

d. Word Normalization

Pada tahap ini seluruh komentar akan dinormalisasikan dari kata-kata yang memiliki singkatan, kata asing, dan kata gaul, hal ini dilakukan agar kata yang akan dihasilkan merupakan kata yang normal yang dapat dimengerti untuk diproses, tahap ini dapat dilihat pada tabel 4.4.

**Tabel 4.4** Proses *World Normalization*

<b>Tokenizing</b>	<b>Normalisasi</b>
['sementara', 'yg', 'kemaren', 'thn', 'sdh', 'di', 'indonesia']	['sementara', 'yang', 'kemaren', 'tahun', 'sudah', 'di', 'indonesia']
['sementara', 'itu', 'berapa', 'lama', 'yang', 'jelas', 'dan', 'pasti', 'dong', 'pak', 'itu', 'yang', 'di', 'sidoharjo', 'sudah', 'thn', 'gk', 'pulang', 'keenakan', 'itu', 'imigran', 'gelap', 'rumah', 'gratis', 'makan', 'gratis', 'uang', 'bulanan', 'juga', 'dapat', 'kerjanya', 'cuma', 'berkembang', 'biak', 'bikin', 'anak']	['sementara', 'itu', 'berapa', 'lama', 'yang', 'jelas', 'dan', 'pasti', 'dong', 'pak', 'itu', 'yang', 'di', 'sidoharjo', 'sudah', 'tahun', 'tidak', 'pulang', 'keenakan', 'itu', 'imigran', 'gelap', 'rumah', 'gratis', 'makan', 'gratis', 'uang', 'bulanan', 'juga', 'dapat', 'kerjanya', 'cuma', 'berkembang', 'biar', 'bikin', 'anak']
['sementara', 'kalo', 'terus', 'datang', 'apa', 'gk', 'bahaya', 'kalo', 'sudah', 'terlalu', 'banyak', 'pak', 'apa', 'sanggup', 'kecuali', 'ketika', 'ada', 'kapal', 'yg', 'datang', 'harus', 'juga', 'ada', 'kapal', 'yg', 'pergi', 'pulangin', 'pengungsi', 'ke', 'negara', 'asalnya']	['sementara', 'kalau', 'terus', 'datang', 'apa', 'tidak', 'bahaya', 'kalau', 'sudah', 'terlalu', 'banyak', 'pak', 'apa', 'sanggup', 'kecuali', 'ketika', 'ada', 'kapal', 'yang', 'datang', 'harus', 'juga', 'ada', 'kapal', 'yang', 'pergi', 'pulangin', 'pengungsi', 'ke', 'negara', 'asalnya']
['sedih', 'karena', 'rohingya', 'cuman', 'minta', 'tolong', 'saja', 'mari', 'pak', 'dukung', 'please']	['sedih', 'karena', 'rohingya', 'cuman', 'minta', 'tolong', 'saja', 'mari', 'pak', 'dukung', 'tolong']
['good', 'pak', 'mari', 'bantu', 'rohingya', 'demi', 'saudara', 'seiman', 'kita', 'kasihan', 'mereka']	['baik', 'pak', 'mari', 'bantu', 'rohingya', 'demi', 'saudara', 'seiman', 'kita', 'kasihan', 'mereka']

#### e. Stopwords Removal

Pada tahap ini komentar akan dihilangkan dari kata-kata yang tidak memiliki makna seperti kata penghubung dan keterangan, tahap ini dilakukan agar data yang dihasilkan merupakan data yang memiliki makna tanpa adanya kata keterangan, kata penghubung dan kata yang tidak memiliki arti. Tahap ini dapat dilihat pada tabel 4.5.

**Tabel 4.5** Proses *Stopwords Removal*

Normalisasi	Stopwords
['sementara', 'yang', 'kemaren', 'tahun', 'sudah', 'di', 'indonesia']	kemaren indonesia
['sementara', 'itu', 'berapa', 'lama', 'yang', 'jelas', 'dan', 'pasti', 'dong', 'pak', 'itu', 'yang', 'di', 'sidoharjo', 'sudah', 'tahun', 'tidak', 'pulang', 'keenakan', 'itu', 'imigran', 'gelap', 'rumah', 'gratis', 'makan', 'gratis', 'uang', 'bulanan', 'juga', 'dapat', 'kerjanya', 'cuma', 'berkembang', 'biar', 'bikin', 'anak']	sidoharjo pulang keenakan imigran gelap rumah gratis makan gratis uang bulanan kerjanya berkembang bikin anak
['sementara', 'kalau', 'terus', 'datang', 'apa', 'tidak', 'bahaya', 'kalau', 'sudah', 'terlalu', 'banyak', 'pak', 'apa', 'sanggup', 'kecuali', 'ketika', 'ada', 'kapal', 'yang', 'datang', 'harus', 'juga', 'ada', 'kapal', 'yang', 'pergi', 'pulang', 'pergi', 'pengungsi', 'ke', 'negara', 'asalnya']	bahaya sanggup kecuali kapal kapal pergi pulang pengungsi negara asalnya
['sedih', 'karena', 'rohingya', 'cuman', 'minta', 'tolong', 'saja', 'mari', 'pak', 'dukung', 'tolong']	sedih rohingya cuman tolong mari dukung tolong
['baik', 'pak', 'mari', 'bantu', 'rohingya', 'demi', 'saudara', 'seiman', 'kita', 'kasihan', 'mereka']	mari bantu rohingya saudara seiman kasihan

#### f. Stemming

Pada tahap terakhir ini, akan dihilangkan imbuhan dari setiap kata, hal ini dilakukan bertujuan agar kata yang akan dihasilkan nantinya merupakan kata yang baku yang tidak terindikasi imbuhan me-kan dan sebagainya, dan hanya menyisakan kata kerja. Proses ini akan mendeteksi masing-masing kata dan

dideteksi apakah kata tersebut memiliki imbuhan atau tidak, jika ada maka akan direplace dengan kata kerjanya saja. tahap ini dapat dilihat pada tabel 4.6.

**Tabel 4.6** Proses *Stemming*

Stopwords	Stemming
kemaren indonesia	['kemaren', 'indonesia']
sidoharjo pulang keenakan imigran gelap rumah gratis makan gratis uang bulanan kerjanya berkembang bikin anak	['sidoharjo', 'pulang', 'enak', 'imigran', 'gelap', 'rumah', 'gratis', 'makan', 'gratis', 'uang', 'bulan', 'kerja', 'kembang', 'bikin', 'anak']
bahaya sanggup kecuali kapal kapal pergi pulangkan pengungsi negara asalnya	['bahaya', 'sanggup', 'kecuali', 'kapal', 'kapal', 'pergi', 'pulang', 'ungsi', 'negara', 'asal']
sedih rohingya cuman tolong mari dukung tolong	['sedih', 'rohingya', 'cuman', 'tolong', 'mari', 'dukung', 'tolong']
mari bantu rohingya saudara seiman kasihan	['mari', 'bantu', 'rohingya', 'saudara', 'iman', 'kasihan']

### 4.1.3 Hasil Analisis Data

#### 4.1.3.1 Pembobotan TF-IDF

Setelah dilakukan *Text Preprocessing* dilakukan pembobotan menggunakan TFIDF, proses ini diawali dengan mengambil seluruh teks yang akan diberikan bobot dan dihitung nilai *Term Frequency* pada setiap kata, dihitung *Document Frequency* dan diakhiri dengan menghitung *Term Frequency-Inverse Document Frequency*. Untuk mempermudah pembaca dalam memahami pembahasan ini, sebagai contoh peneliti akan memberikan perhitungan *tf-idf* pada 3 komentar sebagai berikut:

(Dokumen 1) = “Kau tampung aj di rumh mu kami pun mau makan hidup pun susah jngn di tambh susah lgi”

(Dokumen 2) = ” 😊😊😊😊 lucu kau jokowi tolol“

(Dokumen 3) = “bapak aku akan mendukungmu untuk memegang rasa kemanusiaan kepada sesama. Mari kita bantu sementara mereka pak untuk tinggal sementara “

Setelah melewati tahapan *text preprocessing* maka akan menjadi seperti ini :

(Dokumen 1) =[‘tampung’ ‘makan’ ‘hidup’ ‘susah’ ‘susah’]

(Dokumen 2) =[‘lucu’ ‘jokowi’]

(Dokumen 3) = ['dukung' 'kemanusiaan' 'bantu' 'sementara' 'tinggal' 'sementara']

a. Menghitung Term Frequency

Pada tahap selanjutnya dilakukan perhitungan dengan teknik TFIDF yang bertujuan untuk membentuk *word vector* yang telah diberi nilai. Pada teknik TFIDF terdapat dua kata yaitu TF (*Term Frequency*) dan IDF (*Inverse Document Frequency*). Dimana TF merupakan jumlah sebuah kata dari setiap dokumen, sedangkan IDF berfungsi untuk mengurangi bobot suatu kata jika kemunculannya banyak tersebar di dokumen. Pada tahap ini terlebih dahulu dihitung TF seperti pada Tabel 4.7.

**Tabel 4.7** Contoh Menghitung *Term Frequency*

kata ( <i>term</i> )	TF		
	D1	D2	D3
tampung	1	0	0
makan	1	0	0
hidup	1	0	0
susah	2	0	0
lucu	0	1	0
jokowi	0	1	0
dukung	0	0	1
kemanusiaan	0	0	1
bantu	0	0	1
sementara	0	0	2
tinggal	0	0	1

b. Menghitung Document Frequency

Setelah didapatkan nilai TF, agar dapat masuk kedalam perhitungan IDF. Maka harus ditentukan DF (*Document Frequency*) terlebih dahulu. DF sendiri merupakan banyaknya jumlah dokumen yang mengandung kata tersebut. Perhitungan DF dapat dilihat pada Tabel 4.8.

**Tabel 4.8** Contoh Menghitung *Document Frequency*

kata ( <i>term</i> )	<i>TF</i>			<i>DF</i>
	<b>D1</b>	<b>D2</b>	<b>D3</b>	
tampung	1	0	0	1
makan	1	0	0	1
hidup	1	0	0	1
susah	2	0	0	2
lucu	0	1	0	1
jokowi	0	1	0	1
dukung	0	0	1	1
kemanusiaan	0	0	1	1
bantu	0	0	1	1
sementara	0	0	2	2
tinggal	0	0	1	1

## c. Menghitung Inverse Document Frequency

Setelah didapatkan nilai DF, selanjutnya masuk ke perhitungan IDF. Dikethui jumlah dokumen (D) yang digunakan sebagai contoh diatas sebanyak 3 komentar (D1,D2 dan D3) sehingga dapat diketahui dokumennya menjadi  $D = 3$ . Dimana perhitungan IDF dan Rumusnya dapat dilihat pada Tabel 4.9.

**Tabel 4.9** Contoh Menghitung *Inverse Document Frequency*

kata ( <i>term</i> )	<i>DF</i>	$D/DF$ ( $D=3$ )	<i>IDF</i> ( $\text{Log}(D/DF)$ )	$IDF + 1$
tampung	1	3	$\text{Log } 3 = 0,477$	1,477
makan	1	3	$\text{Log } 3 = 0,477$	1,477
hidup	1	3	$\text{Log } 3 = 0,477$	1,477
susah	2	1,5	$\text{Log } 1,5 = 0,176$	1,176
lucu	1	3	$\text{Log } 3 = 0,477$	1,477
jokowi	1	3	$\text{Log } 3 = 0,477$	1,477
dukung	1	3	$\text{Log } 3 = 0,477$	1,477

kemanusiaan	1	3	$\text{Log } 3 = 0,477$	1,477
bantu	1	3	$\text{Log } 3 = 0,477$	1,477
sementara	2	1,5	$\text{Log } 1,5 = 0,176$	1,176
tinggal	1	3	$\text{Log } 3 = 0,477$	1,477

Setelah nilai TF dan IDF didapatkan, langka selanjutnya menghitung TFIDF dengan mengalikan asil nilai TF dengan nilai IDF, berikut perhitungannya.

**Tabel 4.10** Contoh Menghitung TFIDF

kata (term)	TF			IDF ( $\text{Log}(D/DF)$ )	IDF + 1	TF*IDF		
	D1	D2	D3			D1	D2	D3
tampung	1	0	0	$\text{Log } 3 = 0,477$	1,477	1,477	0	0
makan	1	0	0	$\text{Log } 3 = 0,477$	1,477	1,477	0	0
hidup	1	0	0	$\text{Log } 3 = 0,477$	1,477	1,477	0	0
susah	2	0	0	$\text{Log } 1,5 = 0,176$	1,176	2,352	0	0
lucu	0	1	0	$\text{Log } 3 = 0,477$	1,477	0	1,477	0
jokowi	0	1	0	$\text{Log } 3 = 0,477$	1,477	0	1,477	0
dukung	0	0	1	$\text{Log } 3 = 0,477$	1,477	0	0	1,477
kemanusiaan	0	0	1	$\text{Log } 3 = 0,477$	1,477	0	0	1,477
bantu	0	0	1	$\text{Log } 3 = 0,477$	1,477	0	0	1,477
sementara	0	0	2	$\text{Log } 1,5 = 0,176$	1,176	0	0	2,352
tinggal	0	0	1	$\text{Log } 3 = 0,477$	1,477	0	0	1,477

Berdasarkan contoh perhitungan diatas, peneliti menerapkan TFIDF terhadap 7.547 dataset dan setelah proses pembobotan selesai akan menghasilkan kalimat yang sudah menjadi *array* dalam matrik, dimana setiap baris mewakili setiap dokumen, dan setiap kolom mewakili kata yang ada pada seluruh teks, total kata yang



#### 4.1.3.3 Klasifikasi Naïve Bayes Clasifier

*Dataset* yang telah melalui *text preprocessing* dan *Term Frequeny-Inverse Document Frequency* dilanjutkan ke proses pembelajaran (*learning*) menggunakan algoritma *Naïve Bayes Clasifier*. Sebelum dilakukan *modelling* data dipisah menjadi 70% data latih dan 30% data uji, 70% data latih akan masuk kedalam proses pembelajaran untuk menghasilkan model klasifikasi. Untuk mempermudah pembaca dalam memahami algoritma *Naïve Bayes Clasifier*, berikut peneliti jabarkan contoh perhitungan manualnya :

**Tabel 4.11** Contoh Menghitung *Naive Bayes Clasifier*

Username account	text	preprocessing	label
@deewiefortuna	♥ semangat pak jokowi sesama manusia harus saling bantu	semangat', 'jokowi', 'manusia', 'bantu'	Positif
@henryariesta134	Dikira gampang mengembalikan mereka setelah ditampung sementara. gimana jokowi ini	'kembalikan', 'mereka', 'tampung', 'jokowi'	Negatif
@senorahayu2941	Duhhh jomowi bkin gw ksel dg kebijakan ny!!	'jokowi', 'kesal', 'kebijakan'	Negatif
@ubay6082	pak jokowi jangan pulangkan mereka pak, kasihan mereka harus dibantu anda harus bantu pak	'jokowi', 'pulangan', 'mereka', 'kasihan', 'mereka', 'bantu', 'bantu'	positif

Untuk menghitung naïve bayes clasifier digunakan persamaan sebagai berikut:

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

Keterangan persamaan :

X = Data dengan kelas yang belum diketahui

H = Hipotesis data yang merupakan suatu kelas spesifik

P(H|X)= Probabilitas hipotesis H berdasarkan kondisi X (probabilitas posterior)

P(H) = Probabilitas hipotesis H (probabilitas prior)

P(X|H)= Probabilitas X berdasarkan kondisi pada hipotesis H

P(X) = Probabilitas

a Menghitung Probabilitas Prior

$$P(H) = \frac{\text{total label}}{\text{total data}}$$

$$P(\text{positif}) = \frac{2}{5} = 0.4 \quad P(\text{negatif}) = \frac{3}{5} = 0.6$$

b Menghitung Probabilitas *Likelihood*

$$P(X|H) = \frac{\text{kemuncuan kata disuatu label} + 1}{\text{total keseluruhan kata disuatu label} + \text{total data}}$$

$$P(\text{semangat}|\text{positif}) = \frac{1+1}{11+18} = \frac{2}{29} = 0.0689655172413793$$

$$P(\text{semangat}|\text{negatif}) = \frac{0+1}{7+18} = \frac{1}{25} = 0.04$$

$$P(\text{jokowi}|\text{positif}) = \frac{2+1}{11+18} = \frac{3}{29} = 0.103448275862069$$

$$P(\text{jokowi}|\text{negatif}) = \frac{2+1}{7+18} = \frac{3}{25} = 0.12$$

$$P(\text{manusia}|\text{positif}) = \frac{0+1}{11+18} = \frac{1}{29} = 0.0344827586206897$$

$$P(\text{manusia}|\text{negatif}) = \frac{1+1}{7+18} = \frac{2}{25} = 0.08$$

$$P(\text{bantu}|\text{positif}) = \frac{3+1}{11+18} = \frac{4}{29} = 0.1379310344827586$$

$$P(\text{bantu}|\text{negatif}) = \frac{0+1}{7+18} = \frac{1}{25} = 0.04$$

$$P(\text{kembalikan}|\text{positif}) = \frac{0+1}{11+18} = \frac{1}{29} = 0.0344827586206897$$

$$P(\text{kembalikan}|\text{negatif}) = \frac{1+1}{7+18} = \frac{2}{25} = 0.08$$

$$P(\text{mereka}|\text{positif}) = \frac{2+1}{11+18} = \frac{3}{29} = 0.103448275862069$$

$$P(\text{mereka}|\text{negatif}) = \frac{1+1}{7+18} = \frac{2}{25} = 0.08$$

$$P(\text{tampung}|\text{positif}) = \frac{0+1}{11+18} = \frac{1}{29} = 0.0344827586206897$$

$$P(\text{tampung}|\text{negatif}) = \frac{1+1}{7+18} = \frac{2}{25} = 0.08$$

$$P(\text{kesal}|\text{positif}) = \frac{0+1}{11+18} = \frac{1}{29} = 0.0344827586206897$$

$$P(\text{kesal}|\text{negatif}) = \frac{1+1}{7+18} = \frac{2}{25} = 0.08$$

$$P(\text{kebijakan}|\text{positif}) = \frac{0+1}{11+18} = \frac{1}{29} = 0.0344827586206897$$

$$P(\text{kebijakan}|\text{negatif}) = \frac{1+1}{7+18} = \frac{2}{25} = 0.08$$

$$P(\text{pulangan}|\text{positif}) = \frac{1+1}{11+18} = \frac{2}{29} = 0.0689655172413793$$

$$P(\text{pulangan}|\text{negatif}) = \frac{0+1}{7+18} = \frac{1}{25} = 0.04$$

$$P(\text{kasihan}|\text{positif}) = \frac{1+1}{11+18} = \frac{2}{29} = 0.0689655172413793$$

$$P(\text{kasihan}|\text{negatif}) = \frac{0+1}{7+18} = \frac{1}{25} = 0.04$$

Setelah dihitung probabilitas prior dan likelihood, maka langkah selanjutnya mencoba mengklasifikasikan suatu data berdasarkan perhitungan yang telah dihasilkan di atas, berikut contoh klasifikasi manualnya :

Username account	text	preprocessing	label
@vitaliavardiana968	Mantab pak Jokowi.....bantu aja langsung..... Biar mereka ga kedinginan di laut	'mantap', 'jokowi', 'bantu', 'langsung', 'mereka', 'dingin', 'laut'	?

#### **P(uji|positif)**

$$\begin{aligned} &= P(\text{positif}) \times P(\text{jokowi}|\text{positif}) \times P(\text{bantu}|\text{positif}) \times P(\text{mereka}|\text{positif}) \\ &= 0.4 \times 0.103448275862069 \times 0.1379310344827586 \times 0.103448275862069 \\ &= 5.90430111935709 \text{ (paling tinggi)} \end{aligned}$$

#### **P(uji|negatif)**

$$\begin{aligned} &= P(\text{negatif}) \times P(\text{jokowi}|\text{negatif}) \times P(\text{bantu}|\text{negatif}) \times P(\text{mereka}|\text{negatif}) \\ &= 0.4 \times 0.12 \times 0.04 \times 0.08 \\ &= 0.0001536 \end{aligned}$$

Maka, berdasarkan klasifikasi yang dihitung dapat disimpulkan jika data uji di atas dikategorikan ke dalam sentimen **POSITIF**.

## **4.2 Hasil**

### **4.2.1 Pengujian**

Model yang telah dibuat sebelumnya akan diuji dengan tujuan agar peneliti dapat mengetahui persentase ketepatan dan keakuratan dalam mengklasifikasikan data. dan telah diketahui bersama, sebelumnya data dipisah menjadi dua bagian yaitu 70% data latih dan 30% data uji. Data latih disimpan didalam  $x\_train$  dan  $y\_train$ ,

sedangkan data uji disimpan didalam  $x\_test$  dan  $y\_test$ . Berdasarkan 7.547 data yang dimiliki, 70% data latih berjumlah 5.285 data, dan 30% data uji berjumlah 2.265 data. nantinya, data uji akan di uji apakah label yang di prediksi sesuai berdasarkan label aslinya. *Split* data dapat dilihat pada gambar 4.16.

```

TRAINING
Komentar Training      : 5282 data
Sentimen Training      : 5282 data
TESTING
Komentar Testing       : 2265 data
Sentimen Testing       : 2265 data

```

**Gambar 4.16** Jumlah Data yang akan di Uji

#### 4.2.2 Confussion Matrix

Setelah dilakukan tahap pengujian model, maka akan menghasilkan hasil prediksi benar dan salah, hal ini dapat dilihat dengan menampilkan confusion matrix yang terdiri dari *accuracy\_score*, *preision\_score*, *recall\_score*, *clasification\_report*. Dapat disimpulkan hasil uji model memperoleh akurasi sebesar 0.93 atau 93%. Hasil tersebut dapat dilihat pada Gambar 4.17.

```

Akurasi      : 0.9395143487858719
Precision    : 1.0
Recall       : 0.16463414634146342

Confusion Matrix
[[2101  0]
 [ 137  27]]

Clasification Report
              precision    recall  f1-score   support

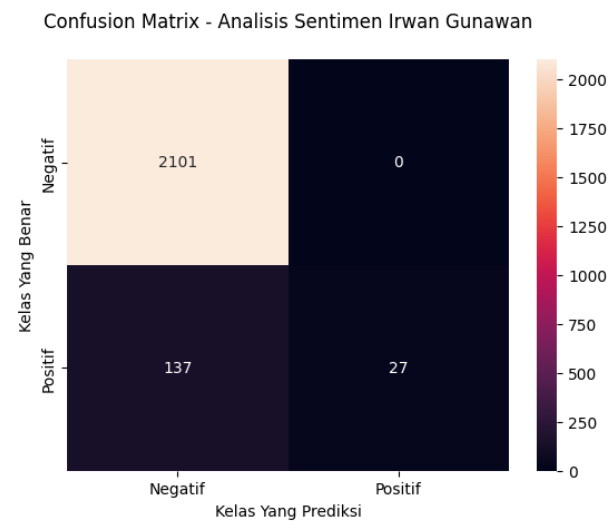
     0         0.94         1.00         0.97         2101
     1         1.00         0.16         0.28          164

 accuracy         0.94         0.58         0.63         2265
 macro avg        0.97         0.58         0.63         2265
 weighted avg     0.94         0.94         0.92         2265

```

**Gambar 4.17** Hasil *Confussion Matrix*

Selain itu, pada gambar diatas, dari total 2.265 data uji didapatkan bahwa 2101 dikategorikan sentimen Negatif dan 27 sentimen Positif. Namun, diketahui 137 yang seharusnya sentimen Positif justru diprediksi menjadi Negatif. Selain itu. Untuk mempermudah pembaca dalam memahami hasil dari klasifikasi yang dihasilkan pada *Confussion Matriix*. Maka, Peneliti juga memvisualisasikannya kedalam bentuk gambar seperti pada Gambar 4.18.



**Gambar 4.18** Visualisasi Confussion Matrix

### 4.2.3 Evaluasi Hasil

Setelah melalui proses pengujian model, selanjutnya pada proses ini akan dilakukan tahap evaluasi model untuk mengetahui performa dari model tersebut. Pada penelitian ini, peneliti melakukan perhitungan performa menggunakan *Confussion Matrix* 3x3 yang terdiri dari nilai *Accuracy*, *Preccision*, *Recall* dan *f1 Score*. Diketahui sebelumnya dari tahap pengujian model didapatkan hasil *multiclass confussion matrix* 3x3 yang ditunjukkan pada Tabel 4.12.

**Tabel 4.12** Hasil *Confussion Matrix*

		<i>False / Predicted Class</i>	
		Negatif	Positif
<i>True Class</i>	Negatif	2101	0
	Positif	137	27

Kemudian peneliti melakukan perhitungan secara manualisasi untuk menghitung performa dari model yang telah dibuat berdasarkan 7547 data yang telah diklasifikasikan menggunakan algoritma *Naïve Bayes Clasifier* pada nilai akurasi yang dihitung dari tabel *multiclass confussion matrix* diatas. Sehingga, hasil perhitungan nilai akurasi secara manual disajikan pada persamaan 8,9,10 dan 11.

$$\begin{aligned}
 \textit{Accuracy} &= \frac{TP+TN}{TP+TN+FP+FN} \times 100\% & (8) \\
 &= \frac{27+2101}{27+2102+0+137} \times 100\% \\
 &= \frac{2128}{2.265} \times 100\% \\
 &= 0,939514349 \times 100\% \\
 &= 93 \%
 \end{aligned}$$

$$\begin{aligned}
 \textit{Precision} &= \frac{27}{27+0} \times 100\% & (9) \\
 &= \frac{27}{27} \times 100\% \\
 &= 1 \times 100\% \\
 &= 100 \%
 \end{aligned}$$

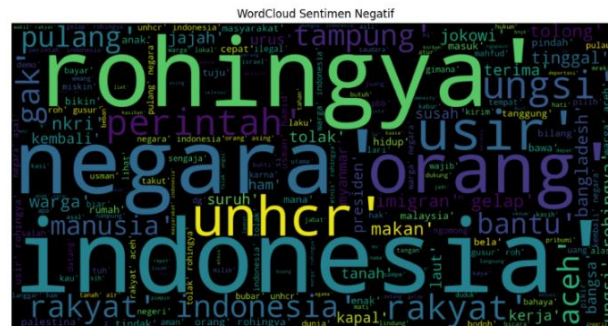
$$\begin{aligned}
 \textit{Recall} &= \frac{27}{27+137} \times 100\% & (10) \\
 &= \frac{27}{164} \times 100\% \\
 &= 0,164634146 \times 100\% \\
 &= 16 \%
 \end{aligned}$$

$$\begin{aligned}
 \textit{f1 Score} &= 2 \times \frac{100 \times 16}{100+16} & (11) \\
 &= 2 \times \frac{1600}{116} \\
 &= 27 \%
 \end{aligned}$$

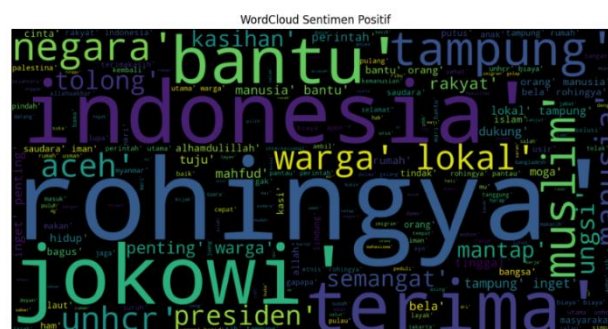
Berdasarkan perhitungan yang telah dilakukan, disimpulkan bahwasannya skor *accuracy* sebesar 93%, *preccision* 100%, *recall* 16%, dan *f1 score* 27%. Dengan begitu, Algoritma *Naïve Bayes Clasifier* yang digunakan dapat dikategorikan baik dalam mengklasifikasikan

#### 4.2.4 Visualisasi

Tahap terakhir yang akan dilakukan dalam penelitian ini yaitu dengan memvisualisasikan hasil dari analisis sentimen menggunakan *wordcloud*. Tujuan dari visualisasi ini adalah untuk mengetahui isu yang sering dibicarakan oleh masyarakat pada topik yang dibahas, sehingga dapat diperoleh informasi yang ditampilkan dalam bentuk gambar yang disajikan pada gambar berikut.



Gambar 4.19 *Wordcoud* pada Sentimen Negatif



Gambar 4.20 *Wordcoud* pada Sentimen Positif



Gambar 4.21 *Wordcoud* pada Sentimen Positif dan Negatif

#### 4.2.5 Penerapan

Penerapan sistem pada penelitian ini bertujuan untuk mengetahui bagaimana pandangan atau sentimen masyarakat indonesia terhadap tindakan pemerintah indonesia untuk menampung sementara pengungsi etnis rohingya pada media sosial youtue. Dengan menerapkan algoritma klasifikasi *Naïve Bayes Clasifier* dan uji tingkat akurasi model. Penerapan sistem ini juga berguna untuk membantu dan mempermudah mahasiswa maupun pelajar dalam melakukan riset analisis sentimen dalam bentuk informasi yang bisa disajikan dalam artikel, jurnal atau media massa.

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil dari proses analisis sentimen pada penelitian ini diperoleh beberapa kesimpulan sebagai berikut ini :

1. Analisis sentimen terhadap tindakan pemerintah Indonesia untuk menampung sementara pengungsi etnis Rohingya menggunakan algoritma *Naïve Bayes Classifier* mendapatkan hasil penerapan klasifikasinya pada data opini masyarakat terhadap topik yang dibahas pada *platform youtube*. Adapun pelabelan menggunakan cara manual dan didapat kelas sentimen Negatif lebih unggul berjumlah 6956 dibanding dengan kelas negatif berjumlah 591. Dengan kata lain masyarakat Indonesia dalam memberikan opini atau tanggapan kurang baik terhadap tindakan pemerintah Indonesia untuk menampung sementara pengungsi etnis Rohingya.
2. Dari penelitian ini juga menunjukkan metode klasifikasi algoritma *Naïve Bayes Classifier* dengan memperhatikan proses pengujian dapat dikategorikan baik dalam mengklasifikasikan opini karena mendapatkan *accuracy* sebesar 93%, *precision* 100%, *recall* 16%, dan *f1 score* 27%. Lalu, dengan memperhatikan pada *Confusion Matrix* bahwasannya data uji yang berhasil diklasifikasikan ke dalam sentimen negatif berjumlah 2101 dan data uji yang berhasil diklasifikasikan ke dalam sentimen positif berjumlah 27, sedangkan data uji yang salah diklasifikasikan berjumlah 137.

#### 5.2 Saran

Berikut adalah saran penulis dalam penelitian ini :

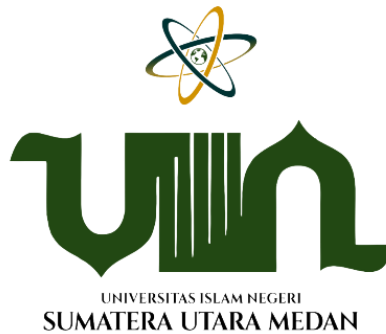
1. Penambahan dataset agar bisa lebih banyak seperti jutaan atau miliaran data.
2. Penambahan kamus kata baik itu kamus bahasa asing, gaul, serta singkatan.
3. Dalam penelitian ini data opini hanya dibatasi untuk menggunakan opini bahasa Indonesia. Maka dari itu, seharusnya dibutuhkan juga pengembangan kembali menggunakan terjemahan bahasa asing.

## DAFTAR PUSTAKA

- Adelia, S., Milanda, E., Santari, J., Kesuma, D. T., Silvia, E., & Kurniawan, F. (2023). Analisis Sentimen Belajar Programming Pada Media Sosial Youtube Menggunakan Algoritma Klasifikasi Naive Bayes. *Jurnal Of Information Technology Ampera*, 4(3).
- Adha, F. (2023). Penanganan Pengungsi Di Indonesia Melalui Kerja Sama Kemenkumham Dengan International Organizatio Of Migration. *Populis: Jurnal Sosia Dan Humaniora*, 8(2), 123–133.
- Aulia, Z. N., Jati, G. K., & Santoso, I. (2023). Analisis Sentimen Tanggapan Public Mengenai E-Tilang Melalui Media Sosial Youtube Menggunakan Algoritma Naive Bayes. *Jurnal Ikraith Informatika*, 7(2).
- Fahlevvi, M. R. (2022). Analisis Sentimen Terhadap Ulasan Aplikasi Pejabat Pengelola Informasi Dan Dokumentasi Kementerian Dalam Negeri Republik Indonesia Di Google Playstore Menggunakan Metode Support Vector Machine. *Jurnal Teknologi Dan Komunikasi*, 4(1), 1–13.
- Firsttama, R. A., Arifiyanti, A. A., & Kartika, D. S. Y. (204 C.E.). Analisis Sentimen Komentar Youtube Konferensi Tingkat Tinggi G20 Menggunakan Metode Naive Bayes. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 6(2), 282–285.
- Furqan, M., & Nasir, A. F. A. (2024). Big Data Approach To Sentiment Analysi In Machine Learning-Based Microblogs: Perspectives Of Religious Moderation Public Policy In Indonesia. *Journal Of Applied Engineering And Technological Science (JAETS)*, 5(2).
- Furqan, M., Sriani, & Sari, S. M. (2022). Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia. *Techno.Com*, 21(1), 52–61.
- Ghozali, M. I., Sugiharto, W. H., & Iskandar, A. F. (2023). Analisis Sentimen Pinjaman Online Di Media Sosial Twitter Menggunakan Metode Naive Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 3(6), 1340–1348.
- Guntara, R. G. (2023). Pemanfaata Google Colab Untuk Aplikasi Pendeteksian Masker Wajah Menggunakan Algoritma Deep Learning Yolov7. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(2), 55–60.
- Hamdi, M., Maulidia, H., & Firlana, H. (2023). Fenomena Pencari Suaka Dan Pengungsi Etnis Rohingya Di Indonesia (Studi Kasus Penangan Rohingya Pada Provinsi Aceh). *Jurnal Ilmiah Kajian Keimigrasian*, 6(1), 55–71.
- Hendra, J., & Nurdin, L. (2020). Eksistensi Media Sosial Youtube, Instagram Dan Whhatsapp Di Tengah Pandemi Covid-19 Dikalangan Masyarakat Virtual Indonesia. *Baituk Ulum : Jurnal Ilmu Perpustakaan Dan Informasi*, 4(1).
- Hermawan, F. (2022). Analisis Minat Masyarakat Pengguna Platform Youtube Sebagai Media Komunikasi Digital Masa Kini. *Jurnal Manajemen*, 14(3), 564–573.
- Hertati, Haerani, E., Novriyanto, & Syafira, F. (2023). Analisis Sentimen Masyarakat Terhadap Kenaikan Biaya Haji Tahun 2023 Menggunakan Metode Naive Bayes Clasifier. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(3), 1578–1584.
- Indradipradana, R. K., & Haridha, F. (2023). Kebijakan Luar Negeri Indonesia:

- Studi Kasus Penerimaan Pengungsi Rohingya Asal Myanmar Tahun 2020-2022. *Indonesian Perspective*, 8(2), 211–236.
- Kurniasih, U., & Suseno, A. T. (2022). Analisis Sentimen Terhadap Bantuan Subsidi Upah (BSU) Pada Kenaikan Harga Bahan Bakar Minyak (BBM). *Jurnal Media Informatika Budidarma*, 6(4), 2335–2340.
- Laurensz, B., & Sedyono, E. (2021). Analisis Sentimen Masyarakat Terhadap Tindakan Vaksinasi Dalam Upaya Mengatasi Pandemi Covid-19 (Analysis Of Public Sentiment On Vaccination In Efforts To Overcome The Covid-19 Pandemic). *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 10(2), 118–123.
- Listiarani, T. (2020). Analisis Kebijakan Luar Negeri Indonesia Dalam Menerima Pengungsi Rohingya Di Indonesia. *Jurnal PIR: Power In International Relations*, 5(1), 19–32.
- Mahadevi, A., Arindah, F., & Sandya, J. M. (2022). Implementasi Hak Asasi Manusia Internasional Dalam Pemenuhan Asas Membership Oleh Myanmar Kepada Etnis Rohingya. *Journal Of Politics And Democracy Studies*, 3(2), 142–157.
- Meilani, N., Furqan, M., & Suhardi. (2024). Analisis Sentimen Pengguna Aplikasi BSI Mobile Akibat Ransomware Menggunakan Algoritma Support Vector Machine. *Infotech: Jurnal Informatika Teknologi*, 5(1), 42–51.
- Mulyani, S., & Novita, R. (2022). Implementasi Algoritma Naive Bayes Classifier Untuk Klasifikasi Sentimen Masyarakat Tentang Depresi Pada Youtube. *Jurnal Teknik Informatika (JUTIF)*, 3(5), 1355–1361.
- Naraswati, N. P. G., Rosmilda, D. C., Desinta, D., Khairi, F., Damaiyanti, R., & Nooraeni, R. (2021). Analisis Sentimen Publik Dari Twitter Tentang Kebijakan Penanganan Covid-19 Di Indonesia Dengan Naive Bayes Classification. *Jurnal Sistem Informasi*, 10(1), 222–238.
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naive Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
- Nugraha, Y. S., Darusalam, U., & Iskandar, A. (2022). Implementasi Algoritma Genetika Pada Perancangan Aplikasi Penjadwal Instalasi Antivirus Berbasis Website Menggunakan Metode Waterfall. *Jurnal JTIK (Jurnal Teknologi Dan Komunikasi)*, 6(2).
- Oktavia, D., Ramadhan, Y. R., & Minarto. (2023). Analisis Sentimen Terhadap Penerapan Sistem E-Tilang Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVN). *Kajian Ilmiah Informatika Dan Komputer*, 4(1), 407–417.
- Prabowo, J. R., Akim, & Sudirman, A. (2022). Peran Tentara Nasional Indonesia Dalam Menanggulangi Pengungsi Dari Luar Negeri: Studi Kasus Pengungsi Rohingya Di Indonesia (2015-2020). *Aliansi: Jurnal Politik, Keamanan Dan Hubungan Internasional*, 1(2), 99–111.
- Pramita, N. T., & Zahidi, M. S. (2023). Analisis Peranan Indonesia Dalam Pemberian Bantuan Kemanusiaan Terhadap Etnis Rohingya Pada Era Jokowi. *Review Of International Relations*, 5(1), 84–106.
- Putra, G. M., & Faiza, D. (2022). Pengendali Suhu, Kelembaban Udara, Dan

- Intensitas Cahaya Pada Greenhouse Untuk Tanaman Bawang Merah Menggunakan Internet Of Things (IOT). *Jurnal Pendidikan Tambusai*, 5(3), 11404–11419.
- Rameliah, D. R., & Sagena, U. W. (2023). Krisis Keamanan Manusia Di Myanmar Dan Penerapan Konsep Responsibility To Protect. *Jurnal Pendidikan Tambusai*, 7(3), 27979–27986.
- Santoso, H., Armansyah, A., & Desliani, D. (2022). Analisis Sentimen Mahasiswa Terkait Pembelajaran Tatap Muka Menggunakan Metode Naive Bayes Classifier. *Techno.Com*, 21(3), 644–654.
- Sari, F. V., & Wibowo, A. (2019). Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naive Bayes Classifier Berbasis Konversi Ikon Emosi. *Jurnal SIMETRIS*, 10(2), 681–686.
- Semendawai, J. N., Febiola, I., Pamungkas, B., & Ruliansyah, M. D. (N.D.). Perancangan Aplikasi Otomatisasi Menggunakan Bahasa Pemrograman Python Pada Aktivitas Monitoring Pemakaian Data Harian Kartu Internet Of Things. *Jurnal Rekayasa Elektro Sriwijaya*, 3(1), 193–198.
- Setiawan, H., & Zufria, I. (2023). Analisis Sentimen Pembatalan Indonesia Sebagai Tuan Rumah Piala Dunia FIFA U-20 Menggunakan Naive Bayes. *Jurnal Media Informatika Budidarma*, 7(3), 1003–1012.
- Sumanjaya, A. A. A., Indriati, & Ridok, A. (2022). Analisis Sentimen Data Tweets Terhadap Penanganan Covid-19 Di Indonesia Menggunakan Metode Naive Bayes Dan Pemilihan Kata Bersentimen Menggunakan Lexicon Based. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 6(4), 1865–1872.
- Supiyandi, Hussein, A., Gunawan, I., & Harjo, W. L. R. (2023). Analisis Klasifikasi Broken Home Pada Anak Menggunakan Metode Naive Bayes Classifier. *JISKA : Jurnal Informatika Sunan Kalijaga*, 8(2), 90–101.
- Susana, H., Suarna, N., Fathurrohman, & Kaslani. (2022). Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet. *Jursistekni : Jurnal Sistem Informasi Dan Teknologi Informasi*, 4(1), 2–9.
- Tinambunan, T. M., & Siahaan, C. (2022). PEMANFAATAN YOUTUBE SEBAGAI MEDIA KOMUNIKASI MASSA DI KALANGAN PELAJAR. *Jurnal Ilmu Komunikasi*, 5(1).
- Wahyuni, W. (2022). Analisis Sentimen Terhadap Opini Feminisme Menggunakan Metode Naive Bayes. *Jurnal Informatika Ekonomi Bisnis*, 4(4), 148–153.
- Yasir, M., & Suraji, R. (2023). Perbandingan Metode Klasifikasi Naive Bayes, Decision, Tree, Random Forest Terhadap Analisis Sentimen Kenaikan Biaya Haji 2023 Pada Media Sosial Youtube. *Jurnal Cahaya Mandalika (JCM)*, 3(2), 180–192.



**BUKU CATATAN KEGIATAN PENELITIAN**  
**(LOG BOOK)**

Judul	Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia Untuk Menampung Sementara Pengungsi Etnis Rohingya Menggunakan Naïve Bayes Clasifier
Kluster	Penelitian
No. Register	-
Nama Peneliti	Irwan Gunawan
Fakultas / Prodi	Sains dan Teknologi / Ilmu Komputer
Alamat / Wa	Jalan Buntu Gang Sapardi No 351 Dusun VIII Desa Bandar Setia Kecamatan Percut Sei Tuan Kabupaten Deli Serdang Provinsi Sumatera Utara Indonesia Kode Pos 20371

UNIVERSITAS ISLAM NEGERI  
SUMATERA UTARA MEDAN

Nama Peneliti : Irwan Gunawan  
NIM : 0701201002  
Fakultas / Prodi : Sains dan Teknologi / Ilmu Komputer  
Judul Penelitian : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia Untuk Menampung Sementara Pengungsi Etnis Rohingya Menggunakan Naïve Bayes Clasifier  
Kuster : Penelitian  
Lokasi Penelitian : Laboratorium Ilmu Komputer Fakultas Sains dan Teknologi UIN Sumatera Utara yang berlokasi di Jalan Lapangan Golf, Desa Durin Jangak, Kecamatan Pancur Batu, Kota Medan  
Rentang Waktu Penelitian : Januari 2024 sampai Mei 2024



LOG BOOK

**Judul Penelitian** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Classifier

1	Hari / Tanggal / Bulan / Tahun	Rabu / 07 Februari 2023
2	Nama Kegiatan / Sub Kegiatan	Pengajuan Judul Penelitian
3	Tujuan Kegiatan /Sub Kegiatan (Sesuai Program)	Untuk mengetahui dan mengevaluasi kelayakan judul dan topik penelitian
4	Uraian Kegiatan	Menemukan topik penelitian yang relevan agar selanjutnya dapat diangkat menjadi judul proposal penelitian serta mendengarkan masukan dan saran dari dosen pembimbing
5	Hasil yang Diperoleh	Diperoleh judul dan topik penelitian yang relevan seputar analisis sentimen
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Judul memenuhi kriteria dan direkomendasikan untuk dilakukan penelitian lebih lanjut
8	Rencaran Kegiatan Selanjutnya	Pembuatan proposal penelitian

Mengetahui / Menyetujui  
Pembimbing



Dr, Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Medan, 07 Februari 2024

Mengetahui /Menyetujui  
Peneliti



Erwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Classifier

1	Hari / Tanggal / Bulan / Tahun	Rabu / 14 Februari 2024
2	Nama Kegiatan / Sub Kegiatan	Pembahasan Awal Proposal Penelitian
3	Tujuan Kegiatan /Sub Kegiatan (Sesuai Program)	Untuk mengetahui dan mengevaluasi kelayakan proposal berdasarkan topik
4	Uraian Kegiatan	Menjelaskan proposal penelitian dihadapan dosen pembimbing dan mendengarkan saran terkait penulisan proposal
5	Hasil yang Diperoleh	Diperoleh saran dan masukan
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Proposal penelitian memenuhi kriteria
8	Rencaran Kegiatan Selanjutnya	Penulisan BAB I

Medan, 14 Februari 2024

Mengetahui / Menyetujui  
Pembimbing



Dr, Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui /Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Clasifier

1	Hari / Tanggal / Bulan / Tahun	Kamis / 22 Februari 2024
2	Nama Kegiatan / Sub Kegiatan	Revisi Proposal Penelitian Tahap I
3	Tujuan Kegiatan /Sub Kegiatan (Sesuai Program)	Menghasilkan propoasl penelitian yang baik sesuai berdasarkan standart penulisan
4	Uraian Kegiatan	Mengoreksi dan mengubah isi proposal pada Latar Belakang, rumusan masalah dan penelitian terkait
5	Hasil yang Diperoleh	Dihasilkan proposal yang baik berdasarkan saran dan masukan pembimbing
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Proposal penelitian memenuhi standart
8	Rencaran Kegiatan Selanjutnya	Penulisan BAB II

Medan, 22 Februari 2024

Mengetahui / Menyetujui  
Pembimbing



Dr, Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui /Menyetujui  
Peneliti



Irfan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Clasifier

1	Hari / Tanggal / Bulan / Tahun	Rabu / 28 Februari 2024
2	Nama Kegiatan / Sub Kegiatan	Revisi Proposal Penelitian Tahap II
3	Tujuan Kegiatan /Sub Kegiatan (Sesuai Program)	Menghasilkan proposal yang baik
4	Uraian Kegiatan	Mengoreksi dan megubah isi pada BAB II, serta menambahkan beberapa rujukan referensi teori
5	Hasil yang Diperoleh	Dihasilkan proposal penelitian yang baik
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Proposal penelitian yang di revisi sesuai
8	Rencaran Kegiatan Selanjutnya	Penulisan BAB III

Medan, 28 Februari 2024

Mengetahui / Menyetujui  
Pembimbing



Dr, Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui /Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Classifier

1	Hari / Tanggal / Bulan / Tahun	Kamis / 29 Februari 2024
2	Nama Kegiatan / Sub Kegiatan	Revisi Proposal Penelitian Tahap III
3	Tujuan Kegiatan /Sub Kegiatan (Sesuai Program)	Menghasilkan proposal penelitian yang baik sesuai dengan standart penulisan untuk segera di ACC seminar proposal
4	Uraian Kegiatan	Mengoreksi dan mengubah tempat penelitian, metodologi penelitian, target penerbit jurnal pada BAB III
5	Hasil yang Diperoleh	Dihasilkan proposal yang baik secara menyeluruh pada BAB I, II, III untuk segera lanjut pada tahap seminar proposal
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Proposal penelitian memenuhi stadart
8	Rencaran Kegiatan Selanjutnya	Seminar Proposal dan Revisi dengan Dosen Penguji

Medan, 29 Februari 2024

Mengetahui / Menyetujui  
Pembimbing



Dr, Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui /Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Clasifier

1	Hari / Tanggal / Bulan / Tahun	Rabu / 27 Maret 2024
2	Nama Kegiatan / Sub Kegiatan	Seminar Proposal dan Revisi dengan Dosen Penguji
3	Tujuan Kegiatan /Sub Kegiatan (Sesuai Program)	Mempresentasikan dan mempertahankan topik
4	Uraian Kegiatan	Menjelaskan isi dan tujuan dari proposal penelitian dihadapan dosen pembing dan penguji
5	Hasil yang Diperoleh	Saran dan masukan (notulensi) perbaikan propoal pada isi dari BAB I, II, dan III
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Proposal penelitian lanjut
8	Rencaran Kegiatan Selanjutnya	Melanjutkan penelitian untuk BAB IV dan V, penelitian mencakup <i>crawling</i> data, <i>text preprocessing</i> , klasifikasi <i>Naïve Bayes Clasifier (modelling)</i> , evaluasi serta penulisan karya ilmiah

Medan, 27 Maret 2024

Mengetahui / Menyetujui  
Pembimbing



Dr, Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui /Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

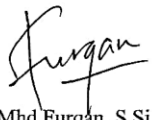
LOG BOOK

**Judul Penelitian** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Clasifier

1	Hari / Tanggal / Bulan / Tahun	Selasa / 01 April 2024
2	Nama Kegiatan / Sub Kegiatan	<i>Crawling Data Youtube</i>
3	Tujuan Kegiatan /Sub Kegiatan (Sesuai Program)	Untuk memperoleh data teks komentar pada penelitian ini untuk diklasifikasikan menggunakan Naïve Bayes Clasifier
4	Uraian Kegiatan	Pengambilan data menggunakan Google Collab
5	Hasil yang Diperoleh	Diperoleh 7.547 teks komentar
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Data mentah diperoleh untuk di proses
8	Rencaran Kegiatan Selanjutnya	Analisis, Pengolahan Data dan Penerapan

Medan, 01 April 2024

Mengetahui / Menyetujui  
Pembimbing



Dr, Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui /Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Clasifier

1	Hari / Tanggal / Bulan / Tahun	Rabu, 10 April 2024
2	Nama Kegiatan / Sub Kegiatan	Analisis, Pengolahan Data dan Penerapan
3	Tujuan Kegiatan /Sub Kegiatan (Sesuai Program)	Memperoleh hasil berupa data teks yang bersih dan telah memiliki label, serta model klasifikasi yang dikembangkan dengan Naïve Baye Clasifier
4	Uraian Kegiatan	Memberikan label secara manual serta menerapkan <i>Text Preprocessing</i> berupa <i>Case Folding, Cleansing, Tokenizing, Word Normalization, Stopwords Removel</i> dan <i>Stemming</i> , serta menerapkan <i>Naïve Bayes</i> Clasifier untuk diklasifikasikan
5	Hasil yang Diperoleh	Diperoleh 92.17% Positif, 7.83 Negatif, model klasifikasi dengan akurasi 93%
6	Hambatan	Tidak Ada
7	Kesimpulan dan Saran	Penelitian telah selesai
8	Rencaran Kegiatan Selanjutnya	Diskusi bersama dosen pembimbing

Mengetahui / Menyetujui  
Pembimbing



Dr. Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Medan, 10 April 2024

Mengetahui /Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002


LOG BOOK

**Judul Penelitian** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Classifier

1	Hari / Tanggal / Bulan / Tahun	Rabu / 08 Mei 2024
2	Nama Kegiatan / Sub Kegiatan	Diskusi Bersama Pembimbing
3	Tujuan Kegiatan /Sub Kegiatan (Sesuai Program)	Mendiskusikan Hasil penelitian yang telah dilaksanakan dan diuji
4	Uraian Kegiatan	Membahas seputar hasil penelitian dan tahap selanjutnya
5	Hasil yang Diperoleh	Mendapatkan hasil diskusi seputar hasil penelitian yang baik
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Hasil penelitian akan dituliskan ke dalam laporan penelitian dan Jurnal
8	Rencaran Kegiatan Selanjutnya	Penulisan Jurnal

Medan, 08 Mei 2024

Mengetahui / Menyetujui  
Pembimbing



Dr, Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui /Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Classifier

1	Hari / Tanggal / Bulan / Tahun	Rabu / 18 Juli 2024
2	Nama Kegiatan / Sub Kegiatan	Penulisan Jurnal
3	Tujuan Kegiatan /Sub Kegiatan (Sesuai Program)	Memaparkan karya ilmiah dalam bentuk teks sesuai dengan template jurnal
4	Uraian Kegiatan	Menulis jurnal sesuai peraturan penerbit
5	Hasil yang Diperoleh	Jurnal telah selesai ditulis sesuai rumah jurnal
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Jurnal siap untuk disubmit
8	Rencaran Kegiatan Selanjutnya	Submit jurnal ke penerbit <b>Journal of Computer Engineering, System and Science (CESS)</b>

Medan, 18 Juli 2024

Mengetahui / Menyetujui  
Pembimbing



Dr, Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui /Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian :** Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Clasifier

1	Hari/ Tanggal / Bulan / Tahun	Senin / 22 Juli 2024
2	Nama Kegiatan / Sub Kegiatan	Submit Jurnal ke Penerbit <b>Journal of Computer Engineering Sytem and Science (CESS)</b>
3	Tujuann Kegiatan / Sub Kegiatan (sesuai program)	Mengirimkan jurnal ke penerbit untuk di review, revisi dan dipublikasikan
4	Uraian Kegiatan	Melakukan submit jurnal ke penerbit serta menunggu proses korespondensi
5	Hasil yang di peroleh	Jurnal telah di upload dan menunggu ulasan oleh pihak editor
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Jurnal sudah diupload
8	Rencana Kegiatan Selanjutnya	Menunggu jurnal untuk di Review

Medan, 22 Juli 2024

Mengetahui / Menyetujui  
Pembimbing



Dr. Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui / Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian :** Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Clasifier

1	Hari/ Tanggal / Bulan / Tahun	Senin / 29 Juli 2024
2	Nama Kegiatan / Sub Kegiatan	Review Jurnal dari Penerbit <b>Journal of Computer Engineering Sytem and Science (CESS)</b>
3	Tujuann Kegiatan / Sub Kegiatan (sesuai program)	Memperhatikan proses review di system OJS di jurnal penerbit
4	Uraian Kegiatan	Peneliti mengecek melalui akun peneliti dan memperhatikan hasil review
5	Hasil yang di peroleh	Jurnal telah di review
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Jurnal telah di review,
8	Rencana Kegiatan Selanjutnya	Menunggu Keputusan Editor

Medan, 29 Juli 2024

Mengetahui / Menyetujui  
Pembimbing



Dr. Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui / Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian :** Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Classifier

1	Hari/ Tanggal / Bulan / Tahun	Selasa / 30 Juli 2024
2	Nama Kegiatan / Sub Kegiatan	Keputusan Editor terkait dari Penerbit <b>Journal of Computer Engineering System and Science (CESS)</b>
3	Tujuan Kegiatan / Sub Kegiatan (sesuai program)	Menerima keputusan dari editor
4	Uraian Kegiatan	Menerima hasil keputusan review
5	Hasil yang di peroleh	Jurnal lolos dari revisi dan diterima untuk di terbitkan pada <b>Vol, 9 No.2 Juli 2024</b>
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Jurnal di terima dan lolos Revisi
8	Rencana Kegiatan Selanjutnya	Menunggu LoA da proses Editing

Medan, 30 Juli 2024

Mengetahui / Menyetujui  
Pembimbing



Dr. Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui / Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian :** Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Clasifier

1	Hari/ Tanggal / Bulan / Tahun	Rabu / 31 Juli 2024
2	Nama Kegiatan / Sub Kegiatan	LoA dan Editing dari Penerbit <b>Journal of Computer Engineering Sytem and Science (CESS)</b>
3	Tujuann Kegiatan / Sub Kegiatan (sesuai program)	Menerima LoA dan jurnal diproses untuk Editing
4	Uraian Kegiatan	Menerima LoA dan Editing
5	Hasil yang di peroleh	Peneliti menerima LoA melalui email korespondensi dan jurnal di editing untuk di terbitkan
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	LoA diterima dan editing
8	Rencana Kegiatan Selanjutnya	Menunggu Jurnal di terbitkan

Medan, 31 Juli 2024

Mengetahui / Menyetujui  
Pembimbing



Dr. Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui / Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

LOG BOOK

**Judul Penelitian :** Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia  
Untuk Menampung Sementara Pengungsi Etnis Rohingya  
Menggunakan Naïve Bayes Classifier

1	Hari/ Tanggal / Bulan / Tahun	Kamis / 01 Juli 2024
2	Nama Kegiatan / Sub Kegiatan	Jurnal telah di terbitkan
3	Tujuann Kegiatan / Sub Kegiatan (sesuai program)	Melihat jurnal yang di terbitkan dan mengecek proses korespondensi akhir
4	Uraian Kegiatan	Membuka system OJS dan melihat jurnal yang diterbitkan pada archive <b>Vol. 9 No.2 Juli 2024</b>
5	Hasil yang di peroleh	Jurnal telah diterbitkan
6	Hambatan	Tidak ada
7	Kesimpulan dan Saran	Proses korespondensi selesai
8	Rencana Kegiatan Selanjutnya	Mendaftar sidang Kolokium

Medan, 01 Juli 2024

Mengetahui / Menyetujui  
Pembimbing



Dr. Mhd Furqan, S.Si, M.Comp, Sc  
NIP. 198008062006041003

Mengetahui / Menyetujui  
Peneliti



Irwan Gunawan  
NIM. 0701201002

## LOA



### Journal of Computer Engineering, System and Science

Jl. William Iskandar Ps. V Medan Estate - Sumatera Utara - Indonesia 20221

Email: [journal\\_CESS@unimed.ac.id](mailto:journal_CESS@unimed.ac.id)

Home Page: <https://jurnal.unimed.ac.id/2012/index.php/CESS>

ISSN: 2502-7131 (Print) | ISSN: 2502-714x (Online)

### LETTER of ACCEPTANCE (LoA)

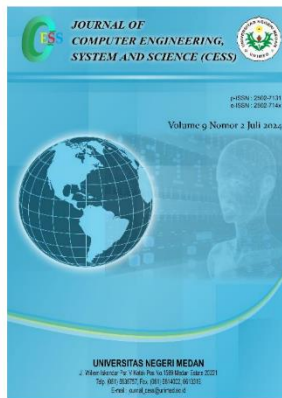
#### Kepada Penulis (Author)

Redaksi CESS (Journal of Computer Engineering, System and Science) menerangkan bahwa artikel dengan rincian berikut:

- Judul** : Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia Untuk Menampung Sementara Pengungsi Etnis Rohingya Menggunakan Naïve Bayes Clasifier
- Title** : *Sentiment Analysis of the Indonesian Government's Actions to Temporarily Accommodate Rohingya Ethnic Refugees Using the Naïve Bayes Clasifier*
- Penulis** : 1. Irwan Gunawan  
2. Mhd. Furqan
- Institusi** : Universitas Islam Negeri Sumatera Utara

Dinyatakan **diterima** karena telah memenuhi kriteria publikasi di Jurnal CESS (Journal of Computer Engineering, System and Science) pada **Volume 9 Nomor 2 Juli 2024**.

Demikianlah surat ini disampaikan, atas partisipasinya kami ucapkan terimakasih.



Medan, 30 Juli 2024

Editor in Chief

Mohamad Ihwani



# JURNAL

CESS (Journal of Computing Engineering, System and Science)  
Vol. 9, No. 2, Juli 2024, pp.737-749

e-ISSN: 2502-714x  
p-ISSN: 2502-7131

Contents list available at [www.jurnal.unimed.ac.id](http://www.jurnal.unimed.ac.id)

**CESS**  
**(Journal of Computing Engineering, System and Science)**  
journal homepage: <https://jurnal.unimed.ac.id/2012/index.php/cess>



## **Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia Untuk Menampung Sementara Pengungsi Etnis Rohingya Menggunakan Naïve Bayes Classifier**

### ***Sentiment Analysis of the Indonesian Government's Actions to Temporarily Accommodate Rohingya Ethnic Refugees Using the Naïve Bayes Classifier***

Irwan Gunawan<sup>1\*</sup>, Mhd. Furqan<sup>2</sup>

<sup>1,2</sup>Prodi Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara  
Jl Lap Golf No 120, Kp Tengah, Kec Pancur Batu, Kab Deli Serdang, Provinsi Sumatera Utara, Indonesia  
email: <sup>1</sup>[irwan0701201002@uinsu.ac.id](mailto:irwan0701201002@uinsu.ac.id), <sup>2</sup>[mfurqan@uinsu.ac.id](mailto:mfurqan@uinsu.ac.id)

#### **ABSTRAK**

Etnis Rohingya merupakan penduduk asli di negara myanmar yang sebagian besar mayoritasnya beragama muslim. Konflik yang terjadi pada etnis tersebut dimulai sejak ditetapkannya kebijakan *Burma Citizen Law* oleh pemerintah myanmar. kebijakan ini berisi terkait penolakan pemerintah myanmar terhadap etnis Rohingya sebagai etnis resmi dan memutuskan jika etnis tersebut tidak termasuk dari negara Myanmar. Indonesia merupakan salah satu negara di ASEAN yang masih menampung sementara pengungsi Rohingya, tindakan ini dilakukan berdasarkan konsep *Human Security* dan mengacu pada Peraturan Presiden Republik Indonesia Nomor 125 Tahun 2016 Tentang Penanganan Pengungsi Dari Luar Negeri Pasal 4 Ayat 2 mengenai koordinasi penanganan pengungsi yang meliputi Penemuan, Penampungan, Pengamanan dan Pengawasan. Akibatnya, terjadinya cemburu sosial yang berdampak pada keberagamannya opini masyarakat dan menjadi isu yang sering dibicarakan. Penelitian ini bertujuan untuk mengetahui kecenderungan opini berdasarkan klasifikasi sentimen yang diperoleh melalui video *YouTube*. Manfaat dari penelitian ini adalah agar pemerintah indonesia dapat mengetahui tindakan tersebut cenderung positif atau negatif. Dalam penelitian ini menerapkan algoritma *Naïve Bayes Classifier* dengan dataset berjumlah 7547 yang dibagi menjadi 6037 data latih dan 1510 data uji. Hasil *Confussion Matrix* pada penelitian ini menunjukkan akurasi 93%.

**Kata Kunci:** *Analisis Sentimen; Menampung; Pemerintah; Rohingya; Penambahan Teks; Youtube; Naive Bayes Clasifier*

#### **ABSTRACT**

The Rohingya ethnic group is a native population in the country of myanmar, where the majority are muslim. The conflict that occurred in this ethnic group began when the Burma

\*Penulis Korespondensi:  
email: [irwan05072002@gmail.com](mailto:irwan05072002@gmail.com)

Citizen Law policy was implemented by the myanmar government. This policy contains the Myanmar government's rejection of the Rohingya ethnicity as an official ethnicity and deciding that this ethnicity is not included in the state of Myanmar. Indonesia is one of the countries in ASEAN which is still temporarily hosting Rohingya refugees. This action is carried out based on the concept of Human Security and refers to Presidential Regulation of the Republic of Indonesia Number 125 of 2016 concerning Handling of Refugees from Abroad Article 4 Paragraph 2 concerning coordination of refugee handling which includes Discovery, Shelter, Security and Supervision. As a result, social jealousy occurs which has an impact on the diversity of public opinion and becomes an issue that is often discussed. To find out this, this research aims to find out opinion trends based on sentiment testimonials obtained through YouTube videos. The benefit of this research is that the Indonesian government can find out whether these actions tend to be positive or negative. In this research, the Naïve Bayes Classifier algorithm was applied with a dataset of 7547 which was divided into 6037 training data and 1510 test data. The results of the Confusion Matrix in this research show that the resulting accuracy level reached 93.

**Keywords:** *Sentiment Analysis; Accommodate; Government's; Rohingya; Text Mining; Youtube; Naive Bayes Classifier*

## 1. PENDAHULUAN

Etnis Rohingya merupakan penduduk asli di negara myanmar yang sebagian besar mayoritasnya beragama Muslim, konflik yang terjadi pada etnis tersebut dimulai sejak ditetapkannya kebijakan *Burma Citizen Law* oleh pemerintah Myanmar pada tahun 1982 [1], kebijakan tersebut membahas terkait penolakan dari pemerintah Myanmar terhadap keberadaan Rohingya sebagai etnis resmi dan mengklaim bahwasanya etnis tersebut bukan bagian dari negara Myanmar. Indonesia termasuk salah satu negara di ASEAN (*Asean South East Nations*) yang menampung pengungsi Rohingya. Penyebaran etnis tersebut mencapai kurang lebih 1000 orang di indonesia [2]. Selain itu, berdasarkan data dari lembaga UNHCR (*United Nations High Commissioner for Refugees*) bulan maret tahun 2021 memperkirakan total pengungsi Rohingya di indonesia mencapai 5% dari 13.497 keseluruhan pengungsi dari berbagai negara yang datang ke indonesia [3].

Penyebab pemerintah indonesia menampung sementara pengungsi rohingya dikarenakan kemanusiaan. Menurut tirza listiarani, Indonesia memiliki kebijakan luar negeri yang terbuka terhadap pengungsi rohingya. dalam perencanaannya dilakukan dengan pendekatan *Human Security* bukan dengan *National Security* [4]. Selain itu, berdasarkan Peraturan Presiden Republik Indonesia Nomor 125 Tahun 2016 Tentang Penanganan Pengungsi Dari Luar Negeri yang terdapat pada Pasal 4 Ayat 2 mengenai koordinasi penanganan pengungsi yang meliputi Penemuan, Penampungan, Pengamanan dan Pengawasan. Pada pasal 26 Ayat 5 pemerintah juga memberikan bantuan kepada pengungsi rohingya mencakup penyediaan air bersih, kebutuhan makanan, pakaian, pelayanan kesehatan dan kebersihan serta fasilitas ibadah.

Akibatnya, terjadi cemburu sosial yang berdampak pada beragamnya opini masyarakat yang menjadi isu yang sering di bicarakan di media sosial. Media sosial telah menjadi sarana komunikasi dan interaksi dalam kehidupan masyarakat dunia [5]. *YouTube* menjadi salah satu media sosial yang menyebarkan berita melalui konten video terkait tindakan pemerintah indonesia yang menampung sementara pengungsi rohingya. Menurut Hermawan, sejak tahun 2022 diperkirakan 139 juta masyarakat indonesia menggunakan *YouTube* [6]. Maka dari itu,

YouTube menjadi media sosial yang sering digunakan oleh masyarakat Indonesia, serta menghasilkan data yang cukup banyak dan komentar pro dan kontra yang dapat di amati untuk di analisis. Maka, untuk mengetahui opini masyarakat terhadap tindakan pemerintah Indonesia untuk menampung sementara pengungsi etnis Rohingya, diperlukan Analisis Sentimen yang bertujuan untuk mengetahui kecenderungan opini berdasarkan klasifikasi sentimen melalui komentar video *youtube*. Analisis sentimen adalah pengolahan bahasa natural yang berperan menganalisis pendapat [7], dalam penerapannya analisis sentimen mengacu pada kemampuan perangkat lunak untuk memahami sikap, pendapat, atau pendirian pembicara atau penulis [8], analisis sentimen juga digunakan untuk menentukan sentimen tentang isu, peristiwa, kepuasan layanan, harga sahan dan analisis persaingan berdasarkan data teks [9].

Analisis sentimen yang dimaksud dalam penelitian ini akan menerapkan algoritma *Naive Bayes Classifier* dalam mengklasifikasikan data teks komentar. Menurut Adelia, *Naive Bayes Classifier* merupakan hasil probabilitas masa depan yang di ambil dari pengalaman masa sebelumnya [10]. Dan *Term Frequency-Inverse Document Frequency* untuk pembobotan kata (*term*). Menurut Mulyani TFIDF termasuk salah satu pembobotan kata terbaik dibandingkan dengan pembobotan kata lainnya [11].

Kajian terkait analisis sentimen menggunakan algoritma *Naive Bayes Classifier* sebelumnya telah dilakukan oleh beberapa penelitian terdahulu. Seperti pada penelitian berjudul "Analisis Sentimen Tanggapan *Public* Mengenai E-Tilang Melalui Media Sosial *Youtube* Menggunakan Algoritma *Naive Bayes*" yang dilakukan oleh [12]. Berdasarkan hasil penelitiannya opini yang dikumpulkan diambil dari media *Youtube*, karena media tersebut menjadi salah satu objek untuk menyampaikan opini karena bersifat *user friendly* dan terbuka dalam mengakses komentar. Namun, data yang digunakan selama penelitian berlangsung hanya berjumlah 500 komentar dan menghasilkan akurasi sebesar 79.44%. Selanjutnya, penelitian yang berjudul "Analisis Sentimen Terhadap Opini Feminisme Menggunakan Metode *Naive Bayes*" yang dilakukan oleh [13], berdasarkan hasil penelitiannya digunakan teknik pembobotan kata menggunakan *Term Frequency-Inverse Document Frequency* dengan pola data 80:20 atau 80% data latih dan 20% data uji yang menghasilkan akurasi sebesar 88%. Selanjutnya, penelitian yang berjudul "Analisis Sentimen Masyarakat Terhadap Tindakan Vaksinasi Dalam Upaya Mengatasi *Pandemic Covid-19*" yang dilakukan oleh [14]. Berdasarkan hasil penelitiannya pemberian label dilakukan secara manual dan menghasilkan akurasi sebesar 85.59%. Penelitian analisis sentimen ini akan menggunakan data berjumlah 7.547 komentar yang diambil dari platform *youtube* dan pemberian label dilakukan secara manual, Lalu sentimen yang akan digunakan hanya Positif dan Negatif, serta pembobotan kata yang akan diterapkan menggunakan teknik *Term Frequency-Inverse Document Frequency* dengan pola 70% data *training* dan 30% data *testing*

## 2. DASAR/TINJAUAN TEORI

### 2.1. Etnis Rohingya

Etnis Rohingya merupakan salah satu etnis dari sejumlah etnis yang berada di negara Myanmar, penduduk Myanmar yang beragama Islam sebagian besar berasal dari etnis Rohingya dan etnis Rohingya termasuk etnis yang menganut agama Islam di negara tersebut [15]. Keberadaan etnis Rohingya di negara Myanmar merupakan minoritas Islam dan minoritas yang terisolasi di Myanmar, Rohingya adalah bangsa yang terbentuk selama peristiwa yang panjang. Mereka ada di wilayah Rakhine sebelum Inggris berkuasa di Rakhine dan Burma

antara tahun 1824 dan 1948 [16]. Pelanggaran dan tindakan kejahatan kemanusiaan yang dilakukan oleh pemerintah Myanmar terhadap etnis rohingya berawal dari di tetapkannya Undang-Undang status kewarganegaraan tahun 1962. Peraturan tersebut menghapus status kewarganegaraan kelompok rohingya sehingga sebagian dari mereka kehilangan hak-hak dan perlindungannya dari negara tersebut [17].

## 2.2. Analisis Sentimen

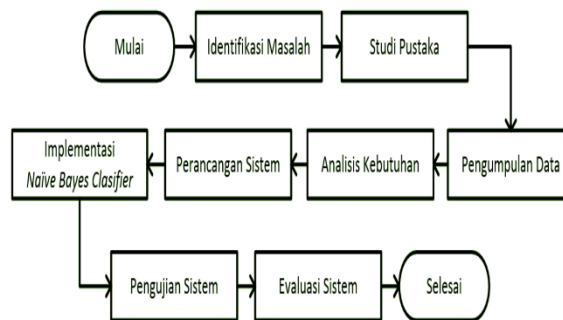
Analisis sentimen merupakan penambangan teks yang bertujuan untuk menganalisis dan mengekstraksi suatu informasi yang bersifat subjektif yang dapat membantu pihak yang membutuhkannya [19]. Analisis sentimen juga digunakan untuk mengekstrak data opini, memahami serta mengolah tekstual data secara otomatis untuk melihat sentimen yang terkandung di sebuah opini. Sentimen yang di kategorikan terbagi menjadi sentimen positif, netral dan negatif [20].

## 2.3. Text Mining

*Text Mining* merupakan proses penambangan yang dilakukan oleh komputer untuk mendapatkan sesuatu yang baru, dan tidak diketahui sebelumnya atau menemukan kembali informasi-informasi yang tersirat secara implisit, temuan tersebut berasal dari informasi yang di ekstrak secara otomatis dari sumber-sumber data teks yang berbeda [20].

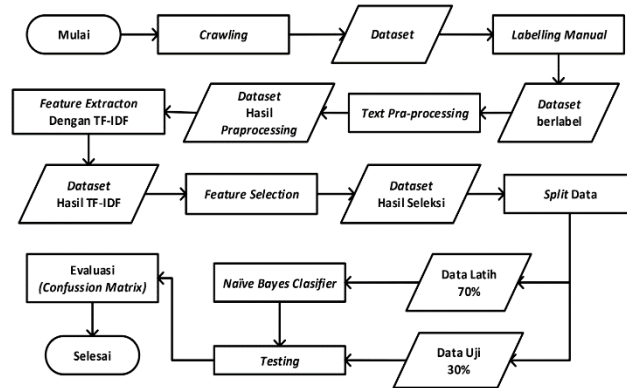
## 3. METODE

Metode penelitian yang digunakan dalam penelitian ini menggunakan metode Kualitatif, dan variabel yang diteliti adalah variabel *review*, karena data yang akan diteliti diperoleh berdasarkan hasil tanggapan masyarakat. Tahapan penelitian yang akan dilaksanakan disajikan dalam bentuk *flowchart* Gambar 1.



**Gambar 1.** Kerangka Penelitian

Perancangan model yang akan menggunakan *Naive Bayes Classifier* juga disajikan dalam *flowchart* sistem Gambar 2.



Gambar 2. Flowchart Sistem Analisis Sentimen

### 3.1 Term Frequency Inverse Document Frequency

TFIDF adalah suatu pembobotan kata untuk menghitung kata umum pada *information retrieval*, pembobotan ini dikenal sebagai gabungan *term frequency (TF)* dan *inverse document frequency (IDF)*, yang mana *TF* merupakan total keberadaan kata pada dokumen. Sedangkan *IDF* adalah cara untuk mengetahui seberapa penting kata pada suatu dokumen [22]

Rumus perhitungan pembobotan *TF-IDF* dalam [23] disajikan dalam persamaan berikut :

Persamaan untuk menghitung *TF (Term Frequency)*

$$TF = \frac{f_{t,d}}{\sum_{t \in d} f_{t,d}}$$

Keterangan persamaan:

TF = *Term Frequency*

d = dokumen

t = *term (kata/token)*

$\int_{t,d}$  = jumlah *term (kata/token)* pada setiap dokumen

$\sum_{t \in d} f_{t,d}$  = jumlah doumen yang memuat t

Persamaan untuk menghitung *IDF (Inverse Document Frequency)*

$$idf_t = \log \frac{N}{dft}$$

Keterangan persamaan:

*idf<sub>t</sub>* = Nilai *idf* dari *term (token) t*

N = Banyaknya dokumen yang tersedia

*dft* = Intensitas kemunculan kata pada dokumen

### 3.2 Naïve Bayes Classifier

*Naïve Bayes Classifier* adalah algoritma klasifikasi yang menerapkan kemungkinan sederhana pada *Teorema Bayes* dan menerapkan aturan ketidaktergantungan (*independent*) pada setiap kondisi [20]. *Naïve Bayes Classifier* termasuk *Supervised Learning*, dikarenakan diperlukan data *training* dalam mengklasifikasikan data kedalam sebuah label [21]. *Naïve Bayes* sering diterapkan dalam memprediksi kemungkinan pada keanggotaan suatu kelas [24].

Rumus perhitungan algoritma *Naïve Bayes Classifier* dalam [25] dijabarkan dalam persamaan berikut :

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

Keterangan persamaan:

- X = Data dengan kelas yang belum diketahui
- H = Hipotesis data yang merupakan suatu kelas spesifik
- P(H|X) = Probabilitas hipotesis H berdasarkan kondisi X (probabilitas posterior)
- P(H) = Probabilitas hipotesis H (probabilitas prior)
- P(X|H) = Probabilitas X berdasarkan kondisi pada hipotesis H
- P(X) = Probabilitas

### 3.3 Confusion Matrix

*Confusion Matrix* adalah tabel yang digunakan untuk menyatakan jumlah klasifikasi jumlah data uji [19]. *confusion matrix* yang dimaksud dalam [19] disajikan pada tabel 1.

Tabel 1. *Confusion Matrix*

		Kelas Prediksi	
		1	0
Kelas sebenarnya	1	TP	FN
	0	FP	TN

Selanjutnya, berikut rumus *confusion matrix* untuk mendapatkan nilai *accuracy*, *precision*, *f1 score* dan *recall* dalam [26] disajikan pada persamaan berikut :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad Recall = \frac{TP}{TP+FN} \times 100\%$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad F1 Score = 2 \times \frac{precision \times recall}{precision+recall}$$

## 4. HASIL DAN PEMBAHASAN

### 4.1. Pengumpulan Data

Adapun cara yang digunakan dalam memperoleh data dilakukan dengan teknik *Crawling Youtube* menggunakan *secret\_key* yang didapat pada layanan *Youtube Data API V3*. Dalam penelitian ini peneliti akan menganalisis data sentimen *youtube* terhadap tindakan pemerintah indonesia untuk menampung sementara pengungsi etnis rohingya. Dalam prosesnya digunakan bahasa pemrograman *python* dan *library googleapiclient* yang memungkinkan pengguna mencari informasi komentar di *platform youtube*. Hasil dari data yang dikumpulkan disimpan kedalam *file* berformat *.xls* dan data yang diambil pada *platform youtube* diperoleh sebanyak 7.547 data, hasil *crawling* dijabarkan dalam tabel 2.

Tabel 2. Hasil *Crawling* Data

No	Author	Updated_at	Like_count	Text	Video_id	Public
1	@yarienaro8606	2024-02-17T07:12:58Z	0	mungkin masih mengkonsultasi dengan pemerintah Bangladesh akan dikembalikan ke coxs bazar di sana sudah dibangun megah penampungan	dgAws5837N0	TRUE

@putuka	2024-01-25T15:26:43Z	0	Kembalikan kemanusiaan bisa ditandang Rohingya minta ini minta tanah katanya	dgAws5837N0	TRUE
...	...	...	...	...	...
@user-tn7t	2023-12-11T09:41:04Z	0	Bangsa kurang ajar	K3BVsk-flk4	TRUE
@wi-toz5	2023-12-11T09:39:57Z	0	Di kembalikan di laut	K3BVsk-flk4	TRUE

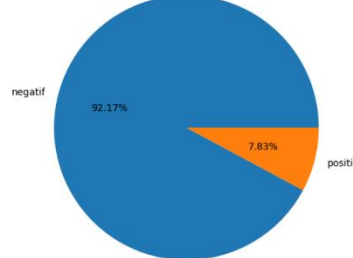
#### 4.2. Pemberian Label

Pemberian label dilakukan secara manual, peneliti mulai membaca seluruh teks komentar. Hal ini dilakukan karena sebelumnya peneliti sudah mencoba memberikan label menggunakan *Lexicon Based* dan *Textblob*. Namun, cara tersebut tidak efektif karena sentimen yang diberikan tidak sesuai dengan isi dari komentarnya. Hasil dari pemberian label dengan cara manual Gambar 3.

id	author	timestamp	text	label
1	@putuka	2024-01-25T15:26:43Z	Kembalikan kemanusiaan bisa ditandang Rohingya minta ini minta tanah katanya	negatif
2	@user-tn7t	2023-12-11T09:41:04Z	Bangsa kurang ajar	negatif
3	@wi-toz5	2023-12-11T09:39:57Z	Di kembalikan di laut	negatif

Gambar 3. Hasil Pemberian Label

Dari pemberian label terhadap 7.547 data komentar diketahui 6956 (92.17%) sentimen Negatif dan 591 (7.83%) sentimen Positif, perbandingan label tersebut pada Gambar 4.



Gambar 4. Perbedaan Label Negatif dan Positif

#### 4.3. Text Preprocessing

Setelah data diberikan label, tahap selanjutnya dilakukan pembersihan teks. Proses ini meliputi *Case Folding*, *Cleansing*, *Tokenizing*, *Word Normalization*, *Stopwords Removal*, dan *Stemming*. Tahap ini bertujuan agar data dapat digunakan tanpa adanya gangguan (*noise*) yang tidak di inginkan. Proses ini dapat mempengaruhi kinerja dari model klasifikasi yang akan dibuat. Adapun 6 tahapan pada *text preprocessing* tersebut dijabarkan sebagai berikut:

##### 4.3.1 Case Folding

Komentar akan diubah menjadi huruf kecil (*lowercase*) dari yang sebelumnya berbentuk huruf kapital (*uppercase*). Tahap ini bertujuan untuk menyeragamkan teks agar tidak ada yang kapital ketika di proses. Hasilnya pada Tabel 3

Tabel 3. Hasil *Case Folding*

Text Asli	Case Folding
y sdh tampung semua demi kemanusiaan. indonesia kn sgt <sup>2</sup> luas, kaya 😊😊😊	y sdh tampung semua demi kemanusiaan. indonesia kn sgt <sup>2</sup> luas, kaya 😊😊😊

##### 4.3.2 Cleansing

Komentar akan dicek apakah terdapat simbol seperti (:/\[!?!)\*&^%\$#@!), angka 1234567890, *link url* seperti *https* atau *www*, karakter khusus dan *mention (@)*. Jika ada maka akan dihapus dan hanya menyisakan teks. Proses ini dilakukan agar teks yang dihasilkan bersih dari gangguan (*noise*). Hasilnya dalam Tabel 4.

Tabel 4. Hasil *Cleansing*

Hasil Case Folding	Cleansing
y sdh tampung semua demi kemanusiaan. indonesia kn sgt <sup>2</sup> luas, kaya 😊😊😊	y sdh tampung semua demi kemanusiaan indonesia kn sgt luas kaya

##### 4.3.3 Tokenizing

Komentar yang masih berbentuk kalimat, akan dipisahkan menjadi satuan kata, hal ini bertujuan untuk memperkecil pola pencarian kata pada tahap *stopwords* dan *stemming*. Hasilnya pada Tabel 5.

Tabel 5. Hasil *Tokenizing*

Hasil Cleansing	Tokenizing
y sdh tampung semua demi kemanusiaan indonesia kn sgt luas kaya	'y' 'sdh' 'tampung' 'semua' 'demi' 'kemanusiaan' 'indonesia' 'kn' 'sgt' 'luas' 'kaya'

##### 4.3.4 Word Normalization

Komentar akan dinormalisasikan dari kata-kata yang memiliki singkatan, kata asing, dan kata gaul. kamus kata digunakan untuk membandingkan kata di *dataset* dengan 375 kamus kata singkatan dan kata normal. Hasilnya pada Tabel 6.

Tabel 6. Hasil *Word Normalization*

Hasil Tokenizing	Word Normalization
'y' 'sdh' 'tampung' 'semua' 'demi' 'kemanusiaan' 'indonesia' 'kn' 'sgt' 'luas' 'kaya'	'iya' 'sudah' 'tampung' 'semua' 'demi' 'kemanusiaan' 'indonesia' 'kan' 'sangat' 'luas' 'kaya'

#### 4.3.5 Stopwords Remove

komentar akan dihilangkan dari kata yang tidak memiliki makna seperti kata penghubung dan keterangan, peneliti menggunakan *library nltk* untuk mengambil *stopwords* yang menyimpan 758 kata bahasa Indonesia. Hasilnya pada Tabel 7.

Tabel 7. Hasil *Stopwords Remove*

Hasil <i>Word Normalization</i>	<i>Stopwords Remove</i>
'iya' 'sudah' 'tampung' 'semua' 'demi'	'tampung' 'semua' 'demi'
'kemanusiaan' 'indonesia' 'kan' 'sangat' 'luas'	'kemanusiaan' 'indonesia' 'luas'
'kaya'	'kaya'

#### 4.3.6 Stemming

Komentar akan dihilangkan dari imbuhan paa setiap kata, bertujuan agar kata yang akan dihasilkan nantinya merupakan kata yang baku yang tidak terindikasi imbuhan didalamnya. Peneliti menggunakan *library Sastrawi* untuk menghilangkan imbuhan secara otomatis. Hasilnya pada Tabel 8.

Tabel 8. Hasil *Stemming*

Hasil <i>Stopwords Remove</i>	<i>Stemming</i>
'tampung' 'semua' 'demi' 'kemanusiaan'	'tampung' 'semua' 'demi' 'manusia'
'indonesia' 'luas' 'kaya'	'indonesia' 'luas' 'kaya'

#### 4.4. Klasifikasi Naïve Bayes Classifier

*Dataset* yang telah melewati tahap *text preprocessing* dan pembobotan TF-IDF dilanjutkan ke proses pembelajaran (*learning*) menggunakan *Naïve Bayes Classifier*. peneliti menggunakan *library sklearn*. Proses diawali dengan menginputkan data teks dan label yang di masukan kedalam *x* dan *y*. Selanjutnya, diberikan *text\_size* 0.3 yang mengindikasikan data uji yang akan digunakan sebesar 30% dan dibuat *x\_train* yang menyimpan teks data latih, *x\_test* yang menyimpan teks data uji, *y\_train* yang menyimpan label data latih, dan *y\_test* yang menyimpan label data uji. Setelah itu, untuk data latih pada *x\_train* dan *y\_train* dimasukan kedalam *MultinomialNB* untuk dilakukan pembelajaran (*learning*) dan disimpan pada *model\_naive\_bayes*, terakhir dilakukan prediksi menggunakan data uji didalam *x\_test* dan hasil pengujiannya disimpan didalam prediksi. Untuk proses ini Gambar 5.

```
x = seleksi
y = y
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.3, random_state = 1)
model_naive_bayes = MultinomialNB().fit(x_train, y_train)
prediksi = model_naive_bayes.predict(x_test)
```

Gambar 5. Klasifikasi *Naïve Baye Clasifier*

#### 4.5. Pengujian Model

Model yang telah dibuat akan diuji dengan tujuan agar peneliti dapat mengetahui persentase ketepatan dan keakuratan dalam mengklasifikasikan data. Berdasarkan 7.547 data, 70% data latih berjumlah 5.285 data, dan 30% data uji berjumlah 2.265 data. Nantinya, 2.265 data uji akan di uji apakah label yang di prediksi sesuai berdasarkan label aslinya.

```
TRAINING
Komentar Training : 5282 data
Sentimen Training : 5282 data
TESTING
Komentar Testing : 2265 data
Sentimen Testing : 2265 data
```

Gambar 6. Pembagian Data

Berdasarkan model analisis sentimen *Naïve Bayes Classifier*, diperoleh nilai akurasi sebesar 0.9395143487858719 atau 93% dari *confusion matrix*. Selain itu, dari total 2.265 data uji, 2101 dikategorikan sentimen Negatif dan 27 sentimen Positif. Namun, diketahui 137 yang seharusnya sentimen Positif justru diprediksi menjadi Negatif. Hasil prediksi pada Gambar 7.

```
Accuracy : 0.9395143487858719
Precision : 1.00
Recall : 0.166341631160342

Confusion Matrix
[[2101  0]
 [ 137  27]]

Classification Report
precision    recall  f1-score   support

0           0.94         1.00         0.97       2101
1           1.00         0.16         0.28         164

accuracy          0.94         0.94       2265
macro avg         0.97         0.58         0.63       2265
weighted avg      0.94         0.94         0.92       2265
```

Gambar 7. Hasil *Confusion Matrix* klasifikasi

#### 4.6. Evaluasi Model

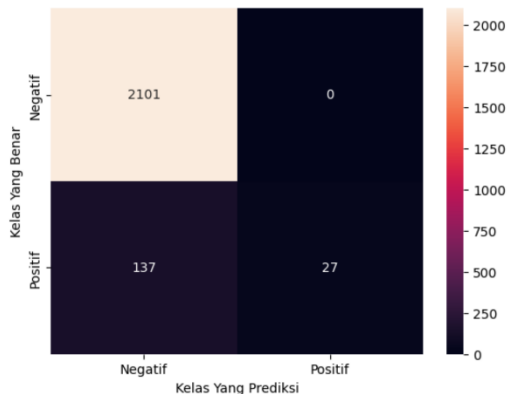
Peneliti ingin mengetahui performa dari model tersebut. Maka dilakukan perhitungan performa menggunakan *Confusion Matrix* 3x3 yang terdiri dari *Accuracy*, *Precision*, *Recall* dan *f1 Score*. Diketahui sebelumnya dari tahap pengujian model didapatkan hasil *confusion matrix* 3x3 pada Tabel 9.

Tabel 9. Hasil *Confusion Matrix*

		False / Predicted Class	
		Negatif	Positif
True Class	Negatif	2101	0
	Positif	137	27

Untuk mempermudah pembaca memahami hasil klasifikasi pada *Confusion Matrix* Peneliti juga memvisualisasikannya kedalam Gambar 8.

Confusion Matrix - Analisis Sentimen Irwan Gunawan



Gambar 8. Visualisasi *Confusion Matrix*

Kemudian peneliti melakukan perhitungan untuk menghitung performa dari model yang telah dibuat mengacu dari tabel *confusion matrix* diatas. Sehingga, hasil perhitungan nilai akurasi secara manual disajikan pada persamaan berikut ini:

$$\begin{aligned}
 \text{Accuracy} &= \frac{TP+TN}{TP+TN+FP+FN} \times 100\% & \text{Recall} &= \frac{27}{27+137} \times 100\% \\
 &= \frac{27+2101}{27+2102+0+137} \times 100\% & &= \frac{27}{164} \times 100\% \\
 &= \frac{2128}{2.265} \times 100\% & &= 0,164634146 \times 100\% \\
 &= 0,939514349 \times 100\% & &= 16\% \\
 &= 93\% & & \\
 \\ 
 \text{Precision} &= \frac{27}{27+0} \times 100\% & \text{f1 Score} &= 2 \times \frac{100 \times 16}{100+16} \\
 &= \frac{27}{27} \times 100\% & &= 2 \times \frac{1600}{116} \\
 &= 100\% & &= 27\%
 \end{aligned}$$

Berdasarkan perhitungan yang dilakukan, disimpulkan bahwasanya skor *accuracy* sebesar 93%, *precision* 100%, *recall* 16%, dan *f1 score* 27%. Dengan memperhatikan hal tersebut. Algoritma *Naïve Bayes Classifier* yang digunakan untuk mengembangkan model klasifikasi dalam analisis sentimen pada penelitian ini dapat dikategorikan baik dalam mengklasifikasikan.

## 5. KESIMPULAN

Berdasarkan penelitian ini diperoleh kesimpulan diantaranya *Naïve Bayes Classifier* dapat diterapkan dalam mengklasifikasikan data opini masyarakat terkait topik yang dibahas di *platform youtube*, diketahui sentimen negatif lebih unggul berjumlah 6956 data dibandingkan dengan sentimen negatif yang hanya berjumlah 591 data. Dengan begitu, masyarakat indonesia tidak setuju atau kurang baik memberikan pendapat terhadap tindakan pemerintah indonesia untuk menampung sementara pengungsi etnis rohingya. Selain itu, pada penelitian ini menunjukan jika *Naïve Bayes Classifier* dapat dikategorikan baik dalam mengklasifikasikan karena mendapatkan skor *accuracy* sebesar 93%, *precision* 100%, *recall* 16%, dan *f1 score* 27%. dan dengan memperhatikan *Confusion Matrix* bahwasanya data uji yang berhasil diklasifikasikan kedalam sentimen negatif berjumlah 2101 data dan data uji yang berhasil diklasifikasikan kedalam sentimen positif berjumlah 27 data, serta data uji yang salah diklasifikasikan berjumlah 137 data.

Adapun saran dari peneliti diantaranya, penambahan dataset agar bisa lebih banyak seperti jutaan atau miliaran data serta diambil dari berbagai *platform* seperti instagram, tiktok, facebook dan sebagainya untuk memperkaya dataset. Selain itu, disarankan untuk menambahkan kamus kata pada tahapan *word normalization*, baik itu kamus bahasa asing, gaul, serta singkatan

## UCAPAN TERIMA KASIH

Terima kasih kepada Allah Swt atas karunianya sehingga penulis mendapatkan berkah dalam menyelesaikan penelitian ini, kepada Bapak Dr. Mhd Furqan, S.Si, S.H M.Comp S.c selaku Dosen Pembimbing, kepada Bapak Ilka Zufria, M.Kom selaku Dosen Penguji, kepada Orang Tua Penulis Bapak Sriadi S.H dan Ibunda Sriani, dan kepada Raisa Azhari selaku sahabat yang terus kebersamai, memberikan support selama penulis melaksanakan penelitian ini.

#### REFERENSI

- [1] N. T. Pramita and M. S. Zahidi, "Analisis Peranan Indonesia Dalam Pemberian Bantuan Kemanusiaan Terhadap Etnis Rohingya Pada Era Jokowi," *Rev. Int. Relations*, vol. 5, no. 1, pp. 84–106, 2023.
- [2] D. R. Rameliah and U. W. Sagena, "Krisis Keamanan Manusia Di Myanmar Dan Penerapan Konsep Responsibility To Protect," *J. Pendidik. Tambusai*, vol. 7, no. 3, pp. 27979–27986, 2023.
- [3] F. Adha, "Penanganan Pengungsi Di Indonesia Melalui Kerja Sama Kemenkumham Dengan International Organizatio Of Migration," *Popul. J. Sosia dan Hum.*, vol. 8, no. 2, pp. 123–133, 2023.
- [4] T. Listiarani, "Analisis Kebijakan Luar Negeri Indonesia Dalam Menerima Pengungsi Rohingya Di Indonesia," *J. PIR Power Int. Relations*, vol. 5, no. 1, pp. 19–32, 2020.
- [5] N. Meilani, M. Furqan, and Suhardi, "Analisis Sentimen Pengguna Aplikasi BSI Mobile Akibat Ransomware Menggunakan Algoritma Support Vector Machine," *Infotech J. Inform. Teknol.*, vol. 5, no. 1, pp. 42–51, 2024.
- [6] F. Hermawan, "Analisis Minat Masyarakat Pengguna Platform Youtube Sebagai Media Komunikasi Digital Masa Kini," *J. Manaj.*, vol. 14, no. 3, pp. 564–573, 2022.
- [7] M. Furqan, Sriani, and S. M. Sari, "Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia," *Techno.Com*, vol. 21, no. 1, pp. 52–61, 2022.
- [8] M. Furqan and A. F. A. Nasir, "Big Data Approach To Sentiment Analysi In Machine Learning-based Microblogs: Perspectives Of Religious Moderation Public Policy In Indonesia," *J. Appl. Eng. Technol. Sci.*, vol. 5, no. 2, 2024.
- [9] H. Setiawan and I. Zufria, "Analisis Sentimen Pembatalan Indonesia Sebagai Tuan Rumah Piala Dunia FIFA U-20 Menggunakan Naive Bayes," *J. Media Inform. Budidarma*, vol. 7, no. 3, pp. 1003–1012, 2023.
- [10] S. Adelia, E. Milanda, J. Santari, D. T. Kesuma, E. Silvia, and F. Kurniawan, "Analisis Sentimen Belajar Programming Pada Media Sosial Youtube Menggunakan Algoritma Klasifikasi Naive Bayes," *J. Inf. Technol. Ampera*, vol. 4, no. 3, 2023.
- [11] S. Mulyani and R. Novita, "Implementasi Algoritma Naive Bayes Clasifier Untuk Klasifikasi Sentimen Masyarakat Tentang Depresi Pada Youtube," *J. Tek. Inform.*, vol. 3, no. 5, pp. 1355–1361, 2022.
- [12] Z. N. Aulia, G. K. Jati, and I. Santoso, "Analisis Sentimen Tanggapan Public Mengenai E-Tilang Melalui Media Sosial Youtube Menggunaka Algoritma Naive Bayes," *J. Ikraith Inform.*, vol. 7, no. 2, 2023.
- [13] W. Wahyuni, "Analisis Sentimen terhadap Opini Feminisme Menggunakan Metode Naive Bayes," *J. Inform. Ekon. Bisnis*, vol. 4, no. 4, pp. 148–153, 2022.
- [14] B. Laurensz and E. Sedyono, "Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19 (Analysis of Public Sentiment on Vaccination in Efforts to Overcome the Covid-19 Pandemic)," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 2, pp. 118–123, 2021.
- [15] R. K. Indradipradana and F. Haridha, "Kebijakan Luar Negeri Indonesia: Studi Kasus Penerimaan Pengungsi Rohingya Asal Myanmar Tahun 2020-2022," *Indones. Perspect.*, vol. 8, no. 2, pp. 211–236, 2023.
- [16] M. Hamdi, H. Maulidia, and H. Firlana, "Fenomena Pencari Suaka dan Pengungsi Etnis Rohingya Di Indonesia (Studi Kasus Penangan Rohingya Pada Provinsi Aceh)," *J. Ilm. Kaji*.

*Keimigrasian*, vol. 6, no. 1, pp. 55–71, 2023.

- [17] A. Mahadevi, F. Arindah, and J. M. Sandya, "Implementasi Hak Asasi Manusia Internasional Dalam Pemenuhan Asas Membership Oleh Myanmar Kepada Etnis Rohingya," *J. Polit. Democr. Stud.*, vol. 3, no. 2, pp. 142–157, 2022.
- [18] J. R. Prabowo, Akim, and A. Sudirman, "Peran Tentara Nasional Indonesia Dalam Menanggulangi Pengungsi Dari Luar Negeri: Studi Kasus Pengungsi Rohingya Di Indonesia (2015-2020)," *Aliansi J. Polit. Keamanan dan Hubunga Int.*, vol. 1, no. 2, pp. 99–111, 2022.
- [19] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform.*, vol. 5, no. 2, pp. 697–711, 2021.
- [20] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi," *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.
- [21] A. A. A. Sumanjaya, Indriati, and A. Ridok, "Analisis Sentimen Data Tweets Terhadap Penanganan Covid-19 Di Indonesia Menggunakan Metode Naive Bayes Dan Pemilihan Kata Bersentimen Menggunakan Lexicon Based," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 4, pp. 1865–1872, 2022.
- [22] M. R. Fahlevvi, "Analisis Sentimen Terhadap Ulasan Aplikasi Pejabat Pengelola Informasi dan Dokumentasi Kementerian Dalam Negeri Republik Indonesia di Google Playstore Menggunakan Metode Support Vector Machine," *J. Teknol. dan Komun.*, vol. 4, no. 1, pp. 1–13, 2022.
- [23] H. Santoso, A. Armansyah, and D. Desliani, "Analisis Sentimen Mahasiswa Terkait Pembelajaran Tatap Muka Menggunakan Metode Naive Bayes Classifier," *Techno.Com*, vol. 21, no. 3, pp. 644–654, 2022.
- [24] Supiyandi, A. Hussein, I. Gunawan, and W. L. R. Harjo, "Analisis Klasifikasi Broken Home pada Anak Menggunakan Metode Naïve Bayes Classifier," *JISKA J. Inform. Sunan Kalijaga*, vol. 8, no. 2, pp. 90–101, 2023.
- [25] N. P. G. Naraswati, D. C. Rosmilda, D. Desinta, F. Khairi, R. Damaiyanti, and R. Nooraeni, "Analisis Sentimen Publik Dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia Dengan Naive Bayes Clasifiation," *J. Sist. Inf.*, vol. 10, no. 1, pp. 222–238, 2021.
- [26] Hertati, E. Haerani, Novriyanto, and F. Syafira, "Analisis Sentimen Masyarakat Terhadap Kenaikan Biaya Haji Tahun 2023 Menggunakan Metode Naive Bayes Clasifier," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 3, pp. 1578–1584, 2023.

## KARTU BIMBINGAN SKRIPSI

P E R	PEMBIMBING I			PEMBIMBING II		
	Tgl.	Materi Bimbingan	T. Tangan	Tgl.	Materi Bimbingan	T. Tangan
III	21 Feb 2024	Rensi bab II, III	FA			
IV	28 Feb 2024	Acc. Sampre	FA			
V	19 Jun 2024	Rensi bab IV dan bab V	FA			
VI	16 Jun 2024	Revisi Jurnal dan Sajian Ke Peneliti	FA			
VII	17 Jun 2024	Rensi Jurnal dan Peneliti	FA			
VIII	15 Jun 2024	Acc Katalogium	FA			

Dibuat dengan Aplikasi Akademik Mahasiswa Fakultas SAINTER UIN-SU Cirebon

24

### KARTU BIMBINGAN SKRIPSI

Semester Ganjil/Cenap Tahun Akademik 2023/2024

Nama : <u>KENAN GUNAWAN</u>	Pembimbing I : <u>Dr. Hnd Fugan, S.Si, SH, M.Com, Sc</u>
NIM : <u>070202002</u>	Pembimbing II : <u>-</u>
Prog. Studi : <u>KeMu Komputer</u>	SK Pembimbing : <u>0.093/ST.V.3/ST.V.1/HM.03/2/2024</u>
Judul Skripsi : <u>Analisis Sistem Terhadap Tindakan Pemerintahan Indonesia Untuk Meningkatkan Semangat Pengabdian Masyarakat Masyarakat Negeri Payes Classifier</u>	

P E R	PEMBIMBING I			PEMBIMBING II		
	Tgl.	Materi Bimbingan	T. Tangan	Tgl.	Materi Bimbingan	T. Tangan
I	14 Februari 2024	Renda Judul	FA			
II	20 Februari 2024	Renda bab I	FA			

Dibuat dengan Aplikasi Akademik Mahasiswa Fakultas SAINTER UIN-SU Cirebon

23

P E R	PEMBIMBING I			PEMBIMBING II		
	Tgl.	Materi Bimbingan	T. Tangan	Tgl.	Materi Bimbingan	T. Tangan
IX						
X						

Madin, 01 April 2024  
An. Dekan  
Kampus/Program Studi  
UIN-SU Cirebon  
  
M. Yem  
NIP. 199506042015031006

Catatan: Pada Saat Bimbingan, Kartu Ini Harus Dijiwi Dan Ditandatangani Oleh Pembimbing

Dibuat dengan Aplikasi Akademik Mahasiswa Fakultas SAINTER UIN-SU Cirebon

25

## DAFTAR RIWAYAT HIDUP



### DATA PRIBADI

Nama : Irwan Gunawan  
Tempat/Tanggal Lahir : Lubuk Pakam, 05 Juli 2002  
Alamat : Jalan Buntu Gang Sapardi Nomor 351 Dusun VIII  
Desa Bandar Setia, Kec Percut Sei Tuan Kab Deli  
Serdang Prov Sumatera Utara, 20371  
Agama : Islam  
Jenis Kelamin : Laki-Laki  
Status : Belum Menikah  
No. HP/Wa : 085372941391  
Gmail : irwan05072002@gmail.com  
Nama Orang Tua : Sriadi,S.H (Ayah) dan Suriani (Ibu)

### PENDIDIKAN FORMAL

1. SD Negeri 104204 Desa Bandar Setia TA. 2008-2014
2. SMP Swasta Prayatna Medan TA. 2014-2017
3. SMA Swasta PAB 1 Medan Estate TA. 2017-2020
4. S1 Program Studi Ilmu Komputer Fakultas Sains dan Teknologi  
Universitas Islam Negeri Sumatera Utara Medan TA. 2020-2024

### PRESTASI

1. Penatara Kader Bela Negara Tingkat SLTA Kodam I/BB TA. 2019
2. Peserta Olimpiade Sains Kabupaten (OSK) Bidang Kimia TA. 2018
3. Juara 3 Lomba Cipta Puisi di UISU TA. 2019