

**Analisis Sentimen Ulasan Jembatan Repo – Repo di Google Maps  
Menggunakan Metode Naive Bayes Dengan Fitur Ekstraksi TF-IDF**

**SKRIPSI**

**Diajukan Oleh:**

**Muhammad Reifin Perdana**

**2011102441136**



**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR**

**JULI 2024**

**Analisis Sentimen Ulasan Jembatan Repo – Repo di Google Maps  
Menggunakan Metode Naive Bayes Dengan Fitur Ekstraksi TF-IDF**

**SKRIPSI**

Diajukan Sebagai Salah Satu Persyaratan Untuk Memperoleh Gelar Sarjana Teknik  
Informatika Fakultas Sains Dan Teknologi Universitas Muhammadiyah Kalimantan Timur

**Diajukan Oleh:**

**Muhammad Reifin Perdana**

**2011102441136**



**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR**

**JULI 2024**

# LEMBAR PERSETUJUAN

## LEMBAR PERSETUJUAN

### Analisis Sentimen Ulasan Jembatan Repo – Repo di Google Maps Menggunakan Metode Naive Bayes Dengan Fitur Ekstraksi TF-IDF SKRIPSI

Diajukan Oleh:

Muhammad Reifin Perdana  
2011102441136

Disetujui untuk diujikan

Pada tanggal 27 Juni 2024


Pembimbing



Rudiman, S.Kom, M.Sc  
NIDN 1105068202

Mengetahui,

Koordinator Tugas Akhir/Skripsi/Tesis/Disertasi



Abdul Rahim, S.Kom., M.Cs  
NIDN 0009047901

# LEMBAR PENGESAHAN

## LEMBAR PENGESAHAN

**Analisis Sentimen Ulasan Jembatan Repo – Repo di Google Maps  
Menggunakan Metode Naive Bayes Dengan Fitur Ekstraksi TF-IDF  
SKRIPSI**

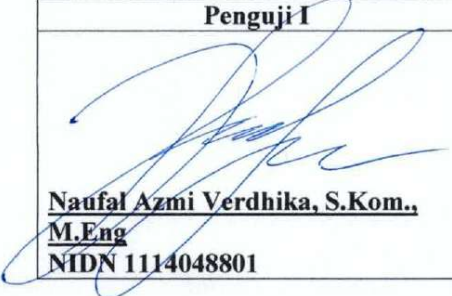
Diajukan oleh:

**Muhammad Reifin Perdana**

2011102441136

Diseminarkan dan Diujikan

Pada Tanggal 3 Juli 2024


Penguji I	Penguji II
 <b><u>Naufal Azmi Verdhika, S.Kom., M.Eng</u></b> NIDN 1114048801	 <b><u>Rudiman, S.Kom., M.Sc</u></b> NIDN 1105068202

Mengetahui,

Ketua

Program Studi Teknik Informatika



  
**Arbansyah, S.Kom., M.TI**

NIDN 1118019203

# PERNYATAAN KEASLIAN PENELITIAN

## PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Reifin Perdana

NIM : 2011102441136

Program Studi : S1 Teknik Informatika

Judul Penelitian : Analisis Sentimen Ulasan Jembatan Repo-Repo di Google Maps  
Menggunakan Metode Naïve Bayes Dengan Fitur Ekstraksi TF-IDF

Menyatakan bahwa Skripsi yang saya tulis ini benar-benar hasil karya saya sendiri, dan bukan merupakan hasil plagiasi/falsifikasi/fabrikasi baik sebagian atau seluruhnya.

Atas pernyataan ini, saya siap menanggung resiko atau sanksi yang dijatuhkan kepada saya apa bila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam skripsi saya ini, atau klaim dari pihak lain terhadap keaslian karya saya ini.

Samarinda 30 Juni 2024

Yang membuat pernyataan



Muhammad Reifin Perdana

NIM: 2011102441136

## **Abstrak**

Penelitian ini menganalisis sentimen ulasan Jembatan Repo-Repo di Google Maps menggunakan metode Naïve Bayes dengan fitur ekstraksi TF-IDF. Tujuan pada penelitian kali ini untuk mengetahui sentimen masyarakat terhadap keberadaan Jembatan Repo-repo serta mengukur akurasi metode Naïve Bayes dalam mengklasifikasikan sentimen menjadi positif, netral dan negatif. Data yang digunakan dalam penelitian ini sebanyak 503 ulasan berbahasa Indonesia yang diambil dari Google Maps.

Proses penelitian meliputi pengumpulan data, labelling data, pra-pemrosesan data, pembobotan katan menggunakan TF-IDF, dan klasifikasi menggunakan Naïve Bayes. Hasil penelitian menunjukkan bahwa metode Naïve Bayes dengan fitur ekstraksi TF-IDF memiliki akurasi sebesar 58%. Dalam mengklasifikasikan sentimen ulasan masyarakat menghargai keberadaan jembatan tersebut, sentimen netral menunjukkan pandangan yang tidak condong ke arah positif maupun negatif, sementara sentimen negatif menunjukkan perlunya perbaikan infrastruktur agar lebih menarik bagi wisatawan.

Kata Kunci: Jembatan Repo-Repo, Ulasan, Analisis Sentimen, Naïve Bayes, TF-IDF

## ***Abstract***

*This research analyzes the sentiment of Repo-Repo Bridge review on Google Maps using the Naïve Bayes method with TF-IDF extraction features. The aim of this research is to determine public sentiment towards the existence of the Repo-Repo Bridge and measure the accuracy of the Naïve Bayes method in classifying sentiment into positive, neutral and negative. The data used in this research were 503 Indonesia language reviews taken from Google Maps.*

*The research process includes data collection, data labeling, data pre-processing, weighting using TF-IDF, and classification using Naïve Bayes. The research results show that the Naïve Bayes method with the TF-IDF extraction feature has an accuracy of 58%. In classifying the sentiment of public reviews appreciating the existence of the bridge, neutral sentiment indicates views that are neither positive or negative, while negative sentiment indicates the need to improve the infrastructure to make it more attractive to tourists.*

*Keywords: Repo-Repo Bridge, Review, Sentiment Analysis, Naïve Bayes, TF-IDF*

## PRAKATA

Puji dan syukur saya panjatkan kepada Allah SWT. Atas ridho-Nya saya dapat menyelesaikan penyusunan skripsi ini. Adapun judul skripsi yang saya ajukan adalah “Analisis Sentimen Ulasan Google Maps Pada Jembatan Repo-Repo Menggunakan Metode Naive Bayes Dengan Fitur Ekstraksi TF-IDF”

Skripsi ini diajukan untuk memenuhi syarat kelulusan mata kuliah Skripsi di Fakultas Sains dan Teknologi, Univeritas Muhammadiyah Kalimantan Timur. Tidak dapat disangka bahwa butuh usaha yang keras dalam penyelesaian pengerjaan skripsi ini. Namun, karya ini tidak akan selesai tanpa orang-orang tercinta di sekeliling saya yang mendukung dan membantu. Terima kasih saya sampaikan kepada:

- a) Bapak Dr. Muhammad Musiyam, M.T. Selaku Rektor Univeritas Muhammadiyah Kalimantan Timur yang telah memberi kesempatan pada peneliti untuk menempuh pendidikan di Universitas Muhammadiyah Kalimantan Timur
- b) Bapak Arbansyah, S.Kom., M.TI. Selaku Ketua Prodi Teknik Informatika yang telah memberikan kesempatan pada penelitian ini dalam menempuh pendidikan di Teknik Informatika.
- c) Bapak Rudiman, M.Sc. Selaku Dosen Pembimbing saya yang telah membagi ilmu, bimbingan dan arahan selama proses perkuliahan.
- d) Orang tua saya yang selalu mendukung, memotivasi dan mendoakan saya sehingga proses pengerjaan proposal ini selesai sampai akhir.
- e) Pacar saya yang bernama Rahmadiani Novita yang telah membantu saya, selalu mendukung saya, selalu menyemangati saya dan mendoakan saya sampai akhirnya saya bisa mengerjakan proposal ini sampai selesai.



f) Teman-teman kuliah saya Ikhsan dan Rayhan yang sudah membantu saya dan mendukung saya, bersama-sama menyelesaikan skripsi ini sampai akhir.

Semoga segala kebaikan dan pertolongan semuanya mendapat berkah dari Allah SWT. Dan akhirnya penulis menyadari bahwa skripsi ini tidak luput dari berbagai kekurangan, karena keterbatasan ilmu yang saya miliki. Untuk itu dengan kerendahan hati saya mengharapkan saran dan kritik yang sifatnya membangun dari semua pihak demi membangun laporan penelitian ini,

Samarinda, .... Juli 2024

Muhammad Reifin Perdana

# DAFTAR ISI

<b>LEMBAR PERSETUJUAN</b> .....	ii
<b>LEMBAR PENGESAHAN</b> .....	iii
<b>PERNYATAAN KEASLIAN PENELITIAN</b> .....	iv
<b>Abstrak</b> .....	v
<b>Abstract</b> .....	vi
<b>PRAKATA</b> .....	vii
<b>DAFTAR ISI</b> .....	ix
<b>DAFTAR GAMBAR</b> .....	xi
<b>DAFTAR TABEL</b> .....	xii
<b>LAMPIRAN</b> .....	xiii
<b>BAB I PENDAHULUAN</b> .....	1
<b>1.1 Latar Belakang</b> .....	1
<b>1.2 Rumusan Masalah</b> .....	4
<b>1.3 Tujuan Penelitian</b> .....	4
<b>1.4 Manfaat Penelitian</b> .....	4
<b>BAB II METODOLOGI PENELITIAN</b> .....	6
<b>2.1 Obyek Penelitian</b> .....	6
<b>2.2 Alat dan Bahan</b> .....	7
<b>2.3 Prosedur Penelitian</b> .....	7
<b>2.3.1 Pengumpulan Data</b> .....	9
<b>2.3.2 Pre-Processing</b> .....	9
<b>2.3.3 Pembobotan Kata (TF-IDF)</b> .....	10
<b>2.3.4 Split Data</b> .....	11
<b>2.3.5 Klasifikasi</b> .....	11
<b>2.3.6 Evaluasi</b> .....	12
<b>2.3.7 Jadwal Penelitian</b> .....	13
<b>BAB III HASIL ANALISIS DAN PEMBAHASAN</b> .....	16
<b>3.1 Pengumpulan Data</b> .....	16
<b>3.2 Pre-Processing</b> .....	17
<b>3.2.1 Case Folding</b> .....	17
<b>3.2.2 Cleaning</b> .....	18
<b>3.2.3 Tokenizing</b> .....	19

3.2.4	Stopword Removal .....	20
3.2.5	Stemming.....	20
3.3	Visualisasi.....	21
3.4	Pembobotan Kata (TF-IDF) .....	22
3.5	Evaluasi .....	23
<b>BAB IV</b>	<b>PENUTUP</b> .....	<b>31</b>
4.1	Kesimpulan .....	31
4.2	Saran.....	32
	Daftar Pustaka .....	33
	Daftar Lampiran .....	36
	<b>RIWAYAT HIDUP</b> .....	<b>47</b>

## DAFTAR GAMBAR

<b>Gambar</b>	<b>Halaman</b>
<b>Gambar 2. 1</b> Alur Penelitian .....	8
<b>Gambar 2. 2</b> Hasil Pengumpulan Data.....	9
<b>Gambar 3. 1</b> Hasil Visualisasi.....	22
<b>Gambar 3. 2</b> Pembobotan Kata (TF-IDF) .....	23
<b>Gambar 3. 3</b> Hasil Akurasi .....	24

## DAFTAR TABEL

<b>Tabel</b>	<b>Halaman</b>
<b>Tabel 2. 1</b> Jadwal Penelitian .....	13
<b>Tabel 3. 1</b> Hasil Pengumpulan Data .....	16
<b>Tabel 3. 2</b> Hasil Case Folding.....	18
<b>Tabel 3. 3</b> Hasil Cleaning .....	18
<b>Tabel 3. 4</b> Hasil Tokenizing.....	19
<b>Tabel 3. 5</b> Hasil Stopword Removal.....	20
<b>Tabel 3. 6</b> Hasil Stemming .....	21

## LAMPIRAN

Lampiran	Halaman
L1 Pengumpulan Data.....	36
L2 Install Library .....	38
L3 Cek Informasi Dataset .....	39
L4 Cek Jumlah Sentimen Positif, Negatif, dan Netral .....	39
L5 Membaca Dataset Kolom Review dan Sentimen .....	39
L6 Pre-Processing Data.....	39
L7 Visualisasi Wordcloud.....	41
L8 Pembobotan Kata (TF-IDF).....	42
L9 Split Data dan Evaluasi Confusion Matrix Metode Naïve Bayes.....	42
L10 Surat Ijin Penelitian .....	44
L11 Kartu Bimbingan .....	45

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Secara geografis, Kabupaten Kutai Kartanegara memiliki luas wilayah 25.988,08 km<sup>2</sup> atau 20,41% dari seluruh wilayah daratan Provinsi Kalimantan Timur (Purwanto, 2022). Kutai Kartanegara dibagi dalam 17 wilayah kecamatan dan 255 desa/kelurahan dengan jumlah penduduk mencapai 626.286 jiwa sensus 2010 dan pada tahun 2020 bertambah menjadi 734.485 jiwa (Novemyleo, 2021).

Kepala Dinas Pariwisata (Dispar) Kukar, Slamet Hadirahardjo mengatakan, Dispar Kukar memiliki sejumlah pekerjaan rumah yang harus dihadapi. Salah satunya adalah mengembangkan objek wisata yang berpotensi di Kukar (Maris, 2023). Kutai Kartanegara memiliki 7 destinasi wisata yaitu, Waduk Panji Sukarame, Ladaya, Museum Kayu, Museum Mulawarman, Pulau Kumala, Jembatan Repo – Repo, dan Taman Kota Raja (Wahyu, 2021).

Pada saat libur akhir pekan pencarian rute alternatif tercepat ke sebuah tempat wisata menjadi masalah yang sangat penting karena banyaknya jumlah wisatawan yang ingin pergi ke tempat wisata untuk menghabiskan waktu berliburnya (Puja Kekal et al., 2021). Pariwisata yang ada di Kabupaten Kutai Kartanegara mempunyai dua pesona yaitu pesona budaya dan pesona alam. Salah satu destinasi wisata yang ada di Kabupaten Kutai Kartanegara ialah Pulau Kumala (Fuad Amirullah et al., 2023).

Untuk menuju ke Pulau Kumala saat ini sudah memiliki akses ke sana dengan menyebrangi sungai mahakam melalui Jembatan Repo-repo yang dalam bahasa Kutai berarti "Gembok", karena di beberapa tempat di pagar jembatan ada tempat khusus disediakan untuk dipasang (Muhammad Bambang Firdaus et al., 2021). Jembatan Repo-Repo merupakan

jembatan yang membentang sepanjang 230 meter dengan lebar hanya 3,5 meter (Anggraini & Susilo, 2022).

Dengan pesatnya perkembangan teknologi informasi, pengunjung di Jembatan Repo-Repo dapat memberikan tanggapan, komentar, maupun ulasan melalui aplikasi peta yang sering digunakan oleh masyarakat yakni Google Maps (Rizki Maulidiah et al., 2023). Analisis sentimen digunakan untuk menganalisis pendapat atau opini dari pengguna Google Maps berdasarkan ulasan (review) pengguna yang diambil dari Google Maps (Doloksaribu & Samuel, 2022).

Analisis sentimen adalah langkah penting dalam mengungkapkan apakah ulasan tersebut bersifat positif, netral dan negatif (Prakoso & Hermawan, 2023). Analisis sentimen ini juga merupakan suatu bidang studi yang mampu menganalisis opini dari seseorang, sentimen, evaluasi, atau pemerintah dalam menganalisis opini masyarakat atas suatu kejadian (Muzaki & Witanti, 2021).

Pembobotan TF-IDF (Term Frequency-Inverse Document Frequency) adalah proses transformasi data dari format data tekstual ke format data numerik untuk melakukan pembobotan pada setiap kata atau fitur. TF-IDF merupakan ukuran statistik yang digunakan untuk mengevaluasi tingkat pentingnya sebuah kata di dalam sebuah dokumen. TF menggambarkan frekuensi kemunculan kata dalam tiap dokumen yang diberikan, menunjukkan seberapa penting kata tersebut di dalam dokumen tersebut. DF menggambarkan frekuensi dokumen yang mengandung kata tersebut, menunjukkan seberapa umum kata tersebut (Septian et al., 2019).

Setelah melakukan tahapan seleksi fitur, dilakukan pembagian data menjadi data training dan data testing. Training merupakan tahap penting pada algoritma Naive Bayes di mana data yang telah melalui proses pembersihan dan pra-pemrosesan disesuaikan untuk



mencapai set data yang telah ditentukan. Proses training digunakan untuk membuat model machine learning yang akan digunakan untuk klasifikasi data. Sementara itu, pengujian data testing berguna untuk mengevaluasi kinerja model yang telah dibuat menggunakan metode Naive Bayes (Isnain et al., 2021).

Naive Bayes Classifier merupakan metode klasifikasi dengan menggunakan probabilitas sederhana yang menerapkan Teorema Bayes dengan asumsi tingkat ketidaktergantungan (independensi) yang tinggi antara kondisi atau kejadian. (Kiedrowsky & Andrianingsih, 2023). Dalam Naïve bayes, probabilitas dihitung untuk mengaitkan suatu kelas dengan atribut yang ada, sehingga dapat menentukan kelas yang memungkinkan (Nurian, 2023).

Pada beberapa penelitian sebelumnya banyak peneliti yang menggunakan metode Naïve Bayes untuk melakukan klasifikasi sentimen diantaranya peneliti pertama seperti yang dilakukan oleh (Putu et al., 2021), analisis sentimen dan pemodelan topik pariwisata lombok menggunakan algoritma Naïve Bayes. Akurasi rata-rata yang diperoleh adalah sebesar 92%. Berdasarkan dari penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma Naïve Bayes terbukti mampu diterapkan dengan baik untuk melakukan analisis sentimen dengan didukung oleh beberapa metode *preprocessing*, pembobotan *term* menggunakan TF-IDF dan parameter tuning algoritma.

Berdasarkan uraian di atas, bertujuan untuk melakukan analisis sentimen terhadap opini publik mengenai adanya Jembatan Repo-Repo di Tenggarong dengan menerapkan metode Naïve Bayes. Tujuannya adalah menggunakan hasil analisis sentimen untuk mengidentifikasi arah kualitas adanya Jembatan Repo – Repo. Hasil yang positif menandakan adanya jembatan tersebut sangat dihargai masyarakat atas kinerja Pemerintah Kabupaten Kutai Kartanegara, tetapi ada juga opini negatif menunjukkan bahwa ada yang harus diperbaiki dan ditingkatkan lagi infrastrukturnya agar wisatawan berminat berkunjung ke tempat wisata tersebut.

## **1.2 Rumusan Masalah**

Dari latar belakang tadi, maka rumusan masalahnya yaitu:

- (i) Bagaimana hasil akurasi analisis sentimen dalam menggunakan metode Naïve Bayes Classifier digunakan dalam sentimen terhadap Jembatan Repo-Repo di Google Maps?
- (ii) Bagaimana sentimen pengguna Google Maps terhadap adanya Jembatan Repo-Repo menggunakan metode Naïve Bayes?

## **1.3 Tujuan Penelitian**

Penelitian ini memiliki beberapa tujuan yaitu :

- (i) Mengetahui tingkat akurasi yang diperoleh menggunakan metode Naïve Bayes dalam memprediksi opini dari masyarakat adanya Jembatan Repo-Repo di Google Maps.
- (ii) Mengetahui sentimen positif, netral dan negatif mengenai Jembatan Repo-Repo di Google Maps.

## **1.4 Manfaat Penelitian**

Dengan adanya tujuan penelitian diatas, maka diharapkan mampu membantu pihak pemerintah daerah yang manfaatnya dapat diperoleh dari penelitian ini adalah :

- (i) Dengan mengetahui tingkat akurasi menggunakan metode Naïve Bayes dalam mengklasifikasi penggunaan Google Maps terhadap ulasan di Jembatan Repo-Repo.
- (ii) Memberikan pengalaman dalam melakukan penelitian yang melibatkan analisis sentimen dan penggunaan metode Naïve Bayes dengan fitur ekstraksi TF-IDF.
- (iii) Memberikan data untuk Pemerintah Kabupaten Kutai Kartanegara untuk meningkatkan kualitas berdasarkan masukan yang diterima dari pengguna Google Maps.

## **BAB II**

### **METODOLOGI PENELITIAN**

#### **2.1 Obyek Penelitian**

Obyek pada penelitian ini adalah menganalisis sentimen masyarakat terhadap adanya Jembatan Repo-Repo. Mengapa objek penelitian ini harus di angkat karena dapat memberikan pandangan masyarakat. Dengan mengidentifikasi harapan, kekhawatiran, dan dukungan, sehingga dapat merumuskan strategi komunikasi dan kebijakan yang lebih efektif dan efisien terhadap masyarakat yang berkunjung.

Penelitian mengenai sentimen terhadap bangunan Jembatan Repo-Repo di Google Maps memiliki dampak dalam memahami pandangan masyarakat yang berkunjung dan bisa mengambil keputusan yang lebih baik. Hasil dari analisis sentimen dapat digunakan untuk merumuskan kebijakan yang lebih responsif. Dengan mengetahui sentimen masyarakat, pemerintah dapat mengambil keputusan yang lebih bijak dan diterima oleh masyarakat yang berkunjung.

Penelitian ini dapat mendorong partisipasi masyarakat yang lebih besar dengan menunjukkan pendapat dan perasaan mereka diperhatikan. Ini dapat meningkatkan rasa memiliki dan memelihara terhadap bangunan Jembatan Repo-Repo yang ada di Kutai Kartanegara. Agar masyarakat yang dari dalam negri maupun luar negri akan mendapatkan kepuasan dan kesenangan saat melewati Jembatan.

Adapun dalam penelitian ini aplikasi sistem informasi yang digunakan dalam mengambil data sentimen merupakan aplikasi Google Maps. Mengapa tidak menggunakan

aplikasi lain atau platform media, karena aplikasi Google Maps mempunyai beberapa keunggulan dalam pengambilan data sentimen yaitu:

- (i) Sumber Data Penting: Ulasan dapat dijadikan sumber data penting untuk penelitian Neuro-Linguistic Programming(NLP), termasuk analisis sentimen.
- (ii) Akses Data: Google Maps adalah salah satu aplikasi yang sangat berguna untuk melakukan pengumpulan data guna menganalisis sentimen masyarakat.
- (iii) Penggunaan Algoritma: Setelah mengumpulkan data, Google Maps memungkinkan algoritma Naive Bayes untuk analisis sentimen masyarakat.

## **2.2 Alat dan Bahan**

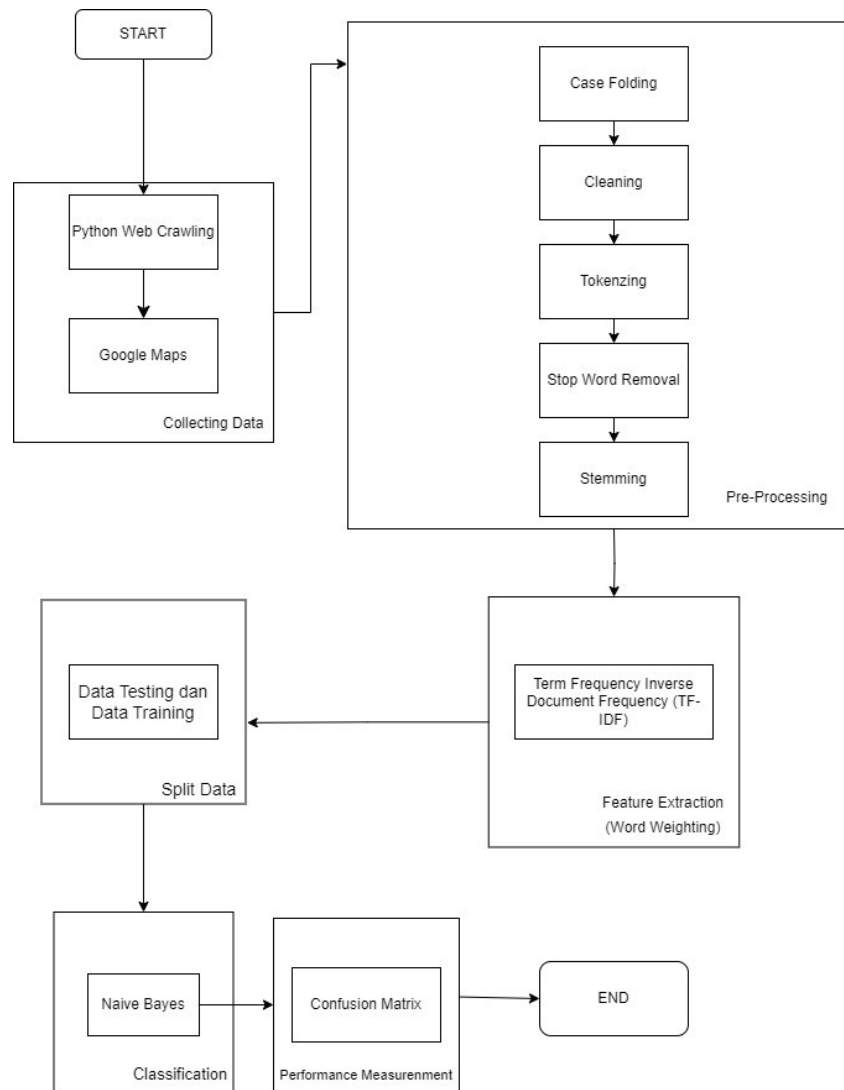
Alat yang harus disiapkan pada penelitian ini berupa tools yang telah disediakan Google yaitu Google Colab atau Google Colaboratory. Salah satu tools pengembangan machine learning yang dinilai user-friendly. Google Colab adalah fasilitas bahasa pemrograman Python dari Google, fasilitas ini memungkinkan para pengguna Google Colab untuk saling berkolaborasi dalam membuat pemrograman dengan menggunakan python.

Bahan yang harus dipersiapkan yaitu berupa data Google Maps, Label Sentimen, Metode Analisis yaitu penelitian ini menggunakan metode Naive Bayes. Kemudian melakukan Preprocessing Data menggunakan tools Google Colab, melakukan tahap pembobotan kata. Setelah tahap-tahap tadi sudah dilakukan peneliti melakukan Evaluasi Model seperti Confusion Matrix untuk mengukur kinerja model yang dilatih.

## **2.3 Prosedur Penelitian**

Dalam penelitian ini, dilakukan teknik web scraping pada Google Maps menggunakan bahasa pemrograman Python kemudian mengambil link Jembatan Repo-Repo pada Google Maps untuk proses analisis lebih lanjut. Metode yang dipakai dalam penelitian ini bertujuan

untuk melakukan analisis sentimen terhadap ulasan pada Google Maps dengan menggunakan metode Naive Bayes. Diagram alur penelitian dapat dilihat pada Gambar 2.1



**Gambar 2.1** Alur Penelitian

Pada penelitian Metode ini terdiri dari beberapa tahapan utama, mulai dari Pengumpulan data melalui scrapping menggunakan Python, Pre-processing data seperti case folding, cleaning, tokenizing, stop word removal, dan stemming, Pembobotan kata menggunakan metode TF-IDF untuk memberikan nilai pada setiap kata berdasarkan frekuensi dan kepentingannya, split data atau pembagian data menjadi data training dan data testing,

klasifikasi menggunakan metode Naive Bayes, serta Evaluasi kinerja model menggunakan confusion matriks. Setiap tahapan akan dijelaskan secara rinci dengan memaparkan teori, konsep, dan algoritma yang digunakan.

### 2.3.1 Pengumpulan Data

Dalam penelitian ini pengambilan data dilakukan melalui proses crawling data ulasan Google Maps dengan memanfaatkan library BeautifulSoup4 yang diimplementasikan dalam bahasa pemrograman Python. BeautifulSoup4 digunakan sebagai alat untuk mengumpulkan data ulasan dari Google berdasarkan identifikasi unik sebuah tempat (`feature_id`) dan menganalisis informasi yang terkandung dalam ulasan tersebut, seperti rating dan sentimen pengguna. Dengan menggunakan BeautifulSoup4, peneliti dapat mengekstrak data ulasan yang relevan, mengidentifikasi sentimen dari ulasan (positif, netral atau negatif), dan mengumpulkan informasi penting lainnya dari ulasan pengguna (Ulfah & Najiah, 2023). Hasil crawling akan ditampilkan pada Gambar 2.2

	name	rating	review	time	sentimen
0	Erik Nawaca Harahap	5	Bagus jembatannya tapi tutup dimalam hari, tap...	31-03-2024	Positif
1	firmansyah zeppy	5	Mantapppp,,, rawat dan jaga aset kembangkan u...	01-03-2024	Positif
2	Mita Hady	4	Jembatan penyebrangan ke pulau kumala. Pas des...	31-05-2023	Positif
3	Ros Aizam	5	Pas kesini lg ada pengecatan bagus banget liat...	02-12-2023	Positif
4	Supri Yanus	5	Pernah kesini pada 7 Maret 2019...Bisa berjala...	31-05-2022	Positif

**Gambar 2. 2** Hasil Pengumpulan Data

### 2.3.2 Pre-Processing

Preprocessing menjadi tahap awal dalam klasifikasi teks untuk mempersiapkan data teks sebelum digunakan pada proses lainnya. Pada tahap ini akan mengubah data teks menjadi bentuk yang lebih baik sehingga menghasilkan informasi teks dengan kualitas yang baik dan siap digunakan pada proses selanjutnya (Khairunnisa et al., 2021). Berikut langkah-langkah yang dilakukan pada tahap preprocessing yaitu (Puspitasari et al., 2023) :

- (i) *Case folding* merupakan tahapan untuk mengubah huruf kapital dalam dokumen menjadi standar, yaitu huruf kecil (lowercase).
- (ii) *Cleaning* merupakan tahapan untuk membersihkan karakter yang tidak dibutuhkan seperti URL, @, #, https: RT (Retweet), angka, simbol, dan emoticon.
- (iii) *Tokenizing* merupakan tahapan untuk memecah kalimat dalam dokumen menjadi kata. Sebuah proses untuk menghilangkan tanda baca, simbol, dan karakter bacaan yang tidak berharga.
- (iv) *Stopword removal* merupakan tahapan untuk menghilangkan sebuah kata informatif rendah. Stopword dilakukan jika kalimat tersebut mengandung kata umum dan tidak signifikan seperti waktu, penghubung, dan sebagainya.
- (v) *Stemming* merupakan tahapan untuk menghapus imbuhan awalan dan akhiran, sehingga menjadi bentuk kata dasar. Stemming dilakukan menggunakan library Sastrawi pada python.

### 2.3.3 Pembobotan Kata (TF-IDF)

Pembobotan kata bertujuan memberikan bobot pada setiap kata yang terdapat dalam, kalimat/dokumen yang sudah melalui tahap pre-processing. Untuk mendapatkan nilai TF-IDF, akan dilakukan perhitungan Term Frequency(TF), dan Inverse Document Frequency (IDF). Term Frequency (TF) merupakan proses untuk menghitung jumlah frekuensi suatu kata terdapat dalam keseluruhan dokumen dan memberikan nilai untuk kata tersebut (Locarso, 2022).

Penggunaan TF dapat menggunakan rumus di bawah ini:

$$tf_{ij} = \frac{f_d(i)}{\max f_d(j)}$$

TF menunjukkan dokumen (d) seberapa banyak kata (t) yang muncul. Dan terkait untuk rumus IDF yaitu:

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

N melambangkan jumlah kata dalam teks, df adalah jumlah teks yang memiliki kata t. Dengan menggabungkan TF dan IDF dalam pengerjaan dapat membantu meningkatkan performa.

Setelah mendapatkan TF dan IDF, kemudia dapat menghitung nilai TF-IDF yang merupakan hasil perkalian dari TF dan IDF. Contoh rumus yang sering digunakan sebagai berikut.

$$tfidf_{t,a} = tf_{t,a} \times idf_a$$

### 2.3.4 Split Data

Splitting data adalah metode yang digunakan untuk membagi dataset menjadi data train dan data test. Fungsi dari data train yaitu data yang akan digunakan untuk dilatih pada saat pembuatan model, dan fungsi dari data test yaitu untuk data yang akan diuji dari hasil prediksi yang dibuat berdasarkan data train. Pada tahap splitting data, peneliti memutuskan untuk melakukan percobaan dengan membagi data menjadi 90:10, 80:20, dan 70:30 dengan urutan data training dan data testing. (Ariyani et al., 2022).

### 2.3.5 Klasifikasi

Salah satu algoritma yang digunakan untuk klasifikasi teks adalah Naive Bayes. Algoritma ini menggunakan perhitungan probabilitas dan statistik dalam pengoperasiannya. Metode klasifikasi Naive Bayes digunakan untuk mengambil keputusan dengan melakukan prediksi suatu kasus berdasarkan hasil dari klasifikasi yang telah diperoleh. Pada penelitian ini, metode Naive Bayes digunakan untuk menentukan sentimen dari data Google Maps (Astari et al., 2020). Berikut ini adalah langkah – langkah rumus Naïve Bayes yaitu:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$



Di mana:

A: Hipotesis data merupakan suatu *class* spesifik.

B: Data dengan kelas yang masih belum diketahui.

$P(A|B)$ : Probabilitas hipotesis berdasar kondisi.

$P(A)$ : Probabilitas hipotesis.

$P(B|A)$ : Probabilitas berdasarkan kondisi pada hipotesis.

$P(B)$ : Probabilitas B.

### 2.3.6 Evaluasi

Evaluasi digunakan untuk melihat kinerja dari model klasifikasi yang digunakan. Metode evaluasi yang digunakan adalah confusion matrix. Berdasarkan confusion matrix pengujian model dilakukan dengan menghitung nilai accuracy, precision, recall, dan juga F-measure (Muslimin & Lusiana, 2023).

Perhitungan untuk metode confusion matrix yang digunakan yaitu nilai skor Accuracy, karena untuk mengukur seberapa sering model klasifikasi membuat prediksi yang benar, rumus berikut digunakan untuk menghitung nilai akurasi:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

Keterangan:

(i) TP = True Positive

(ii) FN = False Negative

(iii) FP = False Positive

(iv) TN = True Negative

### 2.3.7 Jadwal Penelitian

**Tabel 2. 1** Jadwal Penelitian

No	Kegiatan	Bulan/2024					
		Februari	Maret	April	Mei	Juni	Juli
Tahap Pra Penelitian							
1	Menentukan Judul						
2	Identifikasi Masalah						
3	Studi Literatur						
4	Rancangan Metode						
5	Pemilihan Studi Kasus						
6	Menyusun Proposal						
Tahap Penelitian							
1	Pengumpulan Data (Crawling)						
2	Labelling Data						
3	Pre-Processing Data						
4	Pembobotan Kata (TF-IDF)						
5	Split Data						
6	Klasifikasi Support Vector Machine (SVM)						
7	Evaluasi dan Analisis Hasil						
Tahap Akhir Penelitian							
1	Penyusunan Laporan						
2	Seminar Hasil						

## BAB III

### HASIL ANALISIS DAN PEMBAHASAN

#### 3.1 Pengumpulan Data

Pada penelitian ini, digunakan data ulasan terkait hasil ulasan yang diambil dari Google Maps. Data yang berhasil dikumpulkan sebanyak 500 ulasan, dan disimpan dalam format CSV. Proses pada pengumpulan data ini dilakukan dengan metode *crawling* selama 10 menit pada tanggal 18 April 2024, dari pukul 15:45 WITA hingga pukul 15:55 WITA. Hasil Pengumpulan data akan ditampilkan pada Tabel 3.1

**Tabel 3. 1** Hasil Pengumpulan Data

No	Nama	Rating	Review	Time
1	Erik Nawaca Harahap	5	Bagus jembatannya tapi tutup dimalam hari, tapi bisa nongkrong dan minum kopi di pinggir sungai mahakam dan bisa mancing juga..jembatan ini jalam mau kalai ke pulau bisa juga	31-03-2024
2	firmanasyah zeppy	5	Mantapppp,,, rawat dan jaga aset kembangkan untuk pariwisata dan identitas budaya, Jembatan adalah aset kebanggaan masyarakat Kab. Tenggarong	01-03-2024
3	Mita Hady	4	Jembatan penyebrangan ke pulau kumala. Pas desember 2022 kemarin ke sana sedang direnovasi, dicat kembali. Pemandangan cukup bagus... Mudah dimajukan kembali dan diberdayakan kembali pulau kumala ini, dikelola lebih baik lagi biar jadi destinasi utama lagi kalau ke tenggarong. Selain jadi pemasukan daerah, bisa jadi pusat tempat wisata budaya dan edukasi khususnya di Kaltim.	31-05-2023
4	Ros Aizam	5	Pas kesini lg ada pengecatan bagus banget liat pemandangan dari atas tapi saya ada phobia ketinggian jadi pas mau liat ke bawah tu lemes lutut	02-12-2023
5	Supri Yanus	5	Pernah kesini pada 7 Maret 2019...Bisa berjalan kaki menyeberangi jembatan repo	31-05-2022

repo ke Pulau Kumala...Pulau Kumala tempatnya bagus... Wisata yang nyaman buat perorangan maupun keluarga. Fasilitas lengkap... Ada banyak obyek wisata di dalam pulau kumala ini...tempatya teduh dengan rindangnya tanaman... Sangat layak untuk di kunjungi...

---

### 3.2 Pre-Processing

Pada tahap *Pre-Processing* dilakukan untuk mengolah data yang ada sehingga menghindari gangguan pada data-data yang tidak konsisten. Tujuannya agar hasil output dari klasifikasikan memiliki tingkat keakuratan yang tinggi dan tahapan ini dapat mengubah data yang tidak tersusun menjadi data yang tersusun rapi. Dalam text preprocessing terdapat 6 proses yaitu case folding, cleaning, tokenzing, stop word removal dan steaming.

#### 3.2.1 Case Folding

Langkah pertama *preprocessing* pada penelitian ini adalah *case folding*. Tahap case folding ini dilakukan untuk merubah karakter kata di dalam data set yang awalnya huruf besar menjadi huruf kecil. Setelah dilakukan proses pada coddingan maka didapatkan hasil data set yang sudah tidak mempunyai lagi huruf kapital. Pada Tabel 3.2 di bawah hasil sebelum dan sesudah teks yang di konveri menggunakan case folding:

**Tabel 3. 2** Hasil Case Folding

No	Case Folding
1	bagus jembatannya tapi tutup dimalam hari, tapi bisa nongkrong dan minum kopi di pinggir sungai mahakam dan bisa mancing juga..jembatan ini jalam mau kalai ke pulau bisa juga
2	mantapppp,,,, rawat dan jaga aset kembangkan untuk pariwisata dan identitas budaya, jembatan adalah aset kebanggaan masyarakat kab. tenggarong
3	jembatan penyebrangan ke pulau kumala. pas desember 2022 kemarin ke sana sedang direnovasi, dicat kembali. pemandangan cukup bagus... mudahan dimajukan kembali dan diberdayakan kembali pulau kumala ini, dikelola lebih baik lagi biar jadi destinasi utama lagi kalau ke tenggarong. selain jadi pemasukan daerah, bisa jadi pusat tempat wisata budaya dan edukasi khususnya di kaltim.
4	pas kesini lg ada pengecatan bagus banget liat pemandangan dari atas tapi saya ada phobia ketinggian jadi pas mau liat ke bawah tu lemes lutut
5	pernah kesini pada 7 maret 2019...bisa berjalan kaki menyeberangi jembatan repo repo ke pulau kumala...pulau kumala tempatnya bagus... wisata yang nyaman buat perorangan maupun keluarga. fasilitas lengkap... ada banyak obyek wisata di dalam pulau kumala ini...tempatnya teduh dengan rindangnya tanaman... sangat layak untuk di kunjungi...

### 3.2.2 Cleaning

Selanjutnya kita melakukan tahap *cleaning* menggunakan python untuk menghapus atribut data. Seperti penghapusan karakter yang tidak diperlukan dari teks seperti tag html dan emotication yang tercampur dengan komentar dalam teks yang mengganggu proses analisis pada data. Tabel 3.3 menunjukkan hasil sebelum dan sesudah penghapusan karakter yang tercampur pada komentar sebagai berikut:

**Tabel 3. 3** Hasil Cleaning

No	Cleaning
1	bagus jembatannya tapi tutup dimalam hari tapi bisa nongkrong dan minum kopi di pinggir sungai mahakam dan bisa mancing juga jembatan ini jalam mau kalai ke pulau bisa juga
2	mantapppp rawat dan jaga aset kembangkan untuk pariwisata dan identitas budaya jembatan adalah aset kebanggaan masyarakat kab tenggarong
3	jembatan penyebrangan ke pulau kumala pas desember kemarin ke sana sedang direnovasi dicat kembali pemandangan cukup bagus mudahan dimajukan kembali dan diberdayakan kembali pulau kumala ini dikelola lebih baik lagi biar jadi destinasi utama lagi kalau ke tenggarong selain jadi pemasukan daerah bisa jadi pusat tempat wisata budaya dan edukasi khususnya di kaltim jembatan penyebrangan ke pulau kumala

- 4 pas kesini lg ada pengecatan bagus banget liat pemandangan dari atas tapi saya ada phobia ketinggian jadi pas mau liat ke bawah tu lemes lutut
- 5 pernah kesini pada maret bisa berjalan kaki menyeberangi jembatan repo repo ke pulau kumala pulau kumala tempatnya bagus wisata yang nyaman buat perorangan maupun keluarga fasilitas lengkap ada banyak obyek wisata di dalam pulau kumala ini tempatnya teduh dengan rindangnya tanaman sangat layak untuk di kunjungi

### 3.2.3 Tokenizing

Pada tahap selanjutnya adalah *tokenizing* dimana suatu kalimat dari komentar pada Google Maps atau Dataset akan dipecah menjadi potongan kata atau token. Hal ini dilakukan untuk mengetahui kemunculan dari kata tersebut dan mempersiapkan teks agar dapat dianalisis pada tingkat kata. Berikut Tabel 3.4 menunjukkan hasil sebelum dan sesudah *tokenizing*:

**Tabel 3. 4** Hasil Tokenizing

No	Tokenizing
1	['bagus', 'jembatannya', 'tapi', 'tutup', 'dimalam', 'hari', 'tapi', 'bisa', 'nongkrong', 'dan', 'minum', 'kopi', 'di', 'pinggir', 'sungai', 'mahakam', 'dan', 'bisa', 'mancing', 'juga', 'jembatan', 'ini', 'jalam', 'mau', 'kalai', 'ke', 'pulau', 'bisa', 'juga']
2	['mantapppp', 'rawat', 'dan', 'jaga', 'aset', 'kembangkan', 'untuk', 'pariwisata', 'dan', 'identitas', 'budaya', 'jembatan', 'adalah', 'aset', 'kebanggaan', 'masyarakat', 'kab', 'tenggarong']
3	['jembatan', 'penyebrangan', 'ke', 'pulau', 'kumala', 'pas', 'desember', 'kemarin', 'ke', 'sana', 'sedang', 'direnovasi', 'dicat', 'kembali', 'pemandangan', 'cukup', 'bagus', 'mudah', 'dimajukan', 'kembali', 'dan', 'diberdayakan', 'kembali', 'pulau', 'kumala', 'ini', 'dikelola', 'lebih', 'baik', 'lagi', 'biar', 'jadi', 'destinasi', 'utama', 'lagi', 'kalau', 'ke', 'tenggarong', 'selain', 'jadi', 'pemasukan', 'daerah', 'bisa', 'jadi', 'pusat', 'tempat', 'wisata', 'budaya', 'dan', 'edukasi', 'khususnya', 'di', 'kaltim', 'jembatan', 'penyebrangan', 'ke', 'pulau', 'kumala']
4	['pas', 'kesini', 'lg', 'ada', 'pengecatan', 'bagus', 'banget', 'liat', 'pemandangan', 'dari', 'atas', 'tapi', 'saya', 'ada', 'phobia', 'ketinggian', 'jadi', 'pas', 'mau', 'liat', 'ke', 'bawah', 'tu', 'lemes', 'lutut']
5	['pernah', 'kesini', 'pada', 'maret', 'bisa', 'berjalan', 'kaki', 'menyeberangi', 'jembatan', 'repo,repo', 'ke', 'pulau', 'kumala', 'pulau', 'kumala', 'tempatnya', 'bagus', 'wisata', 'yang', 'nyaman', 'buat', 'perorangan', 'maupun', 'keluarga', 'fasilitas', 'lengkap', 'ada', 'banyak', 'obyek', 'wisata', 'di', 'dalam', 'pulau', 'kumala', 'ini', 'tempatnya', 'teduh', 'dengan', 'rindangnya', 'tanaman', 'sangat', 'layak', 'untuk', 'di', 'kunjungi']

### 3.2.4 Stopword Removal

Selanjutnya adalah *Stopword removal* yang bertujuan penghapusan kata-kata dari beberapa potongan teks yang tidak dianggap penting untuk makna teks. Jika terdapat kata yang sesuai dengan kata yang terdapat pada stopwords maka kata tersebut akan dihapus dari dokumen karena dianggap sebagai kata yang tidak penting. Contoh hasil *stopword removal* pada Tabel 3.5

**Tabel 3. 5** Hasil Stopword Removal

No.	Stopword Removal
1	['bagus', 'jembatannya', 'tutup', 'dimalam', 'nongkrong', 'minum', 'kopi', 'pinggir', 'sungai', 'mahakam', 'mancing', 'jembatan', 'jalam', 'kalai', 'pulau']
2	['mantapppp', 'rawat', 'jaga', 'aset', 'kembangkan', 'pariwisata', 'identitas', 'budaya', 'jembatan', 'aset', 'kebanggaan', 'masyarakat', 'kab', 'tenggarong']
3	['jembatan', 'penyebrangan', 'pulau', 'kumala', 'pas', 'desember', 'kemarin', 'direnovasi', 'dicat', 'pemandangan', 'bagus', 'mudah', 'dimajukan', 'diberdayakan', 'pulau', 'kumala', 'dikelola', 'biar', 'destinasi', 'utama', 'tenggarong', 'pemasukan', 'daerah', 'pusat', 'wisata', 'budaya', 'edukasi', 'kaltim', 'jembatan', 'penyebrangan', 'pulau', 'kumala']
4	['pas', 'kesini', 'lg', 'pengecatan', 'bagus', 'banget', 'liat', 'pemandangan', 'phobia', 'ketinggian', 'pas', 'liat', 'tu', 'lemes', 'lutut']
5	['kesini', 'maret', 'berjalan', 'kaki', 'menyeberangi', 'jembatan', 'repo', 'repo', 'pulau', 'kumala', 'pulau', 'kumala', 'tempatny', 'bagus', 'wisata', 'nyaman', 'perorangan', 'keluarga', 'fasilitas', 'lengkap', 'obyek', 'wisata', 'pulau', 'kumala', 'tempatny', 'teduh', 'rindangnya', 'tanaman', 'layak', 'kunjungi']

### 3.2.5 Stemming

Selanjutnya proses terakhir dalam *preprocessing* adalah *stemming*. Ini berguna untuk mengubah kata yang berimbuhan menjadi kata dasar agar memudahkan dalam proses pembobotan nantinya. Tahap *stemming* ini menggunakan *library Sastrawi*, dalam kata lain kata pada dataset dari tahap sebelumnya ke kata dasar sesuai Kamus Besar Bahasa Indonesia. Contoh hasil stemming pada Tabel 3.6

**Tabel 3. 6 Hasil Stemming**

No.	Stemming
1	bagus jembatan tutup malam nongkrong minum kopi pinggir sungai mahakam mancing jembatan jalam kalai pulau
2	mantappppp rawat jaga aset kembang pariwisata identitas budaya jembatan aset bangga masyarakat kab tenggarong
3	jembatan penyebrangan pulau mala pas desember kemarin renovasi cat pandang bagus mudah maju daya pulau mala kelola biar destinasi utama tenggarong pasu daerah pusat wisata budaya edukasi kaltim jembatan penyebrangan pulau mala
4	pas kesini lg cat bagus banget liat pandang phobia tinggi pas liat tu lemes lutut
5	kesini maret jalan kaki rang jembatan repo repo pulau mala pulau mala tempat bagus wisata nyaman orang keluarga fasilitas lengkap obyek wisata pulau mala tempat teduh rindang tanam layak kunjung

### 3.3 Visualisasi

Proses visualisasi ini dilakukan untuk memahami hasil kata-kata atau term yang menghasilkan frekuensi tertinggi dalam data dari penelitian yang dilakukan tadi. Pada gambar 3.1 bisa terlihat hasil yang didapat ada 3 sentimen. Dari ketiga itu sentimen positif, sentimen netral dan sentimen negatif. Semakin besar ukuran kata yang ada di dalam word cloud maka semakin tinggi juga frekuensi kata tersebut.







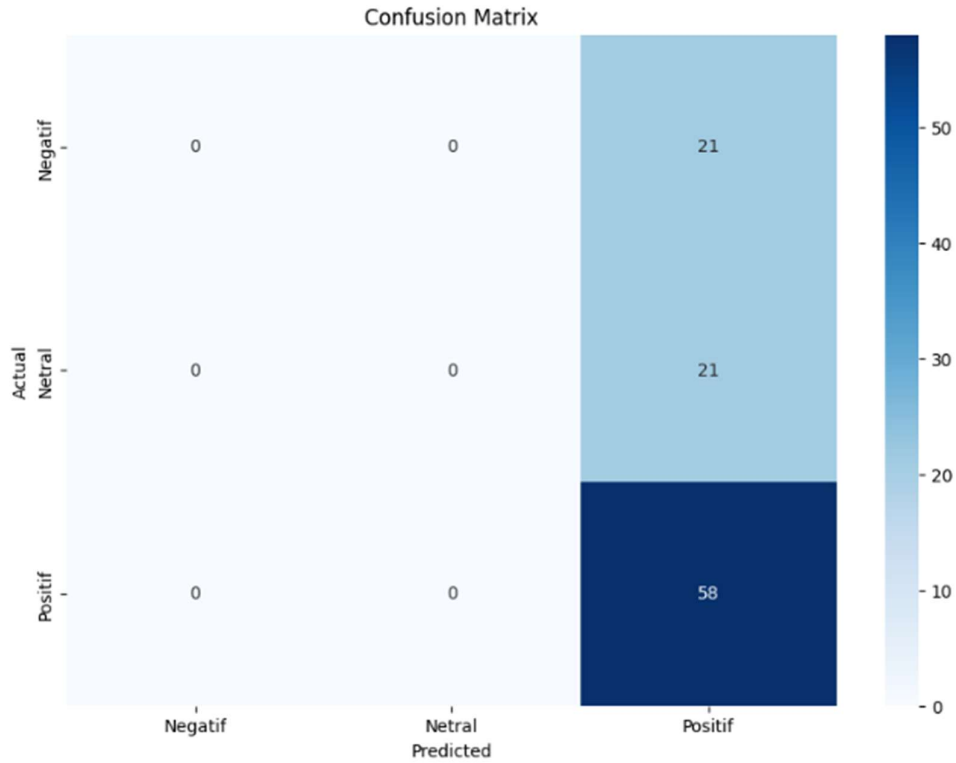
	stemming	total_tfidf
0	bagus jembatan tutup malam nongkrong minum kop...	3.583506
1	mantapppp rawat jaga aset kembang pariwisata i...	3.375408
2	jembatan penyebrangan pulau mala pas desember ...	4.880932
3	pas kesini lg cat bagus banget liat pandang ph...	3.361621
4	kesini maret jalan kaki rang jembatan repo rep...	4.573934
..	...	...
495	salah wisata hibur libur	1.960785
496	bagus mudah akses	1.693713
497	rekreasi keluarga letak kota tenggarong	2.184971
498	pagi	1.000000
499	bawa payung sewa motor jam	2.225833

**Gambar 3. 2** Pembobotan Kata (TF-IDF)

Pada Gambar 3.2 menunjukkan bahwa baris pertama dengan teks “bagus jembatan tutup malam nongkrong minum kopi...” memiliki nilai TF-IDF total sebesar 3.583506. Ini menunjukkan bahwa kata-kata dalam kalimat tersebut memiliki kepentingan dalam teks ulasan yang dianalisis dan pada teks “jembatan penyebrangan pulau mala pas desember...” memiliki nilai TF-IDF total sebesar 4.880932, yang merupakan salah satu nilai tertinggi di antara contoh yang ditampilkan.

### 3.5 Evaluasi

Dalam penelitian pada data yang digunakan sebanyak 500 data yang digunakan dan dibagi menjadi data latih dan data uji dengan rasio 80:20. Sebanyak 400 data yang digunakan untuk data latih dan 100 data sebagai data uji. Akurasi pada model yang dihitung berdasarkan prediksi yang dihasilkan pada data uji, dan ditampilkan pada Gambar 3.3 dibawah.



Accuracy: 0.58

Classification Report:

	precision	recall	f1-score	support
Negatif	1.00	0.00	0.00	21
Netral	1.00	0.00	0.00	21
Positif	0.58	1.00	0.73	58
accuracy			0.58	100
macro avg	0.86	0.33	0.24	100
weighted avg	0.76	0.58	0.43	100

**Gambar 3.3** Hasil Akurasi

Perhitungan nilai akurasi sebagai berikut:

$$Accuracy = \frac{58 + 0}{58 + 0 + 21 + 21}$$

$$= \frac{58}{100} = 0,58\%$$

Pada Gambar 3.3 penerapan algoritma Naïve Bayes dalam penelitian ini menggunakan rasio sebesar 80:20 menghasilkan akurasi sebesar 0,58 %.

## **BAB IV**

### **PENUTUP**

#### **4.1 Kesimpulan**

Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat disimpulkan, bahwa:

- (i) Metode Naïve Bayes terbukti mampu memberikan hasil yang akurat dalam mengklasifikasi sentimen ulasan pengguna Google Maps terhadap Jembatan Repo-Repo. Hasil analisis menunjukkan akurasi sebesar 58%, karena kualitas data yang digunakan memiliki noise atau tidak relevan dan juga tidak seimbang pada data membuat metode Naïve Bayes kesulitan memberikan akurasi yang akurat.
- (ii) Hasil klasifikasi sentimen menunjukkan bahwa sebagian besar ulasan pengguna cenderung positif. Namun, terdapat juga ulasan negatif yang menunjukkan adanya aspek yang perlu diperbaiki terkait infrastruktur dan pelayanan di sekitar Jembatan Repo-Repo.

Secara keseluruhan, penelitian ini berhasil menunjukkan bahwa analisis sentimen dengan metode Naïve Bayes, fitur ekstraksi TF-IDF dapat memberikan hasil yang signifikan dan bermanfaat bagi pengambilan keputusan dalam pengembangan destinasi wisata.

## 4.2 Saran

Berdasarkan penelitian yang sudah dilakukan terdapat beberapa saran, yaitu:

- (i) Disarankan untuk mengumpulkan lebih banyak data ulasan dan melakukan penelitian lanjutan dengan menggunakan metode klasifikasi lain seperti Random Forest, K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM) untuk membandingkan kinerja dan mencari metode yang paling optimal.
- (ii) Pengembangan Infrastruktur yang ada di Jembatan maupun di Pulau Kumala karena untuk menindaklanjuti ulasan negatif dengan memperbaiki infrastruktur dan layanannya untuk meningkatkan kualitas pada Jembatan Repo-Repo.

## Daftar Pustaka

- Anggraini, M. A., & Susilo, B. (2022). *Hidupkan Pulau Kumala Tenggara Kukar, Jembatan Repo-repo akan Dihiasi Lampu Tematik*. *Tribun Kaltim*.  
<https://kaltim.tribunnews.com/2022/10/26/hidupkan-pulau-kumala-tenggara-kukar-jembatan-repo-repo-akan-dihiasi-lampu-tematik>
- Ariyani, V., Putri, P., Prasetijo, A. B., & Eridani, D. (2022). Perbandingan Kinerja Algoritme Naïve Bayes Dan K-Nearest Neighbor (Knn) Untuk Prediksi Harga Rumah. *Jurnal Ilmiah Teknik Elektro*, 4. <https://ejournal.undip.ac.id/index.php/transmisi>
- Astari, N. M. A. J., Dewa Gede Hendra Divayana, & Gede Indrawan. (2020). Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier. *Jurnal Sistem Dan Informatika (JSI)*, 15(1), 27–29.  
<https://doi.org/10.30864/jsi.v15i1.332>
- Doloksaribu, H. P., & Samuel, Y. T. (2022). Komparasi Algoritma Data Mining Untuk. *Jurnal Teknologi Informasi*, 16(1), 1–11.
- Fuad Amirullah, Syariful Alam, & M.Imam Sulisty S. (2023). Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini Twitter Menggunakan Naïve Bayes. *STORAGE: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 2(3), 69–76.  
<https://doi.org/10.55123/storage.v2i3.2293>
- Isnain, A. R., Sakti, A. I., Alita, D., & Marga, N. S. (2021). Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jurnal Data Mining Dan Sistem Informasi*, 2(1), 31. <https://doi.org/10.33365/jdmsi.v2i1.1021>
- Khairunnisa, S., Adiwijaya, A., & Faraby, S. Al. (2021). Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19). *Jurnal Media Informatika Budidarma*, 5(2), 406.  
<https://doi.org/10.30865/mib.v5i2.2835>
- Kiedrowsky, F. F., & Andrianingsih. (2023). Sentiment Analysis Marketplaces Digital menggunakan Machine Learning. *Jurnal JTik (Jurnal Teknologi Informasi Dan Komunikasi)*, 7(3), 493–499. <https://doi.org/10.35870/jtik.v7i3.1002>
- Locarso, G. K. (2022). Analisis Sentimen Review Aplikasi Pedulilindungi Pada Google Play Store Menggunakan Nbc. *Jurnal Teknik Informatika Kaputama (JTik)*, 6(2), 353–361.

<https://jurnal-backup.kaputama.ac.id/index.php/JTIK/article/view/800/598>

- Maris, S. (2023). *Kutai Kartanegara Bakal Kembangkan 8 Objek Wisata Penghasil PAD*. Liputan 6. <https://www.liputan6.com/news/read/5445221/kutai-kartanegara-bakal-kembangkan-8-objek-wisata-penghasil-pad>
- Muhammad Bambang Firdaus, Tejawati, A., Hafizh, M. S., & Anam, M. K. (2021). Panduan Navigasi Virtual Tour Pulau Kumala Kutai Kartanegara. *Metik Jurnal*, 5(1), 1–7. <https://doi.org/10.47002/metik.v5i1.201>
- Muslimin, M., & Lusiana, V. (2023). Analisis Sentimen Terhadap Kenaikan Harga Bahan Pokok Menggunakan Metode Naive Bayes Classifier. *Jurnal Media Informatika Budidarma*, 7(3), 1200–1209. <https://doi.org/10.30865/mib.v7i3.6418>
- Muzaki, A., & Witanti, A. (2021). Sentiment Analysis of the Community in the Twitter To the 2020 Election in Pandemic Covid-19 By Method Naive Bayes Classifier. *Jurnal Teknik Informatika (Jutif)*, 2(2), 101–107. <https://doi.org/10.20884/1.jutif.2021.2.2.51>
- Novemyleo. (2021). *Daftar Kecamatan di Kabupaten Kutai Kartanegara, Provinsi Kalimantan Timur, Indonesia*. Pos Kupang. <https://poskupangwiki.tribunnews.com/2021/12/07/daftar-kecamatan-di-kabupaten-kutai-kartanegara-provinsi-kalimantan-timur-indonesia?page=all>
- Nurian, A. (2023). Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naïve Bayes. *Jurnal Informatika Dan Teknik Elektro Terapan*, 11(3s1), 1613–1621. <https://doi.org/10.23960/jitet.v11i3s1.3348>
- Prakoso, C., & Hermawan, A. (2023). Perbandingan Model Machine Learning dalam Analisis Sentimen Ulasan Pengunjung Keraton Yogyakarta pada Google Maps. *Kajian Ilmiah Informatika Dan Komputer*, 4(3), 1292–1302. <https://doi.org/10.30865/klik.v4i3.1419>
- Puja Kekal, H., Gata, W., Nurdiani, S., Setio Rini, A. J., & Sely Wita, D. (2021). Analisa Pencarian Rute Tercepat Menuju Tempat Wisata Pulau Kumala Kota Tenggarong Menggunakan Algoritma Greedy. *Jurnal Ilmiah Ilmu Komputer*, 7(1), 9–15. <https://doi.org/10.35329/jiik.v7i1.179>
- Purwanto, A. (2022). *Kabupaten Kutai Kartanegara: Dari Kerajaan Tertua Jadi Calon Ibu Kota Negara*. Kompas Pedia.



<https://kompaspedia.kompas.id/baca/profil/daerah/kabupaten-kutai-kartanegara-dari-kerajaan-tertua-jadi-calon-ibu-kota-negara>

- Puspitasari, R., Findawati, Y., Rosid, M. A., Informatika, P. S., & Sidoarjo, U. M. (2023). *Sentiment Analysis of Post-Covid-19 Inflation Based on Twitter Using the K-Nearest Neighbor and Support Vector Machine Analisis Sentimen Terhadap Inflasi Pasca Covid-19 Berdasarkan Twitter Dengan Metode Klasifikasi K-Nearest Neighbor Dan*. 4(4), 1–11.
- Putu, N. L. P. M., Ahmad Zuli Amrullah, & Ismarmiaty. (2021). Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 123–131. <https://doi.org/10.29207/resti.v5i1.2587>
- Rizki Maulidiah, P., Anjani Arifiyanti, A., Pembangunan Nasional, U., Timur Jl Raya Rungkut Madya, J., Anyar, G., & Timur, J. (2023). Analisis Sentimen Ulasan Pengunjung Terhadap Tempat Wisata Religi Walisongo Menggunakan Metode Supervised Learning Dhian Satria Yudha Kartika. *Jurnal Ilmiah Teknik Informatika Dan Komunikasi*, 3(3). <https://doi.org/10.55606/juitik.v3i3.617>
- Septian, J. A., Fachrudin, T. M., & Nugroho, A. (2019). Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor. *Journal of Intelligent System and Computation*, 1(1), 43–49. <https://doi.org/10.52985/insyst.v1i1.36>
- Ulfah, A., & Najjah, I. (2023). Implementasi Web Scraping Pada Situs Jurnal Sinta Menggunakan Framework Selenium Webdriver Python. *JIKA (Jurnal Informatika)*, 7(1), 29. <https://doi.org/10.31000/jika.v7i1.7037>
- Wahyu, A. (2021). *7 Destinasi Wisata Tenggara untuk Liburan bersama Keluarga!* Orami Author. <https://www.orami.co.id/magazine/wisata-tenggarong>

## Daftar Lampiran

### L1 Pengumpulan Data

```
▶ | feature_id = "0x2df67ace57e958ef:0xa9fb2adcb21c8264"  
   | loop = 60  
  
[ ] pip install beautifulsoup4
```

```
from datetime import datetime, timedelta  
  
# nama hari dalam bahasa Indonesia  
nama_hari_indonesia = {  
    "Monday": "Senin",  
    "Tuesday": "Selasa",  
    "Wednesday": "Rabu",  
    "Thursday": "Kamis",  
    "Friday": "Jumat",  
    "Saturday": "Sabtu",  
    "Sunday": "Minggu"  
}  
  
# Fungsi untuk mengonversi teks waktu menjadi timedelta  
def konversi_teks_ke_timedelta(teks):  
    if 'hari' in teks:  
        jumlah = int(teks.split()[0])  
        return timedelta(days=jumlah)  
    elif 'minggu' in teks:  
        if 'seminggu' in teks:  
            return timedelta(weeks=1)  
        else:  
            jumlah = int(teks.split()[0])  
            return timedelta(weeks=jumlah)  
    elif 'bulan' in teks:  
        if 'sebulan' in teks:  
            return timedelta(days=30)  
        else:  
            jumlah = int(teks.split()[0])  
            return timedelta(days=jumlah*30)  
    else:  
        raise ValueError("Format teks waktu tidak dikenali")  
  
# Fungsi untuk mengonversi teks waktu menjadi tanggal dan hari  
def konversi_teks_ke_tanggal(teks):  
    if 'hari' in teks or 'minggu' in teks or 'bulan' in teks:  
        timedelta_lalu = konversi_teks_ke_timedelta(teks)  
        tanggal = datetime.now() - timedelta_lalu
```

```
        nama_hari_indo = nama_hari_indonesia.get(nama_hari, nama_hari)  
        return tanggal, nama_hari_indo  
    elif 'tahun' in teks:  
        if 'setahun' in teks:  
            return datetime.now() - timedelta(days=365), None  
        else:  
            jumlah_tahun = int(teks.split()[0])  
            return datetime.now() - timedelta(days=365 * jumlah_tahun), None  
    else:  
        # Jika format tidak sesuai dengan yang dikenali, kembalikan None  
        return None, None
```

```

import requests
import re
from datetime import datetime, timedelta
from bs4 import BeautifulSoup
import pandas as pd

def get_reviews_data(feature_id):
    token = ''
    headers = {
        "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/101.0.4951.54 Safari/537.36"
    }

    base_url = "https://www.google.com/async/reviewDialog?hl=id&async=feature_id:{feature_id},next_page_token:{token},sort_by:newest,start_index:,associated_topic:,_fmt:pc"

    # List untuk menyimpan DataFrame dari setiap permintaan
    dfs = []

    total_data = 0

    # Lakukan loop permintaan
    for i in range(loop):
        response = requests.get(base_url.format(feature_id=feature_id, token=token), headers=headers)
        soup = BeautifulSoup(response.content, 'html.parser')

        user = []
        location_info = {}
        data_id = ''

        for el in soup.select('.c9QyIf'):
            data_id = soup.select_one('.DvzRnc')['data-fid']
            next_page_token = soup.select_one('.gws-localreviews__general-reviews-block')['data-next-page-token']
            if next_page_token:
                token = next_page_token
            else:
                token = None # Jika token berikutnya kosong, hentikan loop
                break

        for el in soup.select('.gws-localreviews__google-review'):

```

```

# try:
rating_label = el.select_one('span.lTi8oc.z3HNkc')['aria-label']
match = re.search(r'(\d+)(?:,\d+)?', rating_label)

if match:
    rating = match.group(1)
else:
    rating = "0"
# except KeyError:
#     rating = "0"

waktu = el.select_one('.dehysf').text.strip()
# print(waktu)
tanggal, nama_hari_indo = konversi_teks_ke_tanggal(waktu)
# print(tanggal)
# if tanggal is None: # Tambahkan penanganan eksepsi di sini
#     continue # Lewati data yang tidak valid

tanggal_formatted = tanggal.strftime("%d-%m-%Y")
time_formatted = f"{nama_hari_indo} {tanggal_formatted}"
# Determine sentiment
if int(rating) <= 3:
    sentimen = 'Negatif'
elif int(rating) == 4:
    sentimen = 'Netral'
else:
    sentimen = 'Positif'
user.append({
    'name': el.select_one('.TSubDb').text.strip(),
    'rating': rating,
    'review': el.select_one('.Jtu6Td').text.strip(),
    'time': time_formatted,
    'sentimen': sentimen
})

df = pd.DataFrame(user)

```

```

df = pd.DataFrame(user)

# Tambahkan DataFrame ke dalam list
dfs.append(df)
total_data += len(df)

print(f"Loop ke-{i+1}:")
# print(df)
print("-----")

if token is None:
    break # Hentikan loop jika token berikutnya kosong

# Gabungkan semua DataFrame menjadi satu
combined_df = pd.concat(dfs, ignore_index=True)

# Simpan DataFrame gabungan ke dalam file Excel
combined_df.to_excel("googlemaps_data.xlsx", index=False)

print("Total data yang sudah didapat:", total_data)
print("Data telah disimpan dalam googlemaps_data.xlsx")

get_reviews_data(feature_id)

```

## L2 Install Library

```

import re # Modul untuk regular expression
import pandas as pd # Manipulasi data tabular dengan Pandas
import numpy as np # Operasi numerik dengan NumPy
import nltk # Pemrosesan teks dengan NLTK

!pip install Sastrawi # Paket Sastrawi untuk pemrosesan bahasa Indonesia

from Sastrawi.Stemmer.StemmerFactory import StemmerFactory # Stemming bahasa Indonesia dari Sastrawi
from nltk.corpus import stopwords # Stopwords dari NLTK
from nltk.tokenize import word_tokenize # Tokenisasi teks dari NLTK

nltk.download('punkt') # Model tokenisasi teks dari NLTK
nltk.download('stopwords') # Daftar stopwords dari NLTK

```

### L3 Cek Informasi Dataset

```
import pandas as pd

# Mencoba membaca file CSV dengan encoding 'latin1'
data = pd.read_csv("data.csv", encoding='latin1')
data.info() # Menampilkan informasi tentang struktur data

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 502 entries, 0 to 501
Data columns (total 5 columns):
#   Column      Non-Null Count  Dtype
---  -
0   name        502 non-null   object
1   rating      502 non-null   int64
2   review      502 non-null   object
3   time        502 non-null   object
4   sentimen    502 non-null   object
dtypes: int64(1), object(4)
memory usage: 19.7+ KB
```

### L4 Cek Jumlah Sentimen Positif, Negatif, dan Netral

```
sentiment_counts = data['sentimen'].value_counts() # Menghitung Jumlah Kemunculan sentimen positif dan negatif

positif = sentiment_counts.get('Positif', 0)
netral = sentiment_counts.get('Netral', 0)
negatif = sentiment_counts.get('Negatif', 0)

print(f"Total Sentimen Positif: {positif}") # mencetak jumlah total sentimen positif.
print(f"Total Sentimen Netral: {netral}") # mencetak jumlah total sentimen netral.
print(f"Total Sentimen Negatif: {negatif}") # mencetak jumlah total sentimen negatif.
```

### L5 Membaca Dataset Kolom Review dan Sentimen

```
columns = ['review', 'sentimen']
df = pd.DataFrame(data, columns=columns).iloc[0:500] # membatasi DataFrame hanya pada 500 baris pertama.
df.head(5) # Menampilkan 5 baris pertama dari DataFrame df untuk melihat preview data.
```

### L6 Pre-Processing Data

```
def case_folding(text):
    # Mengubah teks menjadi huruf kecil
    if isinstance(text, str):
        lowercase_text = text.lower()
        return lowercase_text
    else:
        return text

df['case_folding'] = df['review'].apply(case_folding)
df.head(5)
```

```

def clean_text(review):
    # Remove URLs
    review = re.sub(r'https?:\/\/\S+', '', review)
    # Remove non-alphabetic characters
    review = re.sub(r'^a-zA-Z', ' ', review)

    # Remove emojis
    emoji_pattern = re.compile("[
        u"\U0001F600-\U0001F64F" # emoticons
        u"\U0001F300-\U0001F5FF" # symbols & pictographs
        u"\U0001F680-\U0001F6FF" # transport & map symbols
        u"\U0001F1E0-\U0001F1FF" # flags (iOS)
        ]+", flags=re.UNICODE)
    review = emoji_pattern.sub(r'', review)

    # Remove digits
    review = re.sub(r'\d+', '', review)
    # Remove single characters
    review = re.sub(r'\b\w\b', '', review)
    # Remove extra spaces
    review = re.sub(r'\s+', ' ', review).strip()

    return review

# Applying the function to the 'case_folding' column
df['cleansing'] = df['case_folding'].apply(clean_text)
df.head(5)

```

```

def tokenize(text):
    return word_tokenize(text)

df['tokenizing'] = df['cleansing'].apply(tokenize)
df.head(5)

```

```

stop_words = stopwords.words('indonesian')

additional_stopwords = [
    'yg', 'b', 'd', 'nya', 'ya', 'sih', 'aku', 'kamu', 'dia', 'mereka',
    'kita', 'saya', 'kalian', 'ini', 'itu', 'loh', 'dong', 'kan', 'uh', 'aa',
    'uh', 'kak', 'ah', 'oh', 'hmm', 'eh', 'aduh', 'ad', 'deh', 'hehe', 'ehh',
    'gk', 'h', 'hh', 'huh', 'udah', 'mau', 'gw', 'lu', 'lo', 'loh', 'pake', 'ntah',
    'ma', 'km', 'kalo', 'kl', 'tuh', 'set', 'dll', 'mcm', 'mn', 'gan', 'bro', 'sis', 'de',
    'btw', 'kayak', 'spt', 'bilang', 'kata', 'tau', 'lihat', 'bikin', 'dan', 'atau',
    'setelah', 'juga', 'ts', 'btul', 'tg', 'tr', 'utk', 'kmrn', 'sy', 'sm', 'dm',
    'hrs', 'dg', 'dl', 'dt', 'sdh', 'tp', 'bsa', 'dan', 'yang'
]

stop_words.extend(additional_stopwords)

def remove_stopwords(text):
    return [word for word in text if word not in stop_words]

df['stopword_removal'] = df['tokenizing'].apply(lambda x: remove_stopwords(x))
df.head(5)

```

```

factory = StemmerFactory()
stemmer = factory.create_stemmer()

def stem_text(tokens):
    return [stemmer.stem(word) for word in tokens]

df['stemming'] = df['stopword_removal'].apply(stem_text)
df['stemming'] = df['stemming'].apply(lambda x: ' '.join(x))
df.head(5)

```

## L7 Visualisasi Wordcloud

```

import pandas as pd
import numpy as np
from PIL import Image
from wordcloud import WordCloud, STOPWORDS, ImageColorGenerator
import matplotlib.pyplot as plt

# Read CSV data
df = pd.read_csv('data.csv')

# Rename the column 'label' to 'sentimen'
df.rename(columns={'label': 'sentimen'}, inplace=True)

# Handle missing values
df['stemming'] = df['stemming'].fillna('').astype(str)

# Combine text based on sentimen
positive_text = ' '.join(df[df['sentimen'] == 'Positif']['stemming'].tolist())
neutral_text = ' '.join(df[df['sentimen'] == 'Netral']['stemming'].tolist())
negative_text = ' '.join(df[df['sentimen'] == 'Negatif']['stemming'].tolist())

# Define stopwords
stopwords = set(STOPWORDS)
stopwords.update(['https', 'co', 'RT', '...', 'amp'])

# Create WordCloud function
def create_wordcloud(text, title):
    wc = WordCloud(stopwords=stopwords, background_color="white", max_words=500, width=800, height=400)
    wc.generate(text)

    plt.figure(figsize=(10, 5))
    plt.imshow(wc, interpolation='bilinear')
    plt.axis("off")
    plt.title(title)
    plt.show()

# Generate and display WordClouds
create_wordcloud(positive_text, 'WordCloud - Positif')
create_wordcloud(neutral_text, 'WordCloud - Netral')
create_wordcloud(negative_text, 'WordCloud - Negatif')

```

## L8 Pembobotan Kata (TF-IDF)

```
import pandas as pd
from sklearn.feature_extraction.text import TfidfVectorizer

# Assuming 'documents' is already defined and contains the list of stemmed texts

# Create and fit TF-IDF vectorizer
vectorizer = TfidfVectorizer()
features = vectorizer.fit_transform(documents)

# Convert the TF-IDF matrix to a DataFrame
tfidf_df = pd.DataFrame(features.toarray(), columns=vectorizer.get_feature_names_out())

# Calculate the total TF-IDF score for each document
tfidf_totals = tfidf_df.sum(axis=1)

# Add the total TF-IDF scores to the original DataFrame
df_preprocessed['total_tfidf'] = tfidf_totals

# Display the DataFrame with total TF-IDF scores
print(df_preprocessed[['stemming', 'total_tfidf']])
```

## L9 Split Data dan Evaluasi Confusion Matrix Metode Naïve Bayes

```
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

# Membagi data menjadi set pelatihan dan pengujian (80% pelatihan, 20% pengujian)
X_train, X_test, y_train, y_test = train_test_split(final_df.drop('sentimen', axis=1),
                                                    final_df['sentimen'],
                                                    test_size=0.2,
                                                    random_state=42)

# Menampilkan jumlah data yang dilatih dan diuji
print(f"Jumlah data pelatihan: {len(X_train)}")
print(f"Jumlah data pengujian: {len(X_test)}")

# Visualisasi distribusi kelas pada data pelatihan
plt.figure(figsize=(14, 6))
plt.subplot(1, 2, 1)
sns.countplot(x=y_train)
plt.title('Distribusi Kelas Data Pelatihan')
plt.xlabel('Sentimen')
plt.ylabel('Frekuensi')

# Visualisasi distribusi kelas pada data pengujian
plt.subplot(1, 2, 2)
sns.countplot(x=y_test)
plt.title('Distribusi Kelas Data Pengujian')
plt.xlabel('Sentimen')
plt.ylabel('Frekuensi')

plt.show()

# Menginisialisasi classifier Naive Bayes
nb_classifier = MultinomialNB()

# Melatih classifier dengan data pelatihan
nb_classifier.fit(X_train, y_train)
```



```
# Prediksi pada set pengujian
y_pred = nb_classifier.predict(X_test)

# Akurasi
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)

# Laporan Klasifikasi
print("\nClassification Report:")
print(classification_report(y_test, y_pred, zero_division=1))

# Confusion Matrix
conf_matrix = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(10, 7))
sns.heatmap(conf_matrix, annot=True, fmt="d", cmap="Blues", xticklabels=['Negatif', 'Netral', 'Positif'], yticklabels=['Negatif', 'Netral', 'Positif'])
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.title('Confusion Matrix')
plt.show()
```

## L10 Surat Ijin Penelitian

 <p><b>UMKT</b> Program Studi Teknik Informatika Fakultas Sains dan Teknologi</p>	Telp. 0541-748511 Fax.0541-766832 Website <a href="http://informatika.umkt.ac.id">http://informatika.umkt.ac.id</a> email: <a href="mailto:informatika@umkt.ac.id">informatika@umkt.ac.id</a>
--	---

*بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ*

Nomor : 055-011/KET/FST.1/A/2024  
Lampiran : -  
Perihal : **Keterangan Pengambilan Data Sekunder**

*Assalamu'alaikum Warrahmatullahi Wabarrakatuh*

Puji Syukur kepada Allah Subhanahu wa ta'ala yang senantiasa melimpahkan Rahmat-Nya kepada kita sekalian. Amin.

Dengan surat ini, kami menerangkan bahwa mahasiswa berikut:

Nama : Muhammad Reifin Perdana  
NIM : 2011102441136  
Program Studi : Teknik Informatika

Melakukan penelitian dengan pengambilan data sekunder di Google Maps data yang diambil yaitu ulasan terkait jembatan repo-repo.

Demikian hal ini disampaikan, atas kerjasamanya kami ucapkan terima kasih.

*Wassalamu'alaikum Warrahmatullahi Wabarrakatuh*

Samarinda, 21 Dzulhijjah 1445 H  
28 Juni 2024 M

Ketua Program Studi S1 Teknik Informatika  
**Arbansyah, S.Kom., M.TI**  
NIDN. 1118019203














Kampus 1 : Jl. Ir. H. Juanda, No.15, Samarinda  
Kampus 2 : Jl. Pelita, Pesona Mahakam, Samarinda

## L11 Kartu Bimbingan

### KARTU KENDALI BIMBINGAN

Nama Mahasiswa : Muhammad Reifin Perdana  
NIM : 2011102441136  
Nama Dosen Pembimbing : Rudiman, S.Kom, M.Sc  
Judul Penelitian : ANALISIS SENTIMEN ULASAN JEMBATAN  
REPO-REPO di GOOGLE MAPS  
MENGUNAKAN METODE NAÏVE BAYES  
DENGAN FITUR EKSTRAKSI TF-IDF

NO	TANGGAL	URAIAN PEMBIMBINGAN	PARAF DOSEN
1	7 Feb 2024	Bimbingan pertama mendiskusikan persetujuan bimbingan dengan dosen Bapak Rudiman S.Kom, M.Sc. dengan topik penelitian Data Sentiment Analyst	
2	14 Feb 2024	Mencari permasalahan topik yang akan diteliti untuk dijadikan judul	
3	22 Feb 2024	Melakukan tahap pencarian data menggunakan python di Google Colab	
4	29 Feb 2024	Menentukan judul dan mulai perencanaan latar belakang dengan syarat ketentuan yang telah ditetapkan oleh dosen pembimbing, yaitu, menggunakan sumber jurnal paling sedikit 13 dan maksimal 15 jurnal untuk BAB 1 dan bersinta minimal sinta 4	
5	9 Maret 2024	Dalam pengerjaan latar belakang harus mengikuti aturan yang telah ditetapkan oleh dospem	
6	13 Maret 2024	Melakukan beberapa revisi terhadap mahasiswa yang dibimbing	
7	18 Maret 2024	Memberikan beberapa saran terhadap canvas pengajuan judul ke prodi	
8	27 Maret 2024	Memberikan beberapa revisi terhadap latar belakang rumusan masalah batasan masalah tujuan penelitian dan manfaat penelitian	

9	5 April 2024	Masuk BAB II menggunakan strategi yang telah diberikan oleh dosen pembimbing menggunakan pembobotan TF-IDF dalam penelitian yang dilakukan	
10	9 April 2024	Memberi masukan dan revisi dan bab II	
11	24 April 2024	Membahas target penelitian dan mencari sumber jurnal serta memberikan revisian pada BAB I dan BAB II	

Dosen Pembimbing

  
Rudiman, S.Kom, M.Sc  
 NIDN. 1105068202

Mengetahui,  
 Ketua Program Studi

  
Arbansyah, S.Kom., M.TI  
 NIDN. 1118019203

## RIWAYAT HIDUP



Muhammad Reifin Perdana, lahir di Tenggarong pada tanggal 17 Juli 2002. Merupakan anak yang lahir dari pasangan Bapak Aripin dan Ibu Yusparini. Pendidikan yang telah di tempuh oleh peneliti yaitu SD Negeri 003 Tenggarong lulus pada tahun 2014. Kemudian dilanjutkan dengan menempuh Pendidikan di SMP Negeri 1 Tenggarong lulus pada tahun 2017. Setelah itu dilanjutkan dengan menempuh Pendidikan di SMA Negeri 1 Tenggarong, lulus pada tahun 2020. Pada tahun 2020 peneliti menempuh Pendidikan Program Studi S1 Teknik Informatika di Universitas Muhammadiyah Kalimantan Timur sebagai syarat kelulusan dan memperoleh gelar Sarjana Komputer (S.Kom).