

**PERBANDINGAN AKURASI ALGORITMA NAÏVE BAYAES DAN REGRESI
LOGISTIK DALAM ANALISIS SENTIMEN WISATAWAN TERHADAP
BUGIS WATERPARK ADVENTURE DI ULASAN GOOGLE MAPS**

SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Mendapatkan Gelar Sarjana (S.Kom)
Program Studi Informatika



JUMARDIN RUSLAN 105841104419

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAKASSAR**

2024

PERBANDINGAN AKURASI ALGORITMA NAÏVE BAYES DAN REGRESI
LOGISTIK DALAM ANALISIS SENTIMEN WISATAWAN TERHADAP BUGIS
WATERPARK ADVENTURE DI ULASAN GOOGLE MAPS

Diajukan Untuk Memenuhi Salah Satu Syarat Guna Memperoleh Gelar
Sarjana Komputer Prodi Informatika Fakultas Teknik Universitas
Muhammadiyah Makassar

Disusun Dan Diajukan Oleh:

JUMARDIN RUSLAN

105841104419

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAKASSAR

2024



UNIVERSITAS MUHAMMADIYAH MAKASSAR
FAKULTAS TEKNIK

GEDUNG MENARA IQRA LT. 3

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

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

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

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

PENGESAHAN

Skripsi atas nama **Jumardin Ruslan** dengan nomor induk Mahasiswa **105 84 11044 19**, dinyatakan diterima dan disahkan oleh Panitia Ujian Tugas Akhir/Skripsi sesuai dengan Surat Keputusan Dekan Fakultas Teknik Universitas Muhammadiyah Makassar Nomor : 066/05/A.5-III/45/2024, sebagai salah satu syarat guna memperoleh gelar Sarjana Komputer pada Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar pada hari Kamis tanggal 29 Februari 2024.

Panitia Ujian : 19 Syaban 1445 H
Makassar, 29 Februari 2024 M

1. Pengawas Umum
 - a. Rektor Universitas Muhammadiyah Makassar
Prof. Dr. H. AMBO ASSE, M.Ag
 - b. Dekan Fakultas Teknik Universitas Hasanuddin
Prof. Dr. Eng. MUHAMMAD ISRAN RAMLI, ST., MT
2. Penguji
 - a. Ketua : Dr. Ir. Zahir Zainuddin, M.Sc.
 - b. Sekretaris : Muhyiddin A.M Hayat, S.Kom., MT.
3. Anggota :
 1. Desi Anggreani, S.Kom., MT.
 2. Rizki Yusliana Bakti ST., MT.
 3. Asyraful Insan Asry, S.Kom., MT.

Mengetahui :

Pembimbing I

Pembimbing II

Fahrim Irahma Rachman S.Kom.,M.T.

Titin Wahyuni S.Pd.,M.T.

Dekan Fakultas Teknik



Dr. H. H. Nurnawaty, ST., MT., IPM.

DEK NBM : 795 108



UNIVERSITAS MUHAMMADIYAH MAKASSAR
FAKULTAS TEKNIK

GEDUNG MENARA IQRA LT. 3

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

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

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

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

HALAMAN PERSETUJUAN

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

Judul Skripsi : **PERBANDINGAN AKURASI ALGORITMA NAÏVE BAYES DAN REGRESI LOGISTIK DALAM ANALISIS SENTIMEN WISATAWAN TERHADAP BUGIS WATERPARK ADVENTURE DI ULASAN GOOGLE MAPS**

Nama : JUMARDIN RUSLAN

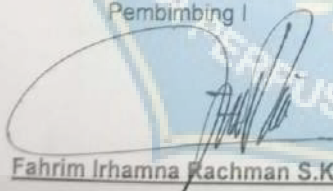
Stambuk : 105841104419


Makassar, 29 Februari 2024

Telah Diperiksa dan Disetujui
Oleh Dosen Pembimbing:

Pembimbing I

Pembimbing II


Fahrir Irhamna Rachman S.Kom.,M.T.


Titin Wahyuni S.Pd.,M.T.

Mengetahui,

Ketua Program Studi Informatika




Muhyiddin A. M. Hayat, S.kom., MT.

NBM : -

ABSTRAK

JUMARDIN RUSLAN. Perbandingan akurasi algoritma naïve bayes dan regresi logistik dalam analisis sentimen wisatawan terhadap bugis waterpark adventure di ulasan google maps (dibimbing oleh Fahrin Irhamna Rachman S.Kom.,M.T. dan Titin Wahyuni S.Pd.,M.T).

Penelitian ini bertujuan untuk membandingkan akurasi dua algoritma klasifikasi, yaitu Naïve Bayes dan Regresi Logistik, dalam menganalisis sentimen wisatawan terhadap Bugis Waterpark Adventure berdasarkan ulasan Google Maps. Metode analisis sentimen ini menggunakan data ulasan yang dikumpulkan dari platform Google Maps. Dalam penelitian ini, dilakukan tahap preprocessing untuk membersihkan data, seperti cleaning, casefolding stopwords, dan tokenizing. Selanjutnya, dilakukan pembagian dataset menjadi data latih dan data uji dengan skenario 90:10 untuk melatih dan menguji model dengan 3 kategori ulasan yaitu positif, netral dan negatif. Hasil dari perbandingan akurasi kedua algoritma menunjukkan performa yang berbeda dalam mengklasifikasikan sentimen wisatawan terhadap Bugis Waterpark Adventure. Naïve Bayes menunjukkan tingkat akurasi 77%, sementara Regresi Logistik menunjukkan akurasi yang lebih tinggi dalam menganalisis sentimen wisatawan yaitu 82%. Kesimpulannya, Regresi Logistik dapat menjadi pilihan yang lebih baik dalam menganalisis sentimen wisatawan terhadap destinasi wisata seperti Bugis Waterpark Adventure berdasarkan ulasan Google Maps.

Kata Kunci : Analisis sentimen, Akurasi algoritma, Bugis Waterpark Adventure ,Naïve Bayes, Regresi Logistik, Ulasan Google Maps, Wisatawan

ABSTRACT

JUMARDIN RUSLAN. *Comparison of the accuracy of the Naïve Bayes algorithm and logistic regression in analyzing tourist sentiment towards Bugis Waterpark Adventure in Google Maps reviews (supervised by Fahrini Irhamna Rachman S.Kom.,M.T. and Titin Wahyuni S.Pd.,M.T).*

This research aims to compare the accuracy of two classification algorithms, namely Naïve Bayes and Logistic Regression, in analyzing tourist sentiment towards Bugis Waterpark Adventure based on Google Maps reviews. This sentiment analysis method uses review data collected from the Google Maps platform. In this research, a preprocessing stage was carried out to clean the data, such as cleaning, casefolding stopwords, and tokenizing. Next, the dataset was divided into training data and test data with a 90:10 scenario to train and test the model with 3 review categories, namely positive, neutral and negative. The results of comparing the accuracy of the two algorithms show different performance in classifying tourist sentiment towards Bugis Waterpark Adventure. Naïve Bayes shows an accuracy rate of 77%, while Logistic Regression shows higher accuracy in analyzing tourist sentiment, namely 82%. In conclusion, Logistic Regression can be a better choice in analyzing tourist sentiment towards tourist destinations such as Bugis Waterpark Adventure based on Google Maps reviews.

Keywords: *Sentiment analysis, Algorithm accuracy, Bugis Waterpark Adventure, Naïve Bayes, Logistic Regression, Google Maps reviews, Tourists*

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Assalamu'alaikum Warahmatullahi Wabaraktuh

Segala puji bagi Allah Subhanallahu wa Ta'ala atas limpahan rahmat dan karunianya sehingga penulis dapat menyelesaikan Laporan Tugas Akhir yang berjudul “PERBANDINGAN AKURASI ALGORITMA *NAÏVE BAYAES* DAN *REGRESI LOGISTIK* DALAM ANALISIS SENTIMEN WISATAWAN TERHADAP BUGIS *WATERPARK ADVENTURE* DI ULASAN *GOOGLE MAPS*”. Shalawat beserta salam senantiasa penulis panjatkan kepada Nabi Muhammad SAW, yang telah membawa kita dari zaman jahiliah menuju zaman yang serba modern seperti saat ini.

Tak lupa penulis mengucapkan banyak-banyak terima kasih kepada seluruh pihak yang telah membantu dan memberikan wejangan dalam penyusunan Skripsi ini, terutama kepada :

1. Ibu **Dr.Ir.Hj Nurnawati, S.T.,M.T.,I.P.M**, selaku Dekan Fakultas Teknik.
2. Bapak **Muh. syafaat S Kuba, S.T.,M.T**, selaku Wakil Dekan Fakultas Teknik.
3. Bapak **Muhyiddin AM Hayat S.Kom.,M.T**, selaku Ketua Prodi Informatika.
4. Bapak **Fahrim Irhamna Rachman S.Kom.,M.T**, selaku Dosen Pembimbing 1 Skripsi.
5. Ibu **Titin Wahyuni S.Pd.,M.T**, selaku Dosen Pembimbing 2 Skripsi.
6. Dosen dan Staf Fakultas Teknik Universitas Muhammadiyah Makassar.
7. Teman-teman Khususnya Angkatan 2019 Fakultas Teknik, Universitas Muhammadiyah Makassar, terima kasih atas dukungan dan doanya.
8. Dan terkhusus kepada Orang tua penulis dan keluarga yang telah memberikan wejangan dan motivasi baik secara moril maupun materil.

Demikian laporan skripsi ini, dan penulis sadar bahwa laporan ini masih banyak kekurangan di dalamnya oleh karena itu penulis mengharapkan kritik dan saran yang sifatnya membangun dari pembaca atas laporan ini. Akhir kata penulis ucapkan terima kasih .

Billahi fisabililhaq, fastabiqul khairat.

Wassalamualaikum Wr.Wb.

Makassar, 16 Februari 2024

Penulis

Jumardin Ruslan



DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN PENGESAHAN	iii
LEMBAR PERSETUJUAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR LAMPIRAN	xiii
DAFTAR ISTILAH	xiv
BAB I PENDAHULUAN	1
A. Latar Belakang.....	1
B. Rumusan Masalah.....	2
C. Tujuan Penelitian.....	3
D. Manfaat Penelitian.....	3
E. Ruang Lingkup Penelitian.....	3
F. Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA	5
A. Landasan Teori.....	5
B. Penelitian Terkait.....	15
C. Kerangka Berpikir.....	17
BAB III METODOLOGI PENELITIAN	19
A. Tempat Penelitian.....	19
B. Alat Dan Bahan.....	19
C. Perancangan Sistem.....	20
D. Teknik Pengujian Sistem.....	21

E. Teknik Analisis Data.....	22
BAB IV HASIL DAN PEMBAHASAN	25
A. Pengambilan Data	25
B. Pelabelan Data.....	26
C. Tahap <i>preprocessing</i>	27
D. Klasifikasi <i>Naïve Bayes</i>	30
E. Klasifikasi <i>Regresi Logistik</i>	33
F. Hasil klasifikasi <i>Naïve Bayes</i> dan <i>Regresi Logistik</i>	36
G. Pengujian <i>Confusion Matrix</i>	38
H. Hasil Perhitungan <i>Confusion Matrix</i>	52
BAB V PENUTUP.....	53
A. Kesimpulan	53
B. Saran.....	53
DAFTAR PUSTAKA	54
LAMPIRAN.....	56



DAFTAR GAMBAR

Gambar 1. Contoh Confusion matrix	9
Gambar 2. Tahap proses preprocessing	10
Gambar 3. Contoh pelabelan sentimen	15
Gambar 4. Kerangka berpikir.....	17
Gambar 5. Flowchart system.....	20
Gambar 6. Confusion Matrix	21
Gambar 7. Scraping Data Ulasan Menggunakan Instant Data Scraper	25
Gambar 8. Dataset Ulasan.....	26
Gambar 9. Pemanggilan Model Naive Bayes	31
Gambar 10. Hasil pengujian Naive Bayes	32
Gambar 11. Hasil Prediksi Naive Bayes Dari Data Uji	33
Gambar 12. Pemanggilan Model Regresi Logistik.....	34
Gambar 13. Hasil pengujian Regresi Logistik	35
Gambar 14. Hasil Prediksi Regresi Logistik dari data uji.....	36
Gambar 15. Confusion matrix pada data uji Naïve Bayes	48
Gambar 16. Confusion matrix pada data uji Regresi Logistik.....	48
Gambar 17. Confusion matrix Naïve Bayes	49
Gambar 18. Confusion matrix Regresi Logistik	49
Gambar 19. Hasil Confusion matrix Naïve Bayes	49
Gambar 20. Hasil Confusion matrix Regresi Logistik.....	50

DAFTAR TABEL

Tabel 1. Tahap Pelabelan	26
Tabel 2. Tahap Cleaning	27
Tabel 3. Tahap Transform Cases	28
Tabel 4. Tahap Stopword	29
Tabel 5. Tahap Tokenizing	30
Tabel 6. Hasil pengujian Naïve Bayes pada data uji.....	37
Tabel 7. Hasil pengujian Regresi Logistik pada data uji	37
Tabel 8. Hasil klasifikasi dengan labelling manual dan confusion matrix	38
Tabel 9. Hasil Klasifikasi Naïve Bayes Yang True Positif.....	38
Tabel 10. Hasil Klasifikasi Naïve Bayes Yang True Negatif	41
Tabel 11. Hasil Klasifikasi Naïve Bayes Yang False Positif	42
Tabel 12. Hasil Klasifikasi Naïve Bayes Yang False Negatif	42
Tabel 13. Hasil Klasifikasi Regresi Logistik Yang True Positif.....	43
Tabel 14. Hasil Klasifikasi Regresi Logistik Yang True Negatif	46
Tabel 15. Hasil Klasifikasi Regresi Logistik Yang False Positif.....	47
Tabel 16. Hasil Klasifikasi Regresi Logistik Yang False Negatif	47
Tabel 17. Keterangan Confusion matrix Naïve Bayes dan Regresi Logistik	50
Tabel 18. Perbandingan Naïve Bayes dan Regresi Logistik.....	52

DAFTAR LAMPIRAN

Lampiran 1. Proses pengumpulan data	56
Lampiran 2. Data ulasan setelah dikumpulkan	56
Lampiran 4. Hasil Data Ulasan Setelah Dilakukan Proses Cleaning.....	57
Lampiran 3. Pelabelan data ulasan positif, netral, negatif	57
Lampiran 5. Hasil Data Ulasan Setelah Dilakukan Proses Case Folding	58
Lampiran 6. Hasil Data Ulasan Setelah Dilakukan Proses Stopword.....	58
Lampiran 7. Hasil Data Ulasan Setelah Tahap Preprocessing	59
Lampiran 8. Hasil Data Ulasan Setelah Dilakukan Proses Tokenizing	59
Lampiran 9. Proses klasifikasi Naive Bayes	60
Lampiran 10. Proses Klasifikasi Regresi Logistik	61
Lampiran 11. Hasil Pengujian Naive Bayes Pada Data Uji	61
Lampiran 12. Hasil Pengujian Regresi Logistik Pada Data Uji.....	62
Lampiran 13. Source Code.....	63
Lampiran 14. Surat Keterangan Bebas Plagiasi.....	72
Lampiran 15. Hasil Scan Plagiasi Per Bab.....	73

DAFTAR ISTILAH

- Algoritma klasifikasi* : Adalah proses menemukan sekumpulan pola atau fungsi yang mendeskripsikan dan membedakan data dengan tujuan agar model dapat digunakan untuk memprediksi kelas data 'suatu objek yang belum diketahui.
- Naïve bayes classification* : Adalah algoritma yang digunakan untuk nilai probabilitas tertinggi guna menempatkan data uji pada kategori tertinggi. *Naïve Bayes Classification* merupakan suatu metode pengklasifikasian teks. Keunggulan *NBC* adalah algoritmanya yang sederhana dan sangat tepat.
- Regresi Logistik* : *Regresi logistik* merupakan suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel respon (y) yang bersifat biner dengan variabel (x) yang bersifat polikot. Hasil dari variabel respon yang terdiri dari 2 kategori yaitu sukses dan gagal yang dinotasikan dengan $y=1$ (sukses) dan $y=0$ (gagal).
- Confusion matrix* : Adalah pengukuran penentuan klasifikasi dilakukan atas kinerja klasifikasi yang telah dilakukan. Untuk mengukur penentuan klasifikasi adalah dengan mengetahui jumlah tiap kelas-kelas prediksi dan kelas nyata yang tersusun dari TP, yaitu jumlah *tweet* dengan sentimen positif yang telah benar di golongan positif, dan TN yaitu *tweet* yang

bersentimen positif. perasaan yang diprediksi negatif.

Scikit-learn : Adalah perpustakaan *open-source* yang sangat populer untuk pemrosesan data dan pembelajaran mesin di lingkungan pemrograman *Python*. *Scikit-learn* dirancang untuk memudahkan penggunaan algoritma pembelajaran mesin dalam berbagai konteks. *Library* ini mencakup berbagai algoritma klasifikasi, *regresi*, *klustering*, dan pengolahan data lainnya. Dengan antarmuka yang intuitif, pengguna dapat dengan mudah menerapkan dan menyesuaikan model sesuai kebutuhan.

Machine learning : Adalah Pembelajaran mesin, sekumpulan algoritma pemrograman yang digunakan untuk mengoptimalkan kinerja komputer atau sistem berdasarkan sampel data yang sudah ada sebelumnya. Pembelajaran mesin memiliki 7 langkah, antara lain mengumpulkan persiapan data masukan, menganalisis data masukan, manusia, melatih algoritma, menguji algoritma penggunaannya.

Supervised learning : Adalah jenis pendekatan dalam pembelajaran mesin di mana algoritma atau model dilatih menggunakan data yang telah dilabeli. Artinya, setiap contoh data dalam dataset pelatihan telah diberikan label atau jawaban yang benar, sehingga algoritma dapat belajar untuk menghasilkan prediksi atau output

yang sesuai ketika diberikan data baru yang belum pernah dilihat sebelumnya.

- Flowchart* : Adalah *Modelling* atau proses merancang dan membangun sebuah sistem yang bertujuan untuk memenuhi kebutuhan dan memecahkan sebuah masalah. *Flowchart* melibatkan pemilihan teknologi yang tepat, arsitektur sistem, desain antarmuka pengguna, pemilihan metode dan algoritma pemrograman, dan pengujian sistem secara menyeluruh.
- Pelabelan : Adalah mengacu pada sekumpulan kepribadian maupun variable yang digunakan guna mengenali variable maupun bagian tertentu dari berkas maupun informasi. Pelabelan menggambarkan Langkah awal dalam analisis sentimen, Dimana bacaan diklasifikasikan ataupun dilabeli bersumber pada perasaan ataupun sentimen yang terdapat didalamnya.
- Data : Merupakan sekelompok informasi atau fakta mentah yang dapat berupa simbol, angka, kata-kata, atau citra. Informasi ini diperoleh melalui proses pengamatan atau pencarian dari sumber-sumber tertentu.

BAB I

PENDAHULUAN

A. Latar Belakang

Bugis *Waterpark Adventure*, lokasi wisata terkenal di Sulawesi selatan, menawarkan berbagai wahana menarik, termasuk seluncuran air, kolam renang, dan aliran sungai yang memanjakan mata. Namun, di era komputerisasi, ulasan online telah berubah menjadi sumber informasi yang penting, khususnya yang berkaitan dengan industri pariwisata. Mengingat semakin populernya platform seperti *Google Maps*, anggapan konsumen pada dasarnya dapat mempengaruhi citra dan popularitas destinasi wisata, termasuk Bugis *Waterpark Adventure*.

Dalam konteks ini, Penelitian yang sah terhadap sentimen pengunjung sangat penting untuk memahami kesan individu terhadap keberatannya. Salah satu anggapan yang sering muncul dalam analisis sentimen diketahui bahwa anggapan informasi yang baik umumnya berisi ulasan yang baik, begitu pula sebaliknya. Namun, secara umum hal ini tidak sering terjadi. Ada situasi di mana ulasan positif mungkin menyertakan analisis atau ketidakpuasan, sementara ulasan negatif mungkin juga menyertakan pengakuan atau sudut pandang positif. Memperkuat alasan bahwa ulasan bagus belum tentu baik dan ulasan negatif belum tentu buruk menjadi fokus utama dalam kasus ini.

Dua algoritma yang diterapkan dalam Penelitian opini diketahui bahwa *Naive Bayes* dan *Regresi Logistik*. Kedua perhitungan ini memiliki pendekatan dan sorotan untuk menampilkan informasi dan mengantisipasi opini dari teks ulasan. Mengingat kerumitan ini, pemahaman tingkat ketepatan antara *Naive Bayes* dan *Regresi Logistik* dalam analisis sentimen terhadap Bugis *Waterpark Adventure* sangatlah penting.

Analisis sentimen, juga disebut pengelompokan penilaian, diketahui bahwa bidang studi yang menyelidiki sudut pandang, opini, penilaian, keputusan, mentalitas, dan perasaan individu terhadap elemen-elemen seperti objek, administrasi, asosiasi, orang, isu, peristiwa, tema, dan kualitasnya (Liu, 2012). Analisis sentimen merupakan siklus terprogram untuk memahami, memisahkan dan menangani informasi pesan yang

ditentukan untuk mendapatkan data tentang perasaan yang terkandung dalam suatu penjelasan atau penilaian (Khofifah et al., 2022). Besarnya pengaruh dan manfaat ujian perasaan membawa perkembangan pesat dalam komposisi dan penyempurnaan aplikasi berbasis ujian pendapat (Lorosae & Prakoso, 2018).

Dalam analisis sentimen, ada beberapa perhitungan yang sering digunakan antara lain *Naïve Bayes* dan *Regresi Logistik*, dimana perhitungan tersebut dapat digunakan dalam analisis sentimen. Dimana *Naïve Bayes* yang keunggulan dalam efisiensi dan kemudahan pada pengklasifikasian teks, khususnya dalam aplikasi fungsional langsung, misalnya, memisahkan klasifikasi berita atau memilah spam (Sang-Bum Kim, Kyoung-Soo Han, Hae-Chang Rim, 2006). Sementara itu, *Regresi Logistik* merupakan suatu model statistik yang digunakan untuk menilai apakah variabel independen memiliki pengaruh pada variabel dependen yang bersifat *biner*.

Dari latar belakang di atas penelitian ini dibuat untuk mengetahui akurasi perbandingan dari kedua algoritma yaitu *naïve bayes* dan *regresi logistik (Logistic regression)* pada analisis sentimen wisatawan Bugis waterpark adventure di ulasan Google maps.

B. Rumusan Masalah

Berdasarkan latar belakang penelitian di atas, peneliti dapat merumuskan masalah sebagai berikut:

1. Bagaimana penerapan algoritma *naïve bayes* dan *regresik logistik* pada analisis sentimen wisatawan terhadap bugis *waterpark adventure* di ulasan *google maps*?
2. Apakah ada perbedaan antara akurasi algoritma *Naïve Bayes* dan *Regresi Logistik* dalam menganalisis sentimen ulasan wisatawan terhadap Bugis *Waterpark Adventure* di ulasan *Google Maps*?
3. Bagaimana tingkat akurasi dari algoritma *naïve bayes* dan *regresik logistik* dalam analisis sentimen wisatawan terhadap bugis *waterpark adventure* di ulasan *google maps*?

C. Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah penelitian di atas, peneliti dapat memberikan tujuan penelitian sebagai berikut:

1. Untuk mengetahui Bagaimana penerapan algoritma *naïve bayes* dan *regresik logistik* pada analisis sentimen wisatawan terhadap bugis *waterpark adventure* di ulasan *google maps*?
2. Mengetahui perbedaan antara akurasi algoritma *Naïve Bayes* dan *Regresi Logistik* dalam menganalisis sentimen ulasan wisatawan terhadap Bugis *Waterpark Adventure* di ulasan *Google Maps*.
3. Untuk mengetahui tingkat akurasi dari algoritma *naïve bayes* dan *regresik logistik* dalam analisis sentimen wisatawan terhadap bugis *waterpark adventure* di ulasan *google maps*.

D. Manfaat Penelitian

Adapun beberapa Manfaat penelitian ini bagi peneliti sebagai berikut :

- a. Penelitian ini dapat memberikan pemahaman perbedaan tentang bagaimana algoritma *naïve bayes* dan *regresik logistik* bekerja dalam analisis sentimen wisatawan terhadap bugis *waterpark adventure* di ulasan *google maps*.
- b. Penelitian ini dapat memberikan pemahaman untuk menentukan algoritma yang lebih optimal dalam analisis sentimen.

E. Ruang Lingkup Penelitian

Ruang lingkup penelitian adalah batasan atau ruang lingkup yang dijadikan fokus penelitian. Dengan bantuan daerah penelitian, wilayah atau bidang yang akan dipelajari ditentukan dan batas-batas objek atau fenomena yang diamati ditentukan.

1. Penelitian ini membutuhkan data ulasan wisatawan pada bugis *waterpark adventure* di *google maps*.

2. Penelitian ini menggunakan algoritma *naïve bayes* dan *regresi logistik* sebagai perbandingan pada analisis sentimen ulasan wisatawan pada bugis *waterpark adventure* di *google maps*.

F. Sistematika Penulisan

Sistematika penulisan dalam penyusunan laporan proposal penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menguraikan tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini menguraikan mengenai studi pustaka terhadap teori-teori yang digunakan dalam penelitian serta aspek-aspek yang mendukung dalam penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini berisi penjelasan tentang langkah-langkah yang digunakan dalam penelitian, dimulai dari pengidentifikasian masalah, pengumpulan data, analisa algoritma, dan alat ukur yang digunakan dalam penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi tentang hasil penelitian, pembahasan, dan hasil implementasi sistem.

BAB V PENUTUP

Bab ini berisi kesimpulan dan saran berdasarkan hasil penelitian.

BAB II

TINJAUAN PUSTAKA

A. LANDASAN TEORI

1. Analisis Sentimen

Analisis sentimen ataupun opinion mining merujuk pada bidang yang luas dalam pengolahan bahasa natural, komputasi linguistik, serta *text mining*. Tujuannya adalah untuk menganalisis pendapat, sentimen, evaluasi, sikap, penilaian, dan emosi seseorang terhadap suatu topik, produk, layanan, organisasi, individu, atau kegiatan tertentu, baik dalam konteks bicara maupun tulisan. Tugas utama dalam analisis sentimen adalah untuk mengklasifikasikan polaritas teks dalam dokumen, kalimat, atau fitur/tingkat aspek. Ini dilakukan dengan menentukan apakah pendapat yang terkandung dalam dokumen, kalimat, atau fitur tersebut bersifat positif atau negatif (Fazrin et al., 2022). Sentiment analysis membantu dalam memahami persepsi dan penilaian orang terhadap berbagai hal. bisa diklasifikasikan ke dalam kelas sentimen bersifat positif serta negatif.

- a. Sentimen Positif: Menurut Kamus Besar Bahasa Indonesia (KBBI) sentimen positif merupakan reaksi atau sikap yang meningkatkan nilai seseorang atau sesuatu.
- b. Sentimen Negatif: Menurut Kamus Besar Bahasa Indonesia (KBBI) sentimen negatif merupakan reaksi atau sikap yang menurunkan nilai seseorang atau sesuatu, jadi kalimat bersentimen negatif akan menyebabkan penyurutan nilai pandang terhadap sesuatu, sehingga membentuk *tren down*. Umumnya kalimat bersentimen negatif ditandai dengan penggunaan kata negasi. Negasi merupakan sesuatu yang dikenal dalam semua bahasa dan biasanya negasi digunakan untuk mengubah polaritas dari suatu pernyataan. (L. Ardiani, H. Sujaini, 2020).

Tugas dasar analisis sentimen adalah mengelompokkan teks menjadi suatu kalimat atau dokumen, untuk menentukan apakah pendapat yang terkandung dalam kalimat atau dokumen tersebut positif atau negatif.

Selain itu, analisis perasaan juga dapat mencakup identifikasi emosi seperti seperti kebahagiaan, kesedihan. Hal ini juga memungkinkan kita untuk mencari pandangan tentang produk, merek, atau individu, dan menentukan apakah pandangan tersebut bersifat positif atau negatif dalam konteks daring. Sentimen atau ekspresi emosional sering kali terkait dengan topik tertentu, sehingga pernyataan yang sama mungkin memiliki makna yang berbeda ketika dibahas dalam konteks yang berbeda. Oleh karena itu, dalam beberapa penelitian, tahap awal melibatkan identifikasi elemen-elemen yang sedang dibahas sebelum proses analisis pendapat dimulai (DARUSSALAM, 2021).

2. Algoritma Klasifikasi

Klasifikasi adalah proses menemukan sekumpulan pola atau fungsi yang mendeskripsikan dan membedakan data dengan tujuan agar model dapat digunakan untuk memprediksi kelas data 'suatu objek yang belum diketahui (Han et al., 2012).

3. *Naïve bayes*

Teorema Bayes merupakan teorema yang berkaitan dengan konsep probabilitas. Secara umum, *Teorema Bayes* dapat dinyatakan sebagai persamaan berikut (Siang, 2005).

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \dots\dots\dots(1)$$

Klasifikasi :

A : ilustrasi informasi yang label kelasnya tidak dikenal.

B : kelas-kelas hasil klasifikasi.

(A|B) : probabilitas terbentuknya *A* bila *B* tidak dikenal.

(B|A) : probabilitas terbentuknya *B* bila *A* tidak dikenal.

P(A) : probabilitas *prior* *A* yang mendahului terbentuknya *B*. di ucap "*prior*", sebab nilainya dapat diperoleh ciri butuh memikirkan informasi maupun penimba *B* terlebih dulu *P(A)* pula berarti probabilitas ini diperoleh dari informasi ilustrasi yang sudah dikenal berkelas *A*.

(B) : probabilitas *pior* *B*, serta berperan selaku *normalizing contant*.

Algoritma *Naïve Bayes Classifier* merupakan algoritma yang digunakan untuk nilai probabilitas tertinggi guna menempatkan data uji pada kategori tertinggi (Feldman, Ronen, 2007). Metode *Naïve Bayes Classification* merupakan suatu metode pengklasifikasian teks. Keunggulan *NBC* adalah algoritmanya yang sederhana dan sangat tepat. Ada dua langkah dalam klasifikasi. Langkah pertama adalah pelatihan pada *tweet* yang kategorinya sudah diketahui. Dalam algoritma *NBC*, setiap dokumen diwakili oleh salah satu atribut “1, a_3 , ..., a_n ” dimana a_1 adalah kata pertama, kata kedua, dan seterusnya (Falahah & Nur, 2015). Sedangkan waktu, V adalah sekumpulan kategori *tweet*. Pada disaat klasifikasi algoritma akan mencari probabilitas paling tinggi dari seluruh bagian kategori dokumen yang diujikan. Adapun persamaan *VMAP* adalah sebagai berikut.

$$V_{MAP} = \frac{\arg \max}{v_j=V} P(v_j) \prod_i P(a_i | v_j) \dots\dots\dots (2)$$

Nilai $P(v_j)$ dihitung pada sesaat *training*, didapat dengan rumus sebagai berikut:

$$P(v_j) = \frac{|doc\ j|}{|training|} \dots\dots\dots (3)$$

Dimana $|doc\ j|$ merupakan jumlah *tweet* yang memiliki kategori j dalam *training*. Sedangkan $|training|$ merupakan jumlah *tweet* dalam yang digunakan untuk *training*. Untuk setiap probabilitas kata a_i untuk setiap kategori ($a_i | v_j$), dihitung pada saat *training*.

$$P(a_i | v_j) = \frac{n_i+1}{|n+ Kosakata|} \dots\dots\dots (4)$$

dimana n_i adalah jumlah kemunculan kata a_i dalam *tweet* yang berkategori v_j dan $|kosakata|$ adalah banyaknya kata dalam *training*.

4. Regresi Logistik

Regresi logistik biner merupakan suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel respon (y) yang bersifat biner dengan variabel (x) yang bersifat polikotomus (David W. Hosmer, Jr., Stanley Lemeshow, 2013). Hasil dari variabel respon yang terdiri dari 2 kategori yaitu sukses dan gagal yang dinotasikan dengan y=1 (sukses) dan y=0 (gagal). Oleh karena itu, variabel y mengikuti distribusi Bernoulli untuk setiap observasi tunggal. Fungsi probabilitas untuk setiap observasi dapat dituliskan pada persamaan sebagai berikut.

$$f(y) = \pi^y(1 - \pi)^{1-y}; y = 0,1 \dots\dots\dots(5)$$

dimana y adalah variabel respon jika y=0 maka f(y) = 1 - π dan jika y = 1 maka f(y) = π. Fungsi regresi logistiknya dapat dituliskan pada persamaan selanjutnya.

$$f(z) = \frac{1}{1 + e^{-z}} \text{ ekuivalen } f(z) = \frac{e^{-z}}{1 + e^{-z}}$$

dimana, $z = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$ dimana p ialah jumlah variabel prediktor. Z adalah antara $-\infty$ dan $+\infty$ sehingga nilai f adalah antara 0 dan 1 untuk setiap nilai z yang diberikan. Hal ini menampilkan jika model *logistik* menggambarkan resiko probabilitas sesuatu objek sehingga model tersebut dapat dituliskan ke dalam persamaan.

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 x_1)}}{1 + e^{(\beta_0 + \beta_1 x_1)}} \dots\dots\dots(6)$$

β_p adalah parameter estimasi ke-p. Untuk memudahkan estimasi parameter *regresi*, model *Regresi Logistik* pada persamaan tersebut dapat diurutkan menggunakan transformasi logit $\pi(x)$ sehingga diperoleh persamaan berikut;

$$g(x) = \ln \left(\frac{\pi(x)}{1 - \pi(x)} \right) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \dots\dots\dots(7)$$

Model tersebut merupakan fungsi linier dari parameter-parameternya. Pada *regresi logistik*, variabel respon diekspresikan sebagai $y = \pi(x) + \varepsilon$ dimana ε mempunyai salah satu dari kemungkinan dua nilai yaitu $y = \pi(x) + \varepsilon$ dengan peluang $\pi(x)$ jika $y = 1$ dan $\varepsilon = -\pi(x)$ dengan peluang $1 - \pi(x)$ jika dan mengikuti distribusi binomial dengan rata-rata nol dan varians $(\pi(x))(1 - \pi(x))$.

5. *Confusion matrix*

pengukuran penentuan klasifikasi dilakukan atas kinerja klasifikasi yang telah dilakukan. Untuk mengukur penentuan klasifikasi adalah dengan mengetahui jumlah tiap kelas kelas prediksi dan kelas nyata yang tersusun dari TP yaitu jumlah *tweet* dengan sentimen positif yang telah benar di golongan positif, TN yaitu *tweet* yang bersentimen positif. perasaan yang diprediksi negatif (Kurniawan, 2017).

Berikut ini adalah *matrix konfusi* yang memiliki empat nilai tersebut.

Kelas Aktual	Kelas Prediksi	
	Positif	Negatif
Positif	TP	FN
Negatif	FP	TN

Gambar 1. Contoh *Confusion matrix*

Matrix yang sering digunakan untuk menghitung akurasi adalah akurasi, spesifisitas, dan sensitivitas/recall. Akurasi adalah persentase dokumen yang teridentifikasi dengan benar terhadap total dokumen dalam proses klasifikasi. Akurasi digunakan untuk menghitung keakuratan klasifikasi suatu dokumen yang berisi data seimbang di setiap kategori. Berikut rumus untuk menghitung presisi, sensitivitas, dan sensitivitas/recall.

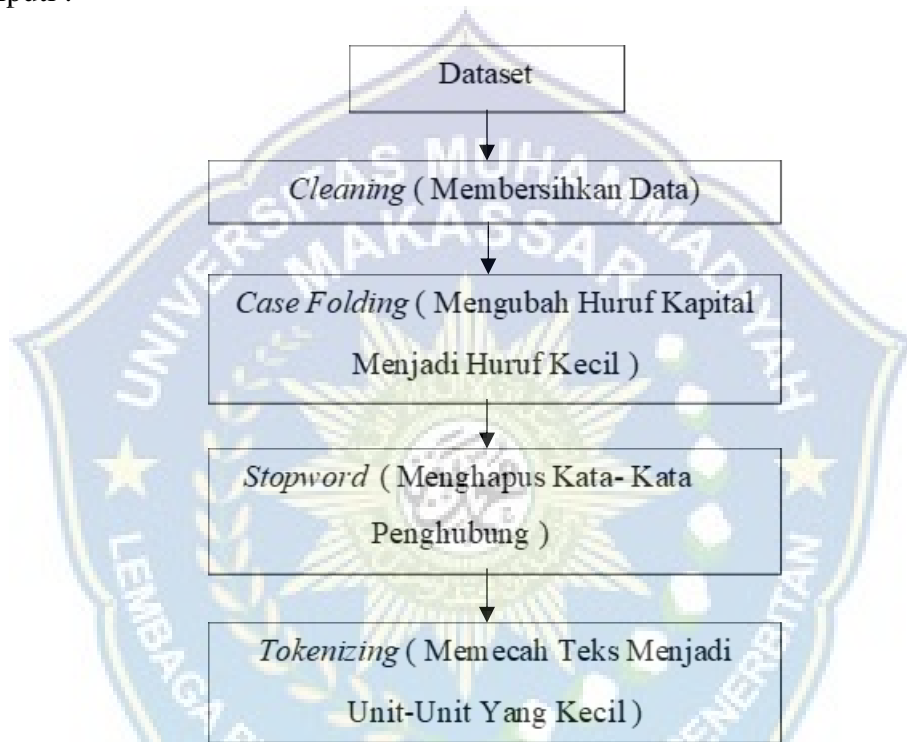
$$accuracy = \frac{TP+TN}{Total} \dots\dots\dots(8)$$

$$precision = \frac{TP}{TP+FP} \dots\dots\dots(9)$$

$$reccal = \frac{TP}{TP+FN} \dots\dots\dots(10)$$

6. *Preprocessing* data

Data yang diperoleh selama *web scraping* memiliki ketidakteraturan dan selalu sewenang-wenang atau semi terstruktur. Oleh karena itu, sebelum data diproses berikut proses untuk dimasukkan ke dalam model, data harus dibersihkan terlebih dahulu melalui *preprocessing* agar data terstruktur (Harjanta, 2015). Langkah *preprocessing* meliputi :



Gambar 2. Tahap proses *preprocessing*

a. *Cleaning*

cleaning adalah proses menyiapkan data sehingga dapat digunakan untuk melakukan analisis sentimen secara akurat dan efisien. Langkah pembersihan bertujuan untuk membersihkan data dari data yang tidak relevan, tidak valid atau tidak diperlukan. Langkah pembersihan, membersihkan informasi dari tanda baca, *hashtag* atau penyebutan.

b. *Case folding*

Dengan fungsi *case folding*, peneliti dapat mengubah semua huruf secara otomatis menjadi huruf kecil atau dalam teks adalah huruf kapital untuk pencarian semua huruf diubah menjadi huruf kecil karena Sebagian besar teks adalah opini, kebanyakan ditulis dengan huruf kecil semua.

c. *Stopwords*

Dengan fitur ini, sebelum diklasifikasikan, teks yang tidak terkait dengan analisis sentimen dihilangkan sehingga dimensi teks tersebut diperkecil tanpa mengurangi sentimentalitas teks tersebut (Harjanta, 2015). Preposisi dan konjungsi merupakan kandidat besar dalam daftar stop word yang sebaiknya dihilangkan. Untuk dokumen berbahasa Indonesia, contoh linknya adalah “siapa”, “dalam”, “dan”, “itu”, “dengan”. Langkah ini berguna untuk mengurangi jumlah yang akan digunakan.

d. *Tokenize*

Selama tahap ini, kalimat dipisahkan menjadi beberapa bagian sebelum dianalisis lebih detail. Tujuannya adalah mempersiapkan teks sedemikian rupa sehingga dapat diolah dengan lebih mudah, tepat, dan efektif dalam perasaan. Tokenisasi memungkinkan kita memahami teks dalam potongan yang lebih kecil dan menerapkan teknik analisis sentimen yang canggih.

7. *Scikit-Learn*

Scikit-learn adalah perpustakaan *open-source* yang sangat populer untuk pemrosesan data dan pembelajaran mesin di lingkungan pemrograman *Python*. Dengan fokus pada kesederhanaan dan kinerja, *scikit-learn* dirancang untuk memudahkan penggunaan algoritma pembelajaran mesin dalam berbagai konteks. *Library* ini mencakup berbagai algoritma klasifikasi, *regresi*, *klustering*, dan pengolahan data lainnya. Dengan antarmuka yang intuitif, pengguna dapat dengan mudah menerapkan dan menyesuaikan model sesuai kebutuhan.

Scikit – Learn juga merupakan Modul *Python* yang mengintegrasikan berbagai algoritma pembelajaran mesin untuk masalah rata-rata yang diawasi dan tidak diawasi

Paket ini bertujuan untuk menghadirkan pembelajaran mesin kepada non-spesialis menggunakan bahasa umum tingkat tinggi. Penekanannya ditempatkan pada kemudahan penggunaan, kinerja, dokumentasi, dan konsistensi API. Ia memiliki ketergantungan minimal dan didistribusikan di bawah Lisensi BSD Sederhana, mendorong penggunaannya dalam konteks komersial. (Riadi Silitonga et al., 2019).

Library dibentuk diatas *scipy* (*Scientific Python*) yang wajib diinstal saat sebelum memakai *scikit-learn*. Tumpukan ini meliputi :

1. *Numpy* : paket bawah buat *array* n-dimensi.
2. *Scipy* : Pustaka fundamental buat komputasi ilmiah.
3. *Matplotlib* : penggambaran 2D/3D yang komprehensif.
4. *Ipython* : kenaikan pada konsol interaktif.
5. *Sympy* : pemrosesan matematika simbolik.
6. *Pandas* : struktur serta analisis informasi ekstensi ataupun materi buat *scipy* secara universal diucap selaku *scikit*. Materi ini sediakan algoritma Pendidikan mesin serta diberi nama *scikit-learn*.

8. Sumber data

Sumber Data, informasi yang digunakan bersumber dari *Google Maps*, *Google Maps* merupakan layanan *website* yang sediakan pemetaan serta navigasi. Layanan ini dibesarkan oleh industry *Google Inc*. *Google Maps* membolehkan pengguna buat memandang posisi dalam fashion satelit, gambar cuaca, tampilan jalan dalam mode 360° (*street view*), serta menyediakan informasi tentang kondisi lalu lintas secara real time. Selain itu, *Google Maps* juga menawarkan sistem navigasi dengan petunjuk arah yang dapat digunakan untuk merencanakan perjalanan dengan estimasi waktu tiba yang disesuaikan dengan jenis kendaraan yang digunakan. *Google Maps* adalah layanan peta online yang dapat diakses secara gratis dan hanya memerlukan koneksi internet (Sentimen et al., 2022).

9. Bugis *waterpark adventure*

Bugis *Waterpark Adventure* Makassar berdiri sejak tahun 2011. Tempat rekreasi ini mengusung konsep unik yang menggabungkan unsur alam dan budaya lokal yang khas. Oleh karena itu, selain menjadi destinasi untuk berwisata dan bermain air, tempat ini juga dapat dijadikan sebagai sarana wisata edukasi yang mengenalkan budaya dan sejarah setempat. Selain memiliki nama yang mencirikan tempatnya, beberapa wahana dan atraksi yang tersedia di Bugis *Waterpark Adventure* juga diberi istilah khusus yang mencerminkan budaya daerah. Beberapa contohnya adalah Phinisi Slide, Mappetang, dan Tappasorong. Ketika Anda memasuki kawasan Bugis *Waterpark Adventure*, Anda akan segera merasakan atmosfer alami yang khas pedesaan Bugis. Bangunan-bangunan di tempat ini dibangun dengan menggunakan bahan seperti kayu dan bambu dengan atap yang terbuat dari daun rumbia. Tempat wisata ini dibagi menjadi tiga zona yang berbeda, masing-masing memiliki ciri khasnya sendiri. Zona tersebut meliputi Kampong Rilau (Kampung Nelayan), Alabuang (pelabuhan), dan Goa Batu Kalibampa (Goa Kupu-kupu) (Ainun, 2022).

10. *Machine Learning*

Pembelajaran mesin adalah sekumpulan algoritma pemrograman yang digunakan untuk mengoptimalkan kinerja komputer atau sistem berdasarkan sampel data yang sudah ada sebelumnya (Alpaydin, 2020). Pembelajaran mesin memiliki 7 langkah, antara lain mengumpulkan persiapan data masukan, menganalisis data masukan, manusia, melatih algoritma, menguji algoritma penggunaannya (Harrington, 2012). Prinsip dasar pembelajaran mesin adalah menggunakan data untuk membuat model statistik. Model ini digunakan oleh sistem untuk membuat prediksi masa depan berdasarkan data masa lalu yang dimasukkan atau untuk mempelajari konten dalam data. Salah satu manfaat utama pembelajaran mesin adalah kemampuan untuk memodifikasi dan beradaptasi sebagai respons terhadap perubahan data. (Marsland, 2015).

11. *Supervised Learning*

Supervised learning adalah jenis pendekatan dalam pembelajaran mesin di mana algoritma atau model dilatih menggunakan data yang telah dilabeli. Artinya, setiap contoh data dalam dataset pelatihan telah diberikan label atau jawaban yang benar, sehingga algoritma dapat belajar untuk menghasilkan prediksi atau output yang sesuai ketika diberikan data baru yang belum pernah dilihat sebelumnya.

Proses ini melibatkan dua tahap utama: pelatihan, di mana model mempelajari pola dalam data pelatihan, dan evaluasi atau pengujian, di mana model diuji pada dataset yang tidak terlihat selama pelatihan untuk mengevaluasi kinerjanya. Tugas *supervised learning* melibatkan prediksi atau klasifikasi data berdasarkan contoh-contoh yang telah dilabeli, seperti klasifikasi email atau prediksi harga rumah. Keuntungan *supervised learning* termasuk kemampuan untuk membuat prediksi yang akurat, sementara kelemahannya adalah ketergantungan pada data yang telah dilabeli.

12. Teknik Pelabelan

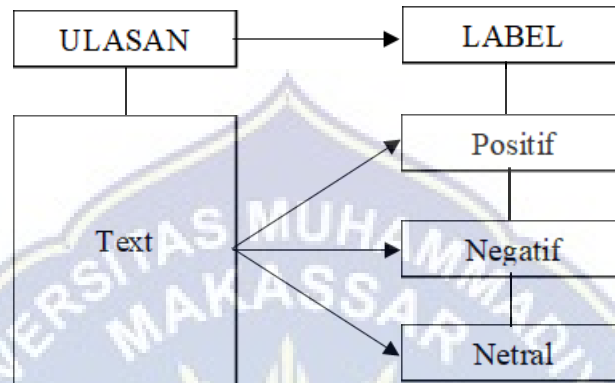
Pelabelan yakni sebutan yang berasal dari proses pemberian label pada informasi. Label mengacu pada sekumpulan kepribadian maupun variable yang digunakan guna mengenali variable maupun bagian tertentu dari berkas maupun informasi (Gunawan et al., 2018). Pelabelan menggambarkan Langkah awal dalam analisis sentimen, Dimana bacaan diklasifikasikan ataupun dilabeli bersumber pada perasaan ataupun sentimen yang terdapat didalamnya. Pelabelan sanggup dicoba dengan Teknik manual maupun secara otomatis.

Pada tahap labeling data, label diberikan secara manual, ini dilakukan dengan menambah label sentiment (positif, negatif, maupun netral) pada data ulasan Bugis *Waterpark Adventure*. Data yang diberikan label diusahakan agar jumlah label positif, negative maupun netral seimbang.

Teknik pelabelan merujuk pada proses menentukan label atau kategori yang sesuai untuk setiap contoh data dalam dataset. Pada penelitian ini memiliki 2 dataset yaitu dataset uji (*testing*) dan dataset latih (*training*). Dimana untuk dataset latih

(*training*) memiliki label yaitu positif, netral dan negatif untuk dijadikan data latih pada model atau algoritma untuk dipelajari dan akan di uji ke akurasiannya menggunakan dataset uji (*testing*).

Proses pelabelan dataset pada ulasan dapat dilihat pada gambar dibawah ini:



Gambar 3. Contoh pelabelan sentimen

B. Penelitian Terkait

Peneliti memberikan segudang inspirasi dan informasi untuk penyempurnaan proposal ini dari penelitian sebelumnya. Penelitian yang terkait sebelumnya meliputi:

Penelitian pertama oleh (Fazrin et al., 2022). “Perbandingan *Algoritma K-Nearest Neighbor* dan *Logistic Regression* pada Analisis Sentimen terhadap Vaksinasi Covid-19 pada Media Sosial Twitter dengan Pelabelan Vader dan Textblob” Pada penelitian ini mampu memberikan hasil perbandingan sentimen positif dan sentimen negatif pada kedua algoritma adalah sentimen negatif lebih besar dari sentimen positif. Terdapat 64% sentimen negatif pada hasil klasifikasi menggunakan TextBlob dan terdapat 56% sentimen negatif pada hasil klasifikasi menggunakan VaderSentiment. Hal itu dikarenakan banyaknya masyarakat yang tidak setuju dengan program vaksinasi Covid-19.

Penelitian kedua oleh (Khofifah et al., 2022) “Analisis Sentimen Menggunakan *Naive Bayes* Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan *Google Maps*”. Dalam penelitian ini dapat

memberikan hasil dimana dalam Perhitungan Naïve Bayes yang dilakukan bahwa dari 5 rentang yang dianalisis, terdapat terdapat rentang yang mempunyai bernilai negatif yaitu yang Cibendo dengan hasil d review negatif sebesar 0.550 Pantai Tanjung Baru memperoleh hasil review sebesar 0.650 dan 3 pantai yang bernilai positif yaitu pantai Samudera Baru memperoleh hasil positif 0.850, pantai Sedari 0.700 dan kisaran Pakis memperoleh hasil sebesar 0,650. Setelah dilakukan pengujian di menggunakan aplikasi RapidMiner, Cibendo mendapatkan akurasi 55%, jangkauan Pakis mendapatkan akurasi 65%, Samudera Baru mendapatkan akurasi 85%, Sedari mendapatkan Akurasi 70% dan Tanjung Baru mempunyai akurasi 70%. akurasi 65%.

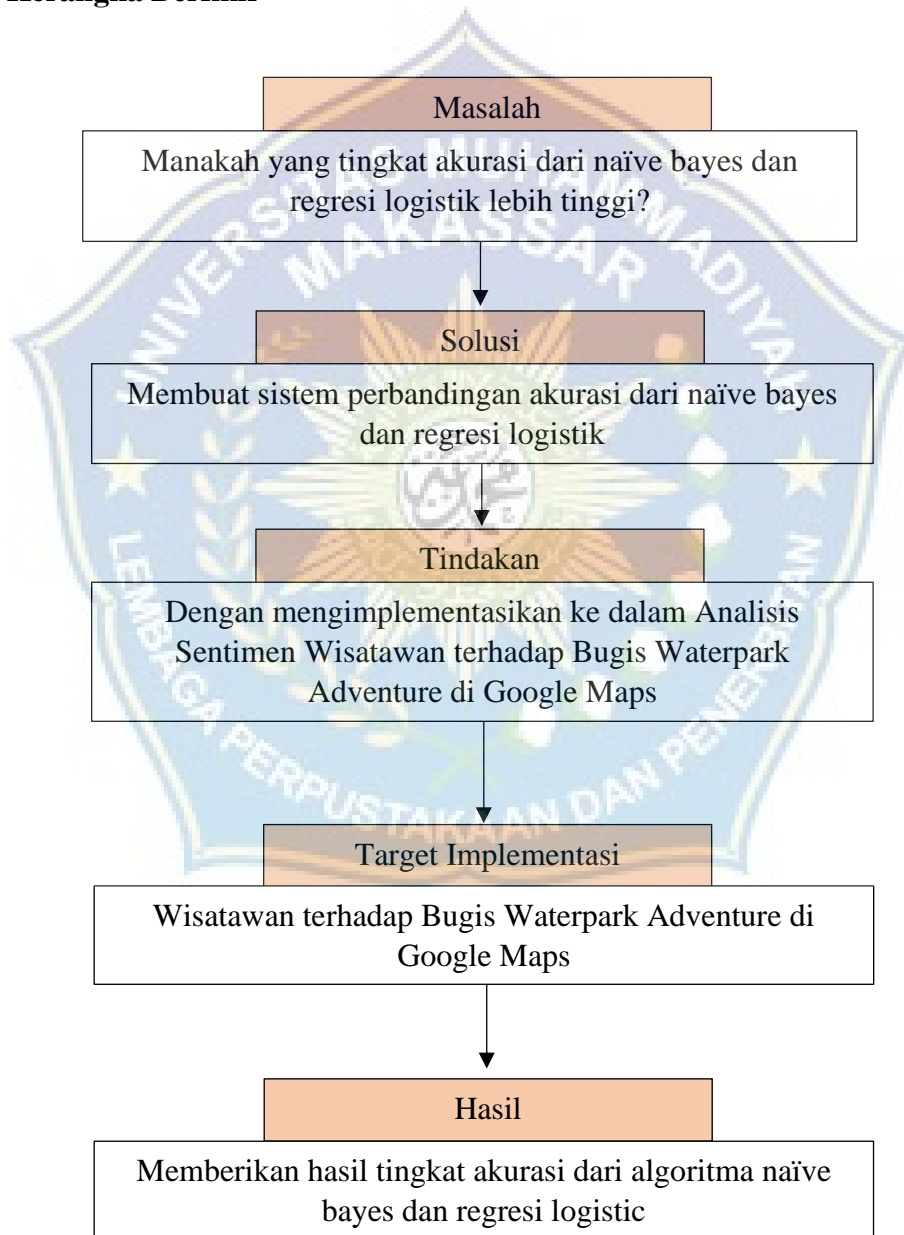
Penelitian ketiga oleh (Sentimen et al., 2022). “Analisis Sentimen Objek Wisata Bali Di *Google Maps* Menggunakan Algoritma *Naive Bayes* Dian” dimana hasil akhir yang di dapat kan yaitu pelaksanaan algoritma Naïve Bayes pada pencarian ini menunjukkan nilai presisi 5 tempat wisata diatas rata-rata dan baik pasti dalam pencarian nilai presisinya cukup baik. Analisis sentimen menggunakan algoritma Naïve Bayes berhasil merangking ulasan wisatawan di *Google Maps*. Hasil pemeringkatan menunjukkan 5 tempat wisata ke Nusa penida merupakan objek wisata yang direkomendasikan. Karena mempunyai akurasi sebesar 94,64%. Diantaranya adalah Garuda Wisnu Kencana dengan akurasi 82,86%, The Edge dengan akurasi 80%, Pandawa dengan akurasi 90,71%, Pura Luhur Uluwutu dengan akurasi 85,54 %.

Penelitian keempat oleh (DARUSSALAM, 2021). “ANALISIS SENTIMEN MENGGUNAKAN *TEXT MINING* DENGAN METODE *NAÏVE BAYES* DAN *REGRESI LOGISTIK*”. Dalam penelitian ini mendapatkan hasil akhir dimana akurasi klasifikasi pada menggunakan Naïve classifier pada data vaksin Covid-19 tercapai. Rerata G dan AUC sebesar 63,75% , dan 62,12% serta regresi logistik biner pada data vaksin Covid19 diperoleh presisi , G dan AUC sebesar 72,6% , 69,9 dan %.

Penelitian kelima oleh (Gunawan et al., 2018) “Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode *Naive Bayes*”. Pada penelitian ini mendapat kan hasil bahwa Metode *Naive Bayes* dapat memprediksi kelas sentimen pada ulasan produk online sesuai dengan sistem yang disiapkan. Sistem analisis sentimen pada

ulasan produk online menggunakan metode *Naive Bayes* menghasilkan nilai akurasi terendah pada pengujian 5 kelas menggunakan dataset 80% latih dan 20% data uji sebesar 52.66%, sedangkan pada pengujian 3 kelas menggunakan dataset 90% data latih dan 10% data uji memiliki akurasi tertinggi sebesar 77.78%.

C. Kerangka Berfikir



Gambar 4. Kerangka berpikir

Pada gambar 4. yaitu diagram kerangka pikir menjelaskan tentang kerangka pikir dari penelitian ini yaitu pertama masalah yaitu menentukan manakah antara algoritma *naïve bayes* dan *regresi logistik* yang memiliki tingkat akurasi paling tinggi dalam analisis sentimen wisatawan terhadap bugis *waterpark adventure* di *google maps*. selanjutnya solusi yang di berikan yaitu membuat sistem perbandingan akurasi dari *naïve bayes* dan *regresi logistik* untuk mengetahui keakurasian nya dalam menganalisis sentimen. Dengan taerget implementasi yaitu Wisatawan terhadap Bugis *Waterpark Adventure* di *Google Maps*, serta penelitian ini menghasilkan tingkat akurasi dari algoritma *naïve bayes* dan *regresi logistik*.



BAB III

METODE PENELITIAN

A. Tempat Penelitian

Tempat penelitian merupakan hal yang sangat penting yang harus ditentukan dalam proses penelitian. faktor ini berperan penting dalam menentukan validitas dan akurasi hasil penelitian yang dilakukan.

Tempat Penelitian.

Tempat penelitian merupakan lokasi atau objek yang akan digunakan pada suatu penelitian. Penentuan lokasi merupakan suatu hal yang penting bagi peneliti dalam proses penelitian sebab akan memudahkan peneliti untuk melakukan sebuah penelitian. Lokasi pengambilan data pada penelitian ini yaitu pada *Google Maps* bagian Komentar pada lokasi bugis *waterpark adventure*.

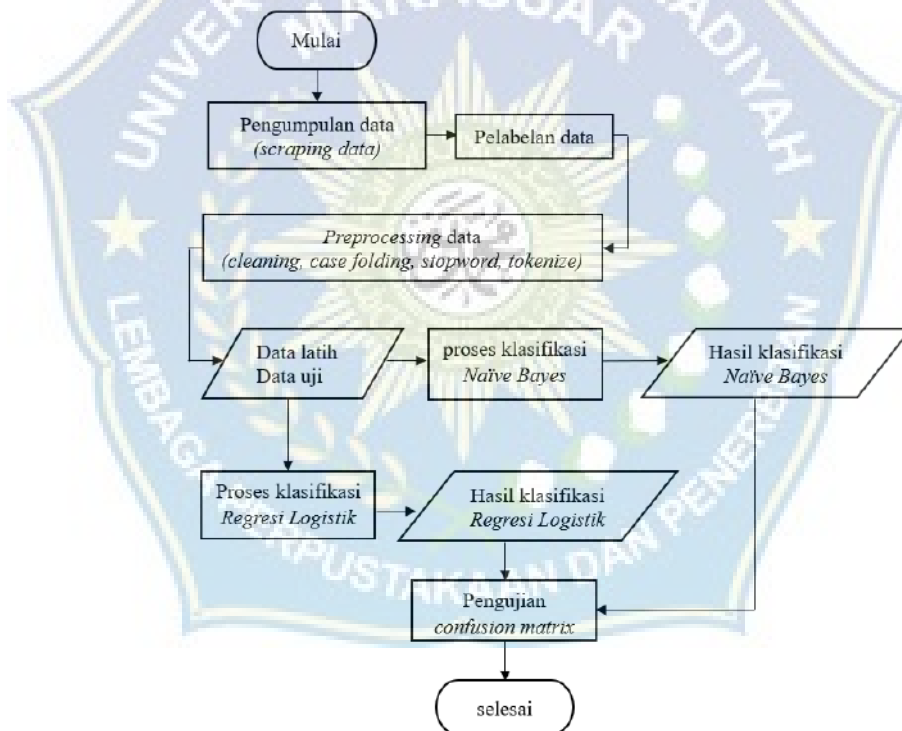
B. Alat dan Bahan

Adapun alat dan bahan yang akan digunakan dalam penelitian ini yaitu:

1. Kebutuhan *Hardware* (Perangkat Keras)
 - a. Laptop asus
 - b. Ram 8 GB
 - c. Sistem operasi Windows 11
2. Kebutuhan *Software* (Perangkat Lunak)
 - a. Data
 - b. *Python*
 - c. *Library Python*
 - d. *Spreadsheed / excel*
 - e. *Google colabulatory / Google colab*

C. Perancangan Sistem

Perancangan sistem merupakan *modelling* atau proses merancang dan membangun sebuah sistem yang bertujuan untuk memenuhi kebutuhan dan memecahkan sebuah masalah tertentu terutama pada penelitian ini. Perancangan sistem melibatkan pemilihan teknologi yang tepat, arsitektur sistem, desain antarmuka pengguna, pemilihan metode dan algoritma pemrograman, dan pengujian sistem secara menyeluruh. Untuk mempermudah dalam pembuatan dan pengembangan sebuah sistem peneliti merancang sebuah *Flowchart*, sehingga dapat dengan mudah memahami alur dari sebuah sistem yang dibangun oleh peneliti dan dibangun dengan terstruktur.



Gambar 5. *Flowchart system*

Proses dimulai dengan mengambil data ulasan dari *google maps*. Setelah data diterima, data akan dilabelkan dengan kategori yang diperlukan, setelah data dilabel data akan dilakukan pra pemrosesan data untuk membersihkan dan mempersiapkan data tersebut agar siap untuk di analisis. Pra pemrosesan biasanya melibatkan

penghapusan data yang tidak relevan atau yang tidak dibutuhkan, normalisasi teks dan pemisahan kata. Data kemudian di bagi jadi dua yaitu data latih dan data uji digunakan untuk menguji kinerja algoritma. Kemudian algoritma *naïve bayes* dan *regresi logistik* masing masing dilatih menggunakan data latih. Setelah pelatihan data maka akurasi algoritma diukur dengan menggunakan data uji. Selanjutnya hasil yang di keluarkan yaitu akurasi dari kedua algoritma untuk mengetahui mana yang memiliki performa lebih baik dalam analisis sentimen khususnya wisatawan terhadap bugis *waterpark adventure* di ulasan *google maps*. Selanjutnya, akan di uji dengan pengujian *confusion matrix*. Algoritma dengan akurasi tertinggi kemudian ditentukan sebagai algoritma paling efektif.

D. Teknik Pengujian Sistem

1. Pengujian *confusion matrix*

Confusion matrix yaitu sesuatu tata cara yang biasa digunakan buat menerapkan perhitungan tingkat akurasi pada informasi mining. *Confusion matrix* muat data tentang klasifikasi yang di prediksi dengan benar oleh suatu system klasifikasi. Ada 3 parameter yang bakal dihitung, ialah akurasi, recall, serta precision (Gunawan et al., 2018).

		Actual Values	
		Positive (1)	Negative (0)
Predictive Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Gambar 6. *Confusion Matrix*

Dalam analisis sentimen, *confusion matrix* digunakan buat menghitung akurasi, presisi, recala, serta f1-score dari model. Akurasi ialah proporsi informasi yang

diklasifikasikan dengan benar dari totalitas informasi. Rumus akurasi biasa dinyatakan dengan berikut :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \dots\dots\dots(11)$$

Presisi memperlihatkan persentase ketepatan hasil prediksi oleh prosedur yang digunakan. Rumus presisi bisa dipaparkan dengan berikut:

$$Precision = \frac{TP}{TP+FP} \times 100\% \dots\dots\dots(12)$$

Recall merupakan rasio informasi yang diklasifikasikan dengan benar selaku positif dari total informasi positif yang didapatkan. Recall mencerminkan keberhasilan model dalam menciptakan Kembali data. Rumus recall bisa diformulasikan sebagai berikut:

$$Reccal = \frac{TP}{TP+FN} \times 100\% \dots\dots\dots(13)$$

F1-score ialah gabungan antara presisi serta recall buat membagikan nilai yang lebih konferensif dalam mengevaluasi kinerja model. Rumusnya bisa dinyatakan dengan berikut:

$$F1 - Score = 2 * \frac{Precision*recall}{Precision+recall} \dots\dots\dots(14)$$

E. Teknik Analisis Data

Teknik analisis data adalah proses pengumpulan, pembersihan, transformasi, dan pengolahan data untuk mendapatkan informasi yang berguna dan bermakna. Tujuan utama dari teknik analisis data adalah untuk mengidentifikasi pola, tren, dan hubungan dalam data sehingga dapat digunakan untuk pengambilan keputusan dan memberikan wawasan yang lebih baik.

Teknik analisis data mencakup berbagai metode dan alat statistik, matematika, atau komputasi yang digunakan untuk memproses dan memanipulasi data, seperti penambangan data, pembelajaran mesin, statistik deskriptif, dan visualisasi data.

Proses analisis data pada penelitian ini memiliki 3 langkah yaitu:

1. Reduksi Data

Menurut Bapak Sugishirono, reduksi data berarti jumlah data yang diperoleh dari lapangan cukup besar, sehingga perlu dicatat secara cermat dan detail. Data yang telah direduksi memberikan gambaran yang lebih jelas dan memudahkan peneliti untuk mengumpulkan dan mengambil data selanjutnya sesuai kebutuhan (Sugiono, 2015).

Dengan mereduksi data, setiap peneliti diarahkan pada tujuannya yang ingin dicapai. Fokus utama penelitian kualitatif adalah hasil oleh karena itu, jika peneliti menemukan *outlier*, *unknown's* dan kualifikasi yang tidak berpola maka peneliti harus berhati-hati dalam mengecilkan data.

2. Display Data

Display data, disebut juga penyajian data, merupakan komponen kedua dari kegiatan analisis data. Penyajian data adalah kumpulan informasi terstruktur yang dijelaskan dalam bentuk narasi yang memungkinkan peneliti menarik kesimpulan penelitian dan mengambil tindakan (Miles dan Huberman, 1992:17). Penyajian data ini merupakan rangkaian yang disusun secara logis dan sistematis untuk membantu peneliti memahami apa yang sedang terjadi dan memungkinkan peneliti melakukan tindakan analitis berdasarkan pemahaman tersebut (Sutopo, 2002: 92).

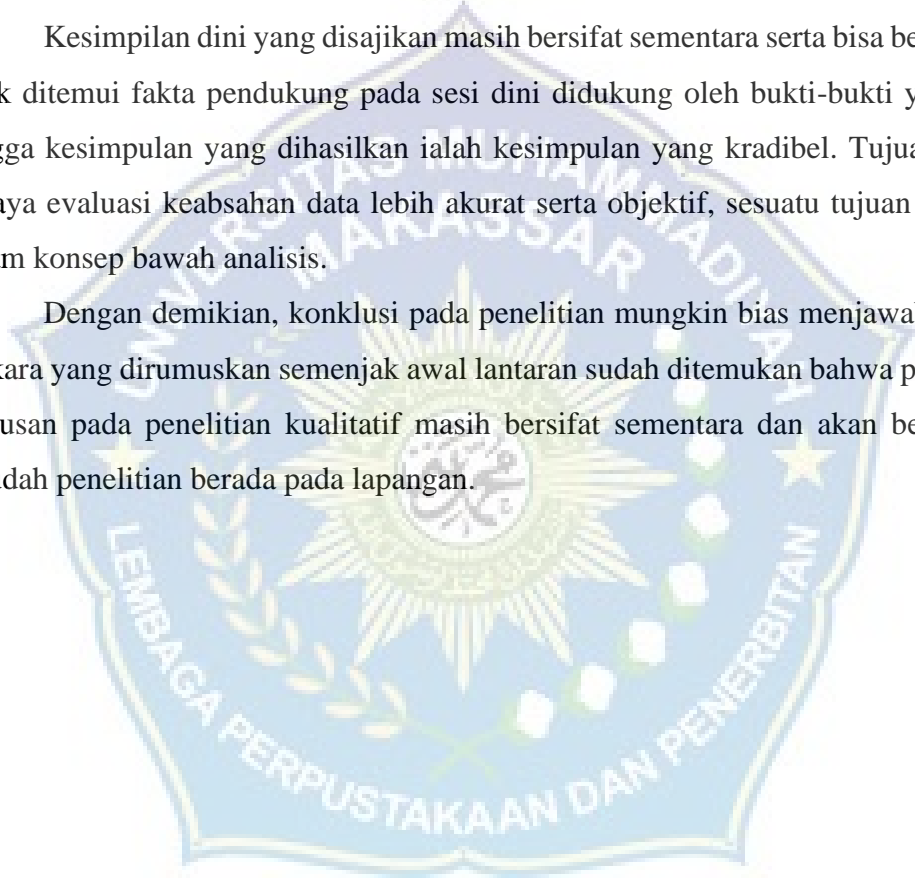
Hasil reduksi disajikan dengan cara tertentu per pola, kategori, fokus dan topik yang ingin dipahami untuk memahami masalah. Lihat data memungkinkan peneliti untuk melihat gambaran besar atau bagian eksklusif dari temuan penelitian mereka. Dalam penelitian kualitatif, penyajian data Sangguo berupa deskripsi singkat, diagram, interaksi antar kategori, flowchart, dan sebagainya. Teks deskriptif adalah jenis yang sering digunakan untuk menyajikan data dalam penelitian kualitatif.

3. Penarikan Kesimpulan

Penarikan kesimpulan ialah tahapan akhir dari metode analisis informasi kualitatif yang dicoba dengan memikirkan hasil reduksi informasi yang masih relevan dengan tujuan analisis yang mau dicapai. Sesi ini bertujuan buat mencari arti dari informasi yang dikumpulkan dengan mencari ikatan persamaan, ataupun perbandingan guna menarik kesimpulan selaku jawaban dari kasus yang terdapat.

Kesimpulan dini yang disajikan masih bersifat sementara serta bisa berganti bila tidak ditemui fakta pendukung pada sesi dini didukung oleh bukti-bukti yang valid, hingga kesimpulan yang dihasilkan ialah kesimpulan yang kradibel. Tujuan validasi supaya evaluasi keabsahan data lebih akurat serta objektif, sesuatu tujuan tercantum dalam konsep bawah analisis.

Dengan demikian, konklusi pada penelitian mungkin bias menjawab rumusan perkara yang dirumuskan semenjak awal lantaran sudah ditemukan bahwa perkara dan rumusan pada penelitian kualitatif masih bersifat sementara dan akan berkembang sesudah penelitian berada pada lapangan.



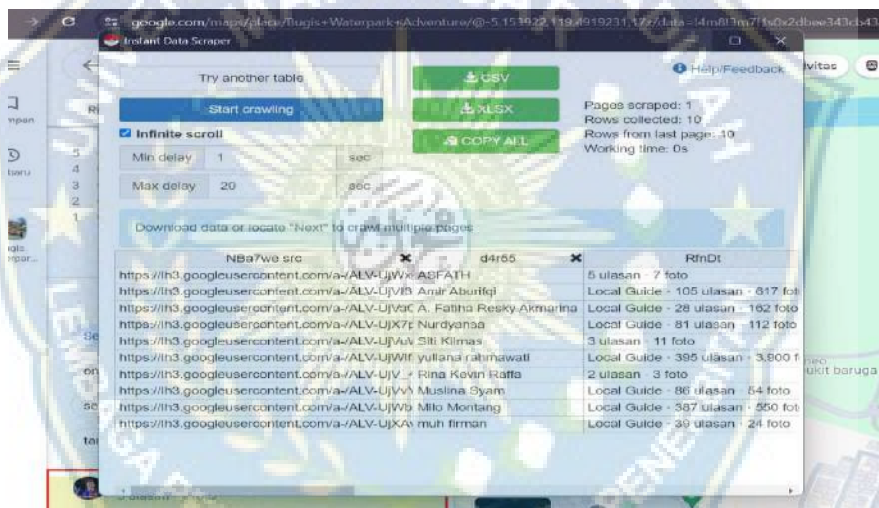
BAB IV

HASIL DAN PEMBAHASAN

A. Pengambilan Data

Pengumpulan data ulasan Bugis *Waterpark Adventure* di *Google Maps* menggunakan *instant data scraper*, proses pengumpulan data ulasan fokus pada pemilihan ulasan yang paling relevan di Bugis *Waterpark Adventure* dengan jumlah data yang di ambil 1000 data ulasan.

Berikut adalah tampilan dari *instant data scraper* dapat dilihat pada gambar dibawah:



Gambar 7. *Scraping* Data Ulasan Menggunakan *Instant Data Scraper*

1. Dataset ulasan

Setelah menyelesaikan proses *scraping* data ulasan dari *Google Maps* Bugis *Waterpark Adventure*, hasil *scraping* akan disimpan dalam format excel.

Berikut adalah hasil *scraping* data ulasan yang telah di simpan ke excel dapat dilihat pada gambar di bawah:



Gambar 8. Dataset Ulasan

B. Pelabelan Data

Proses pelabelan sentimen dilakukan secara manual dengan tujuan mengidentifikasi pola dan ciri-ciri dalam teks yang mencerminkan sentimen positif, negatif, atau netral. Atribut ulasan berisi komentar yang di temukan pada *Google Maps*, sementara atribut label memuat nilai klasifikasi, antara lain positif, negatif, dan netral.

Berikut pelabelan data dapat dilihat pada tabel dibawah ini :

Tabel 1. Tahap Pelabelan

Ulasan	Label
Tempat yang keren untuk sejenak melepas penat refreshing dan juga liburan bersama orang-orang jarak tempuh yang cukup dekat dari pusat kota makassar dan juga memiliki pemandangan yang indah dan asri	Positif
Katanya wahana terbeser di indonesia timur tapi kok wahananya cuman satu-satu yang dibuka dan jembatan kayunya sudah pada lapuk bahjan pada saat saya berkunjung ada anak yang jatuh akibat kayuh patah	Negatif
Tempatnya bagus tapi sayangnya promo2nya tidak konsisten sehingga kadang membingungkan pengunjung	Netral

Setelah melakukan pelabelan secara manual, terdapat sebanyak 726 data yang telah diklasifikasikan dengan jumlah yang seimbang antara label positif, negatif dan netral. Dalam penelitian ini, data dibagi menjadi data latih dan data uji menggunakan *library* dari *scikit-learn*, yakni *train test split*. Proporsi pembagian adalah 10% dari keseluruhan data untuk data uji, dan 90% untuk data latih. Sekitar 653 data latih.

C. Tahap *Preprocessing*

Tahap *preprocessing* adalah tahap awal dalam pengolahan data yang bertujuan untuk membersihkan, memformat, dan mempersiapkan data mentah sebelum dilakukan analisis lebih lanjut.

Berikut merupakan tahapan-tahapan dari proses *preprocessing* :

1. *Cleaning*

Pada tahap *cleaning* akan dilakukan proses untuk penghilangan tanda baca dan karakter yang tidak diperlukan seperti tanda titik, tanda koma, tanda tanya, tanda seru, dan, menghapus hastag dan mention, menghapus emoji, serta menghapus karakter yang tidak relevan.

Tabel 2. Tahap *Cleaning*

Sebelum	Sesudah
Alhamdulillah untuk lepas kejenuhan dan bersantai dengan keluarga mantaplah bugis waterpark is the bestTerutama marketingx yg bernama zaenal mantaplah sangat baik dan ramah 🙏🙏🙏 ...	Alhamdulillah untuk lepas kejenuhan dan bersantai dengan keluarga mantaplah bugis waterpark is the best Terutama marketingx yg bernama zaenal mantaplah sangat baik dan ramah
Tempat wisata yg sangat menarik, jalan bersama keluarga sangat recommended karena bisa puas bermain air mulai dari anak2 sampai dewasa, sangat banyak permainan.. 😊😊 ...	Tempat wisata yg sangat menarik jalan bersama keluarga sangat recommended karena bisa puas bermain air mulai dari anak sampai dewasa sangat banyak permainan

Tempat yang bagus buat main air 😊 bulan oktober juga ada diskon, cuma 66ribu tiap hari termasuk sabtu minggu dalam rangka hut kalla group... yang mau diantar, ayo ajakin aku 🙏 ...	Tempat yang bagus buat main air bulan oktober juga ada diskon cuma ribu tiap hari termasuk sabtu minggu dalam rangka hut kalla group yang mau diantar ayo ajakin aku
---	--

2. Transform Cases

Transform Cases, adalah proses mengubah huruf kapital menjadi huruf kecil dalam teks.

Tabel 3. Tahap *Transform Cases*

Sebelum	Sesudah
Alhamdulillah untuk lepas kejenuhan dan bersantai dengan keluarga mantaplah bugis waterpark is the best Terutama marketingx yg bernama zaenal mantaplah sangat baik dan ramah	alhamdulillah untuk lepas kejenuhan dan bersantai dengan keluarga mantaplah bugis waterpark is the best terutama marketingx yg bernama zaenal mantaplah sangat baik dan ramah
Tempat wisata yg sangat menarik jalan bersama keluarga sangat recomended karena bisa puas bermain air mulai dari anak sampai dewasa sangat banyak permainan	tempat wisata yg sangat menarik jalan bersama keluarga sangat recomended karena bisa puas bermain air mulai dari anak sampai dewasa sangat banyak permainan
Tempat yang bagus buat main air bulan oktober juga ada diskon cuma ribu tiap hari termasuk sabtu minggu dalam rangka hut kalla group yang mau diantar ayo ajakin aku	tempat yang bagus buat main air bulan oktober juga ada diskon cuma ribu tiap hari termasuk sabtu minggu dalam rangka hut kalla group yang mau diantar ayo ajakin aku

3. *stopword*

proses menghapus kata-kata yang umum dan sering muncul dalam teks, seperti kata depan, kata penghubung, dan kata-kata lainnya yang tidak memberikan kontribusi signifikan dalam pemahaman makna teks. Kata-kata tersebut umumnya adalah kata-kata penghubung seperti, "dan", "di", "dari", dan kata-kata lain yang tidak terlalu penting dalam analisis.

Tabel 4. Tahap Stopword

Sebelum	Sesudah
alhamdulillah untuk lepas kejenuhan dan bersantai dengan keluarga mantaplah bugis waterpark is the best terutama marketingx yg bernama zaenal mantaplah sangat baik dan ramah	alhamdulillah lepas kejenuhan bersantai keluarga mantaplah bugis waterpark is the best marketingx yg bernama zaenal mantaplah ramah
tempat wisata yg sangat menarik jalan bersama keluarga sangat recommended karena bisa puas bermain air mulai dari anak sampai dewasa sangat banyak permainan	wisata yg menarik jalan keluarga recommended puas bermain air anak dewasa permainan
tempat yang bagus buat main air bulan oktober juga ada diskon cuma ribu tiap hari termasuk sabtu minggu dalam rangka hut kalla group yang mau diantar ayo ajakin aku	bagus main air oktober diskon ribu sabtu minggu rangka hut kalla group diantar ayo ajakin

4. *Tokenizing*

Tokenizing adalah proses memecah teks menjadi unit-unit yang lebih kecil yang disebut "token". Token ini dapat berupa kata, frasa, atau tanda baca. Pada langkah ini, kalimat dipisahkan menjadi potongan-potongan sebelum dianalisis lebih lanjut.

Tabel 5. Tahap *Tokenizing*

Sebelum	Sesudah
alhamdulillah lepas kejenuhan bersantai keluarga mantaplah bugis waterpark is the best marketingx yg bernama zaenal mantaplah ramah	['alhamdulillah', 'lepas', 'kejenuhan', 'bersantai', 'keluarga', 'mantaplah', 'bugis', 'waterpark', 'is', 'the', 'best', 'marketingx', 'yg', 'bernama', 'zaenal', 'mantaplah', 'ramah']
wisata yg menarik jalan keluarga recomended puas bermain air anak dewasa permainan	['wisata', 'yg', 'menarik', 'jalan', 'keluarga', 'recomended', 'puas', 'bermain', 'air', 'anak', 'dewasa', 'permainan']
bagus main air oktober diskon ribu sabtu minggu rangka hut kalla groub diantar ayo ajakin	['bagus', 'main', 'air', 'oktober', 'diskon', 'ribu', 'sabtu', 'minggu', 'rangka', 'hut', 'kalla', 'group', 'diantar', 'ayo', 'ajakin']

D. Klasifikasi *Naïve Bayes*

Klasifikasi merupakan teknik komputasi untuk mengelompokkan data berdasarkan keterikatan data. Pengelompokkan data dimana data tersebut mempunyai kelas label atau target.

Dalam penelitian ini menggunakan metode naive bayes untuk menimport machine learning naive bayes menggunakan library sklearn. Metode ini hanya memerlukan data yang kecil untuk menghasilkan akurasi. Cara kerja naive bayes yaitu dengan menghitung jumlah label, menghitung jumlah kasus perkelas, kalikan semua variabel kelas dan bandingkan hasil perkelas. Metode ini untuk menghitung berapa nilai accuracy, precision, recall, f1_score.

1. Pemanggilan model *Naïve Bayes*

Pada gambar dibawah adalah pemanggilan dari model *Naïve Bayes* menggunakan *MultinomialNB()* library yang disediakan *scikit-learn*, selanjutnya melatih model menggunakan metode “fit” agar model belajar mengidentifikasi pola

atau hubungan antara vector fitur dan label di data latih Dimana program menggunakan model yang sudah dilatih untuk membuat prediksi pada data uji dan memanggil dari hasil prediksi. Selanjutnya membuat hasil prediksi dari metode yang sebelumnya sudah di ajarkan pada model dari data latih dan data uji.

```
clf = MultinomialNB()
clf.fit(x_train_vec, y_train)

[52] y_pred = clf.predict(x_test_vec)

[53] y_pred
array(['Netral', 'Netral', 'Positif', 'Positif', 'Netral', 'Netral',
       'Positif', 'Positif', 'Negatif', 'Negatif', 'Netral', 'Netral',
       'Netral', 'Positif', 'Negatif', 'Netral', 'Negatif', 'Netral',
       'Positif', 'Negatif', 'Positif', 'Positif', 'Negatif', 'Positif',
       'Positif', 'Positif', 'Negatif', 'Negatif', 'Netral', 'Positif',
       'Positif', 'Positif', 'Positif', 'Positif', 'Netral', 'Netral',
       'Negatif', 'Positif', 'Netral', 'Netral', 'Negatif', 'Positif',
       'Negatif', 'Negatif', 'Positif', 'Negatif', 'Netral', 'Negatif',
       'Negatif', 'Netral', 'Positif', 'Positif', 'Positif', 'Netral',
       'Positif', 'Netral', 'Negatif', 'Positif', 'Negatif', 'Negatif',
       'Netral', 'Negatif', 'Positif', 'Negatif', 'Positif', 'Netral',
       'Positif', 'Positif', 'Netral', 'Positif', 'Negatif', 'Positif',
       'Negatif'], dtype='<U7')
```

Gambar 9. Pemanggilan Model *Naive Bayes*

2. Hasil evaluasi *Naive Bayes*

Pada gambar dibawah merupakan hasil evaluasi *naive bayes* pada data uji dengan akursi yang didapatkan yaitu sebesar 77%, dengan presisi sebesar 75%, recall sebesar 75%, dan f1-score sebesar 75%. Dimana pada akurasi (*accuracy*) yaitu mengukur proporsi prediksi yang benar dari total prediksi, presisi (*precision*) yaitu mengukur proporsi hasil positif yang di prediksi, untuk recall yaitu mengukur proporsi hasil positif yang benar dari total jumlah hasil positif yang seharusnya diprediksi, dan terakhir yaitu, f1-score yaitu ukuran gabungan dari presisi dan recall dan memberikan penambahan menyeluruh tentang kinaerja dari model *Naive Bayes*.

```

# Hitung metrik evaluasi
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average='macro')
recall = recall_score(y_test, y_pred, average='macro')
f1 = f1_score(y_test, y_pred, average='macro')
classification_rep_test = classification_report(y_test, y_pred)

# Tampilkan hasil evaluasi
print(f'Accuracy: {accuracy:.2f}')
print(f'Precision: {precision:.2f}')
print(f'Recall: {recall:.2f}')
print(f'F1-Score: {f1:.2f}')
print("\nClassification Report:")
print(classification_rep_test)

```

```

Accuracy: 0.77
Precision: 0.75
Recall: 0.75
F1-Score: 0.75

Classification Report:

```

	precision	recall	f1-score	support
Negatif	0.67	0.70	0.68	20
Netral	0.71	0.62	0.67	24
Positif	0.87	0.93	0.90	29
accuracy			0.77	73
macro avg	0.75	0.75	0.75	73
weighted avg	0.76	0.77	0.76	73

Gambar 10. Hasil pengujian *Naive Bayes*

3. Hasil Prediksi *Naive Bayes* dari data uji

Pada gambar berikut merupakan hasil prediksi dari 73 data uji yang diberikan sebelumnya, terlihat untuk hasil keluaran pada index ke 2 dengan teks “wahana permainan air bergantian menyala” dengan label asli yaitu “Negatif” dan hasil prediksi label yang dikeluarkan “Positif”.

```

Index: 0
Text: wahana bermain air cocok keluarga sayang harganya mahal
True Label: Netral
Predicted Label: Netral
-----
Index: 1
Text: parkir kendaraan berdebu musim kemarau pastinya becek musim hujan
True Label: Negatif
Predicted Label: Netral
-----
Index: 2
Text: wahana permainan air bergantian menyala
True Label: Negatif
Predicted Label: Positif
-----
Index: 3
Text: wahananya pilihannya cocok liburan keluarga suasanannya adem anak anak main orang tua hehe wahana menantang
True Label: Positif
Predicted Label: Positif
-----
Index: 4
Text: menyenangkan cocok liburan keluarga tpi makanan masuk dalam dan harga makanan mahal
True Label: Negatif
Predicted Label: Netral
-----
Index: 5
Text: keren tempatnya mahal bgt
True Label: Netral
Predicted Label: Netral
-----
Index: 6
Text: tempatnya luass indah
True Label: Positif
Predicted Label: Positif

```



```
Index: 66
Text: bagus tempatnya alami
True Label: Positif
Predicted Label: Positif
-----
Index: 67
Text: rekreasi wahana air seru fasilitas lazy boat saluran air mengalir tenang dimana bersantai dibawa air pelampung peluncur air
True Label: Positif
Predicted Label: Positif
-----
Index: 68
Text: wisata air menari terletak area perumahan buki baruga wahana permainan anak anak orang dewasa tempatnya sejuk tiket masuk terjangkau kekurangan
True Label: Positif
Predicted Label: Netral
-----
Index: 69
Text: pekan ramai pengunjung wahana yg disediakan pengunjung memilih wahana yg cocok liburan keluarga
True Label: Positif
Predicted Label: Positif
-----
Index: 70
Text: sesuai promosi marketingnya menawarkan paket diskon t kanipegawainya ramah sulit informasi yg menyenangkan
True Label: Negatif
Predicted Label: Negatif
-----
Index: 71
Text: menyenangkan penjaganya ramah biayanya terjangkau bersih cantik memanjakan mata
True Label: Positif
Predicted Label: Positif
-----
Index: 72
Text: nanya harga masuknya kesana
True Label: Netral
Predicted Label: Negatif
```

Gambar 11. Hasil Prediksi *Naive Bayes* Dari Data Uji

E. Klasifikasi *Regresi Logistik*

Peneliti akan mengevaluasi akurasi algoritma *Regresi Logistik* dalam melakukan analisis sentimen terhadap ulasan *Bugis Waterpark Adventure*. Dengan data yang sama yang diberikan kepada model *Naive Bayes*, Dimana dilakukan perbandingan dengan cara memberikan data yang sama, memanggil model dari *library Regresi logistik* yang telah disediakan sebelumnya dan membuat tingkat akurasi. Maka hasilnya sebagai berikut.

1. Pemanggilan model *Regresi Logistik*

Pada gambar dibawah yaitu pemanggilan model dari *Regresi Logistik* menggunakan *LogisticRegression()* library yang disediakan *scikit-learn*, selanjutnya malatih model menggunakan model “fit” agar model belajar untuk mengidentifikasi pola atau hubungan antara vector fitur dan label di data latih Dimana program menggunakan model yang sudah dilatih untuk membuat prediksi pada data uji dan

memanggil dari hasil prediksi, selanjutnya membuat hasil prediksi dari metode yang sebelumnya sudah di ajarkan pada model dari data latih dan data uji.

```
logistic_reg = LogisticRegression()  
logistic_reg.fit(x_train_vec, y_train)  
  
[48] y_pred = logistic_reg.predict(x_test_vec)  
  
[49] y_pred  
array(['Positif', 'Netral', 'Positif', 'Positif', 'Positif', 'Negatif',  
      'Positif', 'Positif', 'Netral', 'Negatif', 'Netral', 'Netral',  
      'Netral', 'Positif', 'Negatif', 'Netral', 'Netral', 'Netral',  
      'Positif', 'Negatif', 'Positif', 'Positif', 'Netral', 'Positif',  
      'Positif', 'Netral', 'Negatif', 'Negatif', 'Netral', 'Positif',  
      'Positif', 'Positif', 'Positif', 'Positif', 'Negatif', 'Netral',  
      'Negatif', 'Positif', 'Netral', 'Netral', 'Negatif', 'Positif',  
      'Positif', 'Negatif', 'Positif', 'Negatif', 'Netral', 'Negatif',  
      'Negatif', 'Netral', 'Positif', 'Negatif', 'Positif', 'Netral',  
      'Netral', 'Netral', 'Negatif', 'Positif', 'Negatif', 'Negatif',  
      'Netral', 'Negatif', 'Positif', 'Negatif', 'Positif', 'Netral',  
      'Positif', 'Positif', 'Positif', 'Positif', 'Negatif', 'Positif',  
      'Negatif'], dtype=object)
```

Gambar 12. Pemanggilan Model *Regresi Logistik*

2. Hasil evaluasi *Regresi Logistik*

Pada gambar dibawah merupakan hasil evaluasi *regresi logistik* pada data uji dengan akursi yang didapatkan yaitu sebesar, 82% dengan presisi sebesar 81%, recall sebesar 81%, dan f1-score sebesar 81%. Dimana pada akurasi (*accuracy*) yaitu mengukur proposisi prediksi yang benar dari total prediksi, presisi (*precision*) yaitu mengukur preposisi hasil positif yang diprediksi, untuk recall yaitu mengukur proposisi hasil positif yang benar dari total jumlah hasil positif yang seharusnya di prediksi, dan terakhir yaitu f1-score yaitu ukuran gabungan dari presisi dan recall dan memberikan penambahan menyeluruh tentang kinerja dari model *Regresi Logistik*.

```

# Hitung metrik evaluasi
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average='macro')
recall = recall_score(y_test, y_pred, average='macro')
f1 = f1_score(y_test, y_pred, average='macro')
classification_rep_test = classification_report(y_test, y_pred)

# Tampilkan hasil evaluasi
print(f'Accuracy: {accuracy:.2f}')
print(f'Precision: {precision:.2f}')
print(f'Recall: {recall:.2f}')
print(f'F1-Score: {f1:.2f}')
print("\nClassification Report:")
print(classification_rep_test)

```

```

Accuracy: 0.82
Precision: 0.81
Recall: 0.81
F1-Score: 0.81

Classification Report:

```

	precision	recall	f1-score	support
Negatif	0.71	0.75	0.73	20
Netral	0.81	0.71	0.76	24
Positif	0.90	0.97	0.93	29
accuracy			0.82	73
macro avg	0.81	0.81	0.81	73
weighted avg	0.82	0.82	0.82	73

Gambar 13. Hasil pengujian *Regresi Logistik*

3. Hasil prediksi *Regresi Logistik* dari data uji

Pada gambar berikut merupakan hasil prediksi dari 73 data uji yang diberikan sebelumnya, terlihat untuk hasil keluaran pada index ke 0 dengan teks “wahana bermain air cocok keluarga sayang harganya mahal” dengan label benar yaitu “Netral” dan hasil prediksi label yang dikeluarkan “Positif”

```

Index: 0
Text: wahana bermain air cocok keluarga sayang harganya mahal
True Label: Netral
Predicted Label: Positif
-----
Index: 1
Text: parkir kendaraan berdebu dimusin kemarau pastinya becek dimusin hujan
True Label: Negatif
Predicted Label: Netral
-----
Index: 2
Text: wahana permainan air bergantian menyala
True Label: Negatif
Predicted Label: Positif
-----
Index: 3
Text: wahananya pilihanpokoknya cocok liburan keluarga suasanannya adem anak anak main orang tua hehe wahana menantang
True Label: Positif
Predicted Label: Positif
-----
Index: 4
Text: menyenangkan cocok liburan keluarga tpi makanan masuk dalamdan harga makanan mahal
True Label: Negatif
Predicted Label: Positif
-----
Index: 5
Text: keren tempatnya mahal bgt
True Label: Netral
Predicted Label: Negatif
-----
Index: 6
Text: tempatnya luass indah
True Label: Positif
Predicted Label: Positif

```

```
Index: 66
Text: bagus tempatnya alami
True Label: Positif
Predicted Label: Positif
-----
Index: 67
Text: rekreasi wahana air seru fasilitas lazy boat saluran air mengalir tenang dimana bersantai dibawa air pelampung peluncur air
True Label: Positif
Predicted Label: Positif
-----
Index: 68
Text: wisata air memari terletak area perumahan buki baruga wahana permainan anak anak orang dewasa tempatnya sejuk tiket masuk terjangkau kekurangan
True Label: Positif
Predicted Label: Positif
-----
Index: 69
Text: pekan ramai pengunjung wahana yg disediakan pengunjung memilih wahana yg cocok liburan keluarga
True Label: Positif
Predicted Label: Positif
-----
Index: 70
Text: sesuai promosi marketingnya menawarkan paket diskon t kaitpegawainya ramah sulit informasi yg menyenangkan
True Label: Negatif
Predicted Label: Negatif
-----
Index: 71
Text: menyenangkan penjaganya ramah biayanya terjangkau bersih cantik memanjakan mata
True Label: Positif
Predicted Label: Positif
-----
Index: 72
Text: nanya harga masuknya kesana
True Label: Netral
Predicted Label: Negatif
```

Gambar 14. Hasil Prediksi *Regresi Logistik* dari data uji

F. Hasil klasifikasi *Naïve Bayes* dan *Regresi Logistik*

Setelah menganalisis hasil dari kedua algoritma yaitu *Naïve Bayes* dan *Regresi Logistik*, selanjutnya akan melakukan perbandingan untuk mengetahui algoritma mana yang lebih efektif dalam melakukan analisis sentimen terhadap ulasan *Bugis Waterpark Adventure* pada *Google Maps*.

Berdasarkan hasil klasifikasi antara *Naïve bayes* dan *Regresi logistik* pada sub bab sebelumnya. Dengan jumlah data sebanyak 726, data tersebut kemudian dibagi menjadi dua yaitu data latih dan data uji. Dengan skenario pembagian data 90:10, Dimana 90% data latih dan 10% data uji. Jadi data ulasan bugis waterpark adventure sebanyak 653 data latih dan 73 data uji. Berdasarkan hasil data ulasan bugis waterpark adventure dengan menggunakan naive bayes dan regresi logistik maka didapatkan nilai sebagai berikut :

Hasil pengujian data ulasan dengan menggunakan *Naive Bayes* mendapatkan nilai *accuracy* sebesar 77%.

Tabel 6. Hasil pengujian *Naive Bayes* pada data uji

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
<i>Positif</i>	0,87%	0,93%	0,90%	29
<i>Negatif</i>	0,67%	0,70%	0,68%	20
<i>Netral</i>	0,71%	0,62%	0,67%	24
<i>Marco</i>	0,75%	0,75%	0,75%	73
<i>Average</i>				
<i>Weighted</i>	0,76%	0,77%	0,76%	73
<i>Average</i>				

Berikutnya hasil pengujian data ulasan dengan menggunakan *Regresi Logistik* mendapatkan nilai *accuracy* sebesar 82%.

Tabel 7. Hasil pengujian *Regresi Logistik* pada data uji

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
<i>Positif</i>	0,90%	0,97%	0,93%	29
<i>Negatif</i>	0,71%	0,75%	0,73%	20
<i>Netral</i>	0,81%	0,71%	0,76%	24
<i>Marco</i>	0,81%	0,81%	0,81%	73
<i>Average</i>				
<i>Weighted</i>	0,82%	0,82%	0,82%	73
<i>Average</i>				

Jika dilihat pada table 6 dan 7, menunjukkan bahwa analisis sentimen data ulasan *bugis waterpark adventure* dengan metode *Naive Bayes* dan *Regresi logistic*, metode *Regresi Logistik* lebih mempunyai tingkat kinerja yang cukup bagus, karena memiliki nilai *accuracy* diatas 50%. Dengan true positif rate sebesar 90%. *Macro average* adalah salah satu metode atau cara yang digunakan untuk menghitung rata-rata dari metrik evaluasi pada setiap kelas. *Weighted average* adalah metode

penghitungan rata-rata yang memberikan bobot berbeda pada setiap nilai yang akan dihitung berdasarkan faktor tertentu.

G. Pengujian *confusion matrix*

Pada tahap pengujian menggunakan *confusion matrix*, dimana peneliti akan membandingkan hasil klasifikasi yang dilakukan secara manual dengan hasil klasifikasi yang telah dihasilkan oleh model metode *naive bayes dan regresi logistik*.

Dalam perhitungan secara manual untuk evaluasi hasil prediksi algoritma, terdapat dua algoritma yang dievaluasi, yaitu *Naïve Bayes* dan *Regresi Logistik*, dengan jumlah sebanyak 73 data.

Terdapat 4 istilah sebagai representasi hasil proses klasifikasi pada *confusion matrix*. Keempat istilah tersebut adalah True Positive (TP), True Negative (TN), False Positive (FP) dan False Negative (FN). Berikut adalah Hasil Klasifikasi Dengan Labeling Manual Dan *Confusion Matrix*.

Tabel 8. Hasil klasifikasi dengan labelling manual dan confusion matrix

Ulasan	Label	Label Prediksi	Hasil <i>Confusion Matrix</i>
wisata air menyenangkan	Positif	Positif	<i>True Positif</i>
kecewa ngga sesuai expektasi	Negatif	Negatif	<i>True Negatif</i>
berlibur keluarga teman gathering kantor	Netral	Positif	<i>False Positif</i>
pemandangannya indah tdk kotor	Positif	Negatif	<i>False Negatif</i>

Tabel 9. Hasil Klasifikasi *Naïve Bayes* Yang *True Positif*

No	Ulasan	<i>True Label</i>	<i>Prediksi Naïve Bayes</i>
----	--------	-------------------	-----------------------------

1	wahananya pilihanpokoknya cocok liburan keluarga suasanannya adem anak anak main orang tua hehe wahana menantang	Positif	Positif
2	tempatnya luass indah	Positif	Positif
3	adik suka banget	Positif	Positif
4	wahana water boom mengasikan tersedia restoran toko swalayan beristirahat keluarga yg asri	Positif	Positif
5	harga masuk sesuai fasilitas disediakan tempatnya bersih pilihan wahana bermain	Positif	Positif
6	wisata air menyenangkan	Positif	Positif
7	nya bagus ukuran wisata makanannya murah meriah	Positif	Positif
8	permandian air bagus air jernih pengunjung	Positif	Positif
9	sejuk	Positif	Positif
10	wahana air luas dn lengkap fasilitas x bersih dn makanan nya terjangkau recomended liburan keluarga	Positif	Positif
11	menyenangkan luas	Positif	Positif
12	serubuat liburan keluarga wahana	Positif	Positif
13	arena bermain yg aman anak anak pengawasan yg orang tua khawatir	Positif	Positif
14	sy bgtu sangt puas dengn permainan wahananya yg bagus	Positif	Positif
15	bagus rekomendasi liburan keluarga	Positif	Positif
16	bagus bersih tempatnya air kolamnya jg bersih	Positif	Positif
17	gokil lengkap luas ditambah wahananya	Positif	Positif
18	refreshing warga kota makassar wahana air	Positif	Positif
19	pokoky top petugasy ramah humoris	Positif	Positif
20	destinasi wisata ya sih bagus dn indah	Positif	Positif
21	bagus luas	Positif	Positif
22	rekreasi keluarga recomended ya tempatnya	Positif	Positif
23	wahana air lengkap pulau sulawesi	Positif	Positif

24	bagus tempatnya alami	Positif	Positif
25	rekreasi wahana air seru fasilitas lazy boat saluran air mengalir tenang dimana bersantai dibawa air pelampung peluncur air	Positif	Positif
26	pekan ramai pengunjung wahana yg disediakan pengunjung memlih wahana yg cocok liburan keluarga	Positif	Positif
27	menyenangkan penjaganya ramah biayanya terjangkau bersih cantik memanjakan mata	Positif	Positif
28	wahana bermain air cocok keluarga sayang harganya mahal	Netral	Netral
29	keren tempatnya mahal bgt	Netral	Netral
30	lumayan gasebo nya klu hujan tdk berteduh apalg klu pengunjungnya ramai	Netral	Netral
31	biaya masuk mahal lumayan bagus sih	Netral	Netral
32	smoga tmn tmn berkunjung kesini	Netral	Netral
33	kartu data murah	Netral	Netral
34	cuman sih kolam ombaknya sore kebuka gitu jd yg pergi pagi pulang siang ga bgitu	Netral	Netral
35	wahananya dibatasin jam operasionalnya wahana utama yg daya tarik phinisi slide black hole free fall beroperasi jam	Netral	Netral
36	tempatnya bagus cuman minim petunjuk arah lokasi sistem jam operasi wahananya yg bagus menunggu jam wahana itupun durasi beroperasinya singkat	Netral	Netral
37	recomended liburan bareng kluarga cuman sygnya gk dibolehin bawa makanan hhe	Netral	Netral
38	musim berkunjung	Netral	Netral
39	suasana asri alami cuman perawatan sarana maksimal	Netral	Netral
40	makanan dilarang bawa krn tersedia resto	Netral	Netral
41	promo masuk ramai teman	Netral	Netral
42	ndak lgi deh aja	Netral	Netral

True Positif = 42

Pada tabel 9. Adalah hasil *confusion matrix* data ulasan dengan membandingkan hasil klasifikasi *Naïve Bayes* dengan label manual maka didapatkan nilai *True positif* sebanyak 42 data ulasan.

Tabel 10. Hasil Klasifikasi *Naïve Bayes* Yang *True Negatif*

No	Ulasan	<i>True Label</i>	Prediksi <i>Naïve Bayes</i>
1	trmpatnya terawat kotor	Negatif	Negatif
2	sayangnya akses jalanya rusak sulit nyaman	Negatif	Negatif
3	rebutan benen karna	Negatif	Negatif
4	wahana tutup	Negatif	Negatif
5	lumayan bagus wahanya tdk aktif disayangkan	Negatif	Negatif
6	jln jelek pa lg klau hujan genangan air	Negatif	Negatif
7	kecewa deengan bugis waterpark nga profesional	Negatif	Negatif
8	seru fasilitasnya tdk sesuai	Negatif	Negatif
9	tiket masuk mahal wahana tdk dibuka	Negatif	Negatif
10	tiketnya melonjak banget anak hitung tiket dewasa ngecewain	Negatif	Negatif
11	fasilitas yg rusak	Negatif	Negatif
12	kotor	Negatif	Negatif
13	sesuau promosi marketingnya menawarkan paket diskon t kamipegawainya ramah sulit informasi yg menyenangkan	Negatif	Negatif
14	penjaganya aktif contoh ban milik pelanggan tau dipake tidur tiduran disembunyikan parahnya petugas menegur nggak nyaman	Negatif	Negatif

True Negatif = 14

Pada tabel 10. Adalah hasil *confusion matrix* data ulasan dengan membandingkan hasil klasifikasi *Naïve Bayes* dengan label manual maka didapatkan nilai *True Negatif* sebanyak 14 data ulasan.

Tabel 11. Hasil Klasifikasi *Naïve Bayes* Yang *False Positif*

No	Ulasan	<i>True Label</i>	Prediksi <i>Naïve Bayes</i>
1	wahana permainan air bergantian menyala	Negatif	Positif
2	mengecewakan inidatang ajak keluarga permainan ngak yg aktifban bocor semuamasuk mahal ngak nikmati permainan kecuali berenang	Negatif	Positif
3	parkiran kendaraan berdebu dimusim kemarau pastinya becek dimusim hujan	Negatif	Netral
4	menyenangkan cocok liburan keluarga tpi makanan masuk dalamdan harga makanan mahal	Negatif	Netral
5	harga tiketnya aja mahal rajin hunting voucer biar murah	Negatif	Netral
6	bagus wahananya dibatasi operasionalnya harganya ndk turun	Negatif	Netral
7	assalamualaikum warahmatullahi wabarakatu salam lestari pencinta alam penikmat alam jaga alam nikmati anak cucu	Netral	Positif
8	keamanan wajib pakaian renang turun kolam wahana	Netral	Positif

False Positif = 8

Pada tabel 11. Adalah hasil *confusion matrix* data ulasan dengan membandingkan hasil klasifikasi *Naïve Bayes* dengan label manual maka didapatkan nilai *false positif* sebanyak 8 data ulasan.

Tabel 12. Hasil Klasifikasi *Naïve Bayes* Yang *False Negatif*

No	Ulasan	<i>True Label</i>	Prediksi <i>Naïve Bayes</i>
1	satusatunya permandian oke makassar	Positif	Netral
2	wisata air menari terletak area perumahan buki baruga wahana permainan anak anak orang dewasa tempatnya sejuk tiket masuk terjangkau kekurangan menu kulinernya mahal	Positif	Netral

	pengunjung dilarang membawa makanan minuman		
3	elayanan petugas wahana diperhatikan	Netral	Negatif
4	fasilitas ban nya memadai ban dipake wahananya	Netral	Negatif
5	wahana waktujam bukanya mainnya maksimal krn main full tdk main	Netral	Negatif
6	aturan	Netral	Negatif
7	wahananya banyaknamun wahananya dibatasi keterbatasan perlengkapan wahana karyawan life guardnya mungkim	Netral	Negatif
8	tutup sarana vaksin gratis	Netral	Negatif
9	nanya harga masuknya kesana	Netral	Negatif

False Negatif = 9

Pada tabel 12. Adalah hasil *confusion matrix* data ulasan dengan membandingkan hasil klasifikasi Naïve Bayes dengan label manual maka didapatkan nilai *false negatif* sebanyak 9 data ulasan.

Tabel 13. Hasil Klasifikasi *Regresi Logistik* Yang *True Positif*

No	Ulasan	True Label	Prediksi <i>Regresi Logistik</i>
1	wahananya pilihanpokoknya cocok liburan keluarga suasanannya adem anak anak main orang tua hehe wahana menantang	Positif	Positif
2	tempatnya luass indah	Positif	Positif
3	adik suka banget	Positif	Positif
4	wahana water boom mengasikan tersedia restoran toko swalayan beristirahat keluarga yg asri	Positif	Positif
5	harga masuk sesuai fasilitas disediakan tempatnya bersih pilihan wahana bermain	Positif	Positif
6	wisata air menyenangkan	Positif	Positif

7	nya bagus ukuran wisata makanannya murah meriah	Positif	Positif
8	permandian air bagus air jernih pengunjung	Positif	Positif
9	sejuk	Positif	Positif
10	wahana air luas dn lengkap fasilitas x bersih dn makanan nya terjangkau recomended liburan keluarga	Positif	Positif
11	menyenangkan luas	Positif	Positif
12	serubuat liburan keluarga wahana	Positif	Positif
13	arena bermain yg aman anak anak pengawasan yg orang tua khawatir	Positif	Positif
14	sy bgtu sangt puas dengn permainan wahananya yg bagus	Positif	Positif
15	bagus rekomendasi liburan keluarga	Positif	Positif
16	bagus bersih tempatnya air kolamnya jg bersih	Positif	Positif
17	gokil lengkap luas ditambah wahananya	Positif	Positif
18	refreshing warga kota makassar wahana air	Positif	Positif
19	pokoky top petugasy ramah humoris	Positif	Positif
20	destinasi wisata ya sih bagus dn indah	Positif	Positif
21	bagus luas	Positif	Positif
22	rekreasi keluarga rekomended ya tempatnya	Positif	Positif
23	wahana air lengkap pulau sulawesi	Positif	Positif
24	bagus tempatnya alami	Positif	Positif
25	rekreasi wahana air seru fasilitas lazy boat saluran air mengalir tenang dimana bersantai dibawa air pelampung peluncur air	Positif	Positif
26	wisata air menari terletak area perumahan buki baruga wahana permainan anak anak orang dewasa tempatnya sejuk tiket masuk terjangkau kekurangan menu kulinernya mahal pengunjung dilarang membawa makanan minuman	Positif	Positif

27	pekan ramai pengunjung wahana yg disediakan pengunjung memlih wahana yg cocok liburan keluarga	Positif	Positif
28	menyenangkan penjaganya ramah biayanya terjangkau bersih cantik memanjakan mata	Positif	Positif
29	elayanan petugas wahana diperhatikan	Netral	Netral
30	lumayan gasebo nya klu hujan tdk berteduh apalg klu pengunjungnya ramai	Netral	Netral
31	biaya masuk mahal lumayan bagus sih	Netral	Netral
32	smoga tmn tmn berkunjung kesini	Netral	Netral
33	kartu data murah	Netral	Netral
34	aturan	Netral	Netral
35	assalamualaikum warahmatullahi wabarakatu salam lestari pencinta alam penikmat alam jaga alam nikmati anak cucu	Netral	Netral
36	cuman sih kolam ombaknya sore kebuka gitu jd yg pergi pagi pulang siang ga begitu	Netral	Netral
37	wahananya dibatasin jam operasionalnya wahana utama yg daya tarik phinisi slide black hole free fall beroperasi jam	Netral	Netral
38	tempatnya bagus cuman minim petunjuk arah lokasi sistem jam operasi wahananya yg bagus menunggu jam wahana itupun durasi beroperasinya singkat	Netral	Netral
39	recomended liburan bareng keluarga cuman sygnya gk dibolehin bawa makanan hhe	Netral	Netral
40	musim berkunjung	Netral	Netral
41	suasana asri alami cuman perawatan sarana maksimal	Netral	Netral
42	keamanan wajib pakaian renang turun kolam wahana	Netral	Netral
43	makanan dilarang bawa krn tersedia resto	Netral	Netral
44	promo masuk ramai teman	Netral	Netral
45	ndak lgi deh aja	Netral	Netral

True positif = 45

Pada tabel 13. Adalah hasil *confusion matrix* data ulasan dengan membandingkan hasil klasifikasi *Regresi Logistik* dengan label manual maka didapatkan nilai *True Positif* sebanyak 45 data ulasan.

Tabel 14. Hasil Klasifikasi Regresi Logistik Yang *True Negatif*

No	Ulasan	<i>True Label</i>	Prediksi <i>Regresi Logistik</i>
1	trmpatnya terawat kotor	Negatif	Negatif
2	sayangnya akses jalanya rusak sulit nyaman	Negatif	Negatif
3	wahana tutup	Negatif	Negatif
4	lumayan bagus wahanya tdk aktif disayangkan	Negatif	Negatif
5	harga tiketnya aja mahal rajin hunting voucher biar murah	Negatif	Negatif
6	jln jelek pa lg klau hujan genangan air	Negatif	Negatif
7	kecewa deengan bugis waterpark nga profesional	Negatif	Negatif
8	seru fasilitasnya tdk sesuai	Negatif	Negatif
9	tiket masuk mahal wahana tdk dibuka	Negatif	Negatif
10	mengecewakan inidatang ajak keluarga permainan ngak yg aktifban bocor semuamasuk mahal ngak nikmati permainan kecuali berenang	Negatif	Negatif
11	tiketnya melonjak banget anak hitung tiket dewasa ngecewain	Negatif	Negatif
12	fasilitas yg rusak	Negatif	Negatif
13	penjaganya aktif contoh ban milik pelanggan tau dipake tidur tiduran disembunyikan parahnya petugas menegur nggak nyaman	Negatif	Negatif
14	kotor	Negatif	Negatif
15	sesuau promosi marketingnya menawarkan paket diskon t kamipegawainya ramah sulit informasi yg menyenangkan	Negatif	Negatif

True Negatif : 15

Pada tabel 14. Adalah hasil *confusion matrix* data ulasan dengan membandingkan hasil klasifikasi *Regresi Logistik* dengan label manual maka didapatkan nilai *True Negatif* sebanyak 15 data ulasan.

Tabel 15. Hasil Klasifikasi *Regresi Logistik* Yang *False Positif*

No	Ulasan	<i>True Label</i>	Prediksi <i>Regresi Logistik</i>
1	wahana permainan air bergantian menyala	Negatif	Positif
2	menyenangkan cocok liburan keluarga tpi makanan masuk dalam dan harga makanan mahal	Negatif	Positif
3	parkiran kendaraan berdebu dimusim kemarau pastinya becek dimusim hujan	Negatif	Netral
4	rebutan benen karna	Negatif	Netral
5	bagus wahananya dibatasi operasionalnya harganya ndk turun	Negatif	Netral
6	wahana bermain air cocok keluarga sayang harganya mahal	Netral	Positif

False Positif = 6

Pada tabel 15. Adalah hasil *confusion matrix* data ulasan dengan membandingkan hasil klasifikasi *Regresi Logistik* dengan label manual maka didapatkan nilai *false positif* sebanyak 6 data ulasan.

Tabel 16. Hasil Klasifikasi *Regresi Logistik* Yang *False Negatif*

No	Ulasan	<i>True Label</i>	Prediksi <i>Regresi Logistik</i>
1	satusatunya permandian oke makassar	Positif	Netral
2	keren tempatnya mahal bgt	Netral	Negatif
3	wahananya banyak namun wahananya dibatasi keterbatasan perlengkapan wahana karyawan life guardnya mungkim	Netral	Negatif

4	tutup sarana vaksin gratis	Netral	Negatif
5	fasilitas ban nya memadai ban dipake wahananya	Netral	Negatif
6	wahana waktujam bukanya mainnya maksimal krn main full tdk main	Netral	Negatif
7	nanya harga masuknya kesana	Netral	Negatif

False Negatif = 7

Pada tabel 16. Adalah hasil *confusion matrix* data ulasan dengan membandingkan hasil klasifikasi *Regresi Logistik* dengan label manual maka didapatkan nilai *false positif* sebanyak 7 data ulasan.

Gambar berikut adalah *Confusion matrix* pada hasil prediksi data uji *naive bayes* dan *regresi logistik*:

Confusion Matrix Naive Bayes:

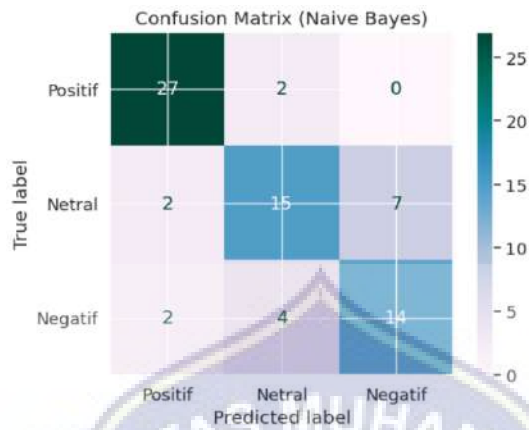
```
[[14  4  2]
 [ 7 15  2]
 [ 0  2 27]]
```

Gambar 15. *Confusion matrix* pada data uji *Naive Bayes*

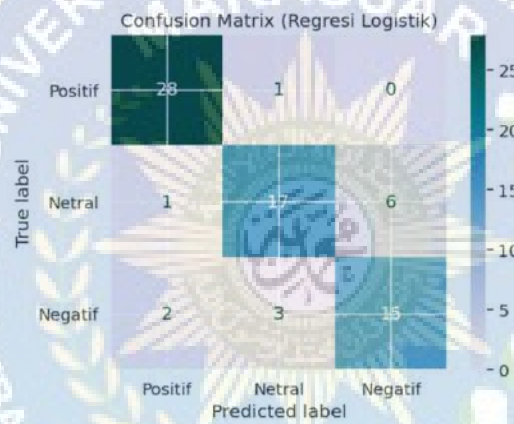
Confusion Matrix Regresi Logistik:

```
[[15  3  2]
 [ 6 17  1]
 [ 0  1 28]]
```

Gambar 16. *Confusion matrix* pada data uji *Regresi Logistik*



Gambar 17. *Confusion matrix Naïve Bayes*



Gambar 18. *Confusion matrix Regresi Logistik*

		PREDIKSI LABEL		
		POSITIF	NETRAL	NEGATIF
TRUE LABEL	POSITIF	(TP) 27	(FN) 2	(FN) 0
	NETRAL	(FP) 2	(TP) 15	(FN) 7
	NEGATIF	(FP) 2	(FP) 4	(TN) 14

Gambar 19. Hasil *Confusion matrix Naïve Bayes*

		PREDIKSI LABEL		
		POSITIF	NETRAL	NEGATIF
TRUE LABEL	POSITIF	(TP) 28	(FN) 1	(FN) 0
	NETRAL	(FP) 1	(TP) 17	(FN) 6
	NEGATIF	(FP) 2	(FP) 3	(TN) 15

Gambar 20. Hasil *Confusion matrix* Regresi Logistik

Keterangan *confusion matrix*:

Tabel 17. Keterangan *Confusion matrix* Naïve Bayes dan Regresi Logistik

<i>Naïve Bayes</i>		<i>Regresi Logistik</i>	
(TP) True Positif	: 42	(TP) True Positif	: 45
(TN) True Negatif	: 14	(TN) True Negatif	: 15
(FP) False Positif	: 8	(FP) False Positif	: 6
(FN) False Negatif	: 9	(FN) False Negatif	: 7

Confusion matrix digunakan untuk menghitung *accuracy*, *presisi*, *recall*, dan *f1-score* dari *naïve bayes* dan *regresi logistik*. Berikut adalah hasil *accuracy* dari *confusion matrix* pada data uji *naïve bayes*:

1. *Naïve Bayes*

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% = \frac{42+14}{42+14+8+9} \times 100\% = 77\%$$

Kemampuan model dalam mengklasifikasikan data prediksi dari *naïve bayes* dengan benar adalah 77%.

$$precision = \frac{TP}{TP+FP} \times 100\% = \frac{42}{42+8} = \frac{42}{50} = 0,84 \times 100\% = 84\%$$

Kemampuan model dalam mengklasifikasikan data prediksi dari *naïve bayes* dengan benar untuk ulasan yang positif adalah 84%.

$$recall = \frac{TP}{TP+FN} \times 100\% = \frac{42}{42+9} = \frac{42}{51} = 0,82 \times 100\% = 82\%$$

Kemampuan model dalam mengklasifikasikan data prediksi dari *naïve bayes* dengan benar untuk ulasan yang positif adalah 82%.

$$F1 - score = 2 * \frac{precision * recall}{precision + recall} = 2 * \frac{84 * 82}{84 + 82} = \frac{13.776}{166} = 82\%$$

Hasil dari *f1-score* adalah 82%

2. Regresi Logistik

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% = \frac{45+15}{45+15+6+7} \times 100\% = 82\%$$

Kemampuan model dalam mengklasifikasikan data prediksi dari *regresi logistik* dengan benar adalah 82%.

$$precision = \frac{TP}{TP+FP} \times 100\% = \frac{45}{45+6} = \frac{45}{51} = 0,77 \times 100\% = 88\%$$

Kemampuan model dalam mengklasifikasikan data prediksi dari *regresi logistik* dengan benar untuk ulasan yang positif adalah 88%.

$$recall = \frac{TP}{TP+FN} \times 100\% = \frac{45}{45+7} = \frac{45}{52} = 0,86 \times 100\% = 86\%$$

Kemampuan model dalam mengklasifikasikan data prediksi dari *regresi logistik* dengan benar untuk ulasan yang positif adalah 86%.

$$F1 - score = 2 * \frac{precision * recall}{precision + recall} = 2 * \frac{88 * 86}{88 + 86} = \frac{15.136}{174} = 86\%$$

Hasil dari *f1-score* adalah 86%

Accuracy adalah proporsi data yang diklasifikasikan dengan benar dari keseluruhan data. *Presisi* adalah seberapa tepat model dalam membandingkan data secara keseluruhan dengan jumlah Positif yang di klasifikasi secara benar. *Recall* adalah sistem melakukan perbandingan jumlah data benar Positif dengan jumlah data yang sebenarnya Positif secara keseluruhan. *F-Score* didapatkan dari penggabungan hasil *presisi* dan *recall* dengan menggunakan rata-rata harmonic *presisi* dan *recall*.

H. Hasil Perhitungan *Confusion Matrix*

Setelah diperoleh hasil klasifikasi dengan metode *naïve bayes* dan *regresi logistik*, ditentukan metode yang paling baik dalam melakukan analisis sentimen terhadap ulasan *bugis waterpark adventure*. Dari pengujian *confusion matrix* yang telah diperoleh, elemen-elemen yang diambil untuk menghitung empat ukuran ketepatan yang digunakan dalam melakukan perbandingan, yaitu *accuracy*, *presisi*, *recall*, dan *f1-score*. Hasil perbandingan dari kedua metode tersebut dapat dilihat pada tabel berikut.

Tabel 18. Perbandingan *Naïve Bayes* dan *Regresi Logistik*

Ukuran	<i>Naïve Bayes</i>	<i>Regresi Logistik</i>
Akurasi	77%	82%
Presisi	84%	88%
Recall	82%	86%
F1-score	82%	86%

pada tabel 18, bahwa tingkat akurasi *Regresi Logistik* dan *naïve bayes*, metode *regresi logistik* lebih menunjukkan kinerja yang lebih baik secara keseluruhan. Dapat dilihat *Regresi Logistik* memiliki tingkat akurasi yang lebih tinggi yaitu (82%) dibandingkan dengan *Naïve Bayes* (77%), ini menunjukkan bahwa *Regresi Logistik* lebih baik dalam memprediksi sentimen pada dataset uji Bugis Waterpark Adventure pada *Google Maps*.

BAB V

PENUTUP

Berdasarkan evaluasi kinerja dua model klasifikasi, *Regresi Logistik* dan *Naïve Bayes*, dalam penelitian membandingkan akurasi algoritma *Naïve Bayes* dan *Regresi Logistik* dalam analisis sentimen wisatawan terhadap Bugis *Waterpark Adventure* di ulasan *Google Maps*, kita dapat menyimpulkan:

A. Kesimpulan :

Berdasarkan evaluasi kinerja dua model klasifikasi, *Naïve Bayes* dan *Regresi Logistik* dalam tugas analisis sentimen terhadap ulasan wisatawan terhadap Bugis *waterpark Adventure*, kita dapat menyimpulkan :

1. *Regresi Logistik* menunjukkan performa yang lebih baik secara keseluruhan dibandingkan dengan *Naïve Bayes*, dengan nilai akurasi, precision, recall, dan f1-score yang lebih tinggi.
2. Performa *Regresi Logistik* lebih stabil dan konsisten, memberikan keunggulan dalam klasifikasi sentimen pada dataset ini.
3. Model *Regresi Logistik* memiliki tingkat akurasi yang memuaskan, mencapai 82%, sedangkan metode *Naïve Bayes* memiliki tingkat akurasi 77%. Menandakan bahwa metode *Regresi Logistik* memiliki kemampuan untuk melakukan prediksi dengan tepat.

B. Saran

Jika memungkinkan, mempertimbangkan untuk meningkatkan ukuran dataset dengan menambahkan lebih banyak sampel data. Ini dapat membantu meningkatkan keakuratan dan generalisasi model.

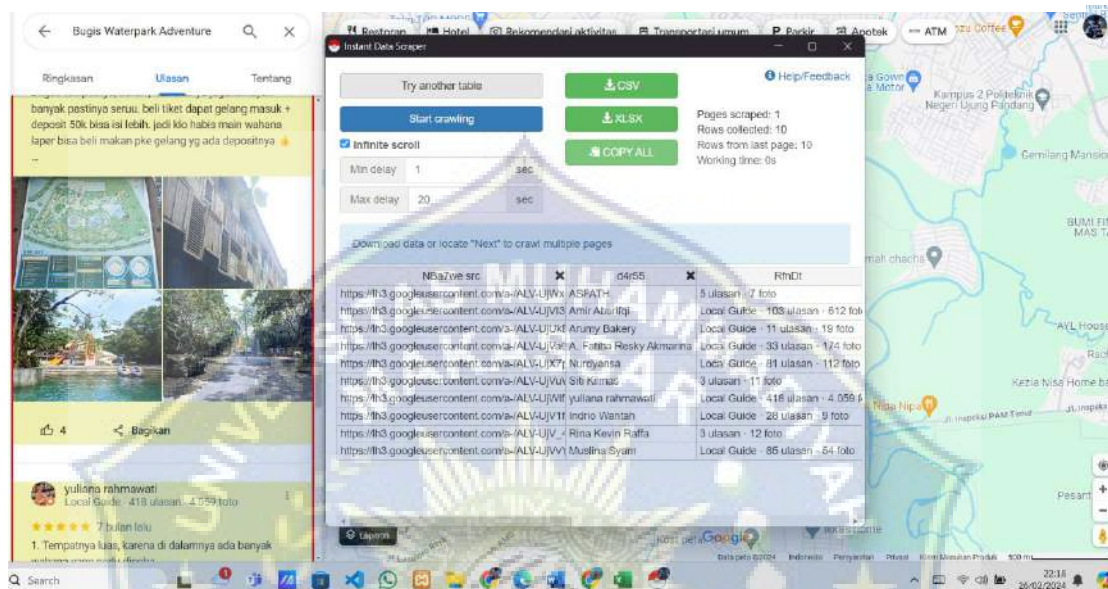
DAFTAR PUSTAKA

- Ainun, N. (2022). *Bugis Waterpark Adventure Makassar*. Detiksulsel.
- Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.
- DARUSSALAM, A. G. (2021). *Analisis Sentimen Menggunakan Text Mining Dengan Metode Naïve Bayes Dan Regresi Logistik*. 4(1), 1–23.
- David W. Hosmer, Jr., Stanley Lemeshow, R. X. S. (2013). *Sturdivant. Applied logistic regression* (3rd, berilus ed.). John Wiley & Sons. <https://doi.org/0470582472>
- Falahah & Nur, D. D. A. (2015). Pengembangan Aplikasi Sentiment Analysis Menggunakan Metode Naïve Bayes. *Seminar Nasional Sistem Informasi Indonesia*.
- Fazrin, F., Nurul Prastiwi, O., & Andeswari, R. (2022). Perbandingan Algoritma K-Nearest Neighbor dan Logistic Regression pada Analisis Sentimen terhadap Vaksinasi Covid-19 pada Media Sosial Twitter. *EProceedings of Engineering*, 10(2), 1596–1604.
- Feldman, Ronen, and J. S. (2007). *The text mining handbook: advanced approaches in analyzing unstructured data*. Cambridge university press.
- Gunawan, B., Pratiwi, H. S., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 4(2), 113. <https://doi.org/10.26418/jp.v4i2.27526>
- Han, J., Kamber, M., & Pei, J. (2012). Third Edition : Data Mining Concepts and Techniques. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.
- Harjanta, A. T. J. (2015). Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining. *Jurnal Informatika Upgris*, 1–9.
- Harrington, P. (2012). *Machine learning in action*. Simon and Schuster.
- Khofifah, W., Rahayu, D. N., & Yusuf, A. M. (2022). Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps. *Jurnal Interkom: Jurnal*

- Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 16(4), 28–38.
<https://doi.org/10.35969/interkom.v16i4.192>
- Kurniawan, T. (2017). Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Media Mainstream Menggunakan Naïve Bayes Classifier Dan Support Vector Machine. *Institut Teknologi Sepuluh Nopember*.
- L. Ardiani, H. Sujaini, and T. T. (2020). *Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak*. 8, 183.
<https://doi.org/10.26418/justin.v8i2.36776>.
- Liu. (2012). *Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies*. 1–167.
- Lorosae, T. A., & Prakoso, B. D. (2018). *ANALISIS SENTIMEN BERDASARKAN OPINI MASYARAKAT PADA*. 25–30.
- Marsland, S. (2015). *Machine learning: an algorithmic perspective*. CRC press.
- Riadi Silitonga, Y., Munawar, & Noor Hapsari, I. (2019). Analisis Dan Penerapan Datamining Untuk Mendeteksi Berita Palsu (Fake News) Pada Social Media Dengan Memanfaatkan Modul Scikit Learn. *Undergraduate Theses of Information Systems*.
- Sang-Bum Kim, Kyoung-Soo Han, Hae-Chang Rim, S. H. M. (2006). *Some Effective Techniques for Naive Bayes Text Classification*. 18.
<https://doi.org/10.1109/TKDE.2006.180>
- Sentimen, A., Wisata, O., Di, B., Maps, G., Siti Utami, D., Utami, D. S., & Erfina, A. (2022). Analisis Sentimen Objek Wisata Bali Di Google Maps Menggunakan Algoritma Naive Bayes. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 6(1), 418–427.
- Siang, J. J. (2005). *Jaringan Syaraf Tiruan dan Pemrogramannya Menggunakan MATLAB*.

LAMPIRAN

Lampiran 1. Proses pengumpulan data



Lampiran 2. Data ulasan setelah dikumpulkan



Lampiran 3. Pelabelan data ulasan positif, netral, negatif

ULASAN	LABEL
Tempat rekreasi bersama keluarga baik juga dijadikan tempat outing untuk acara kantor	Positif
Sangat menyenangkan berlibur disini	Positif
Keren klaw kesini harus coba semua wahana Pokokx MANTAP deh	Positif
Bangga dan senang banget di makassar sudah ada wahana bermain untuk keuarga kunjungi tidak perlu lagi jauh2 ke jakarta	Positif
yang mau rekreasi keluarga yang rekomended ya disini tempatnya	Positif
Kami Losari Taylor crew benar benar bersenang senang Terimakasih Bugis Waterpark	Positif
Uhhuyyyy... Tempat ini saya cocok untuk liburan keluarga ataupun liburan musim panas cuk. Selain di sani banyak wahana, pengunjungnya pun sangat ramah tamah. Di tempat ini kali	Positif
Keren dengan semua wahana air nya.. Ada kolam ombaknya	Positif
tempat rekreasi yang sangat bagus dan lumayan dekat	Positif
Mantapp. Bagus. Sangat Recommended.	Positif
Tempatnya seru, cocok untuk family gathering dan tempat rekreasi. Tempatnya ketat karena dimana mana ada penjaga tapi itu membuktikan kualitas suatu tempat	Positif
Karena di situ banyak wahana air yang asik di naikkan mulai dari indoor sampai outdoor. Intinya nggak nyesel deh dan satu lagi di situ terdapat ombak buatan yang keren.	Positif
Harga masuk sesuai dengan fasilitas yang disediakan. Secara umum tempatnya bersih, banyak pilihan wahana bermain.	Positif
Salah satu tempat wisata di makassar yang sangat cocok buat liburan keluarga. Wahananya banyak dan ada kolam ombaknya. Pokoknya tempat ini sangat bagus buat kalian yang ingin	Positif
Wahananya banyak dan juga seru, airnya jernih, tempatnya jg besar. Menyenangkan banget deh klu masuk disana. Pelayanannya jg baik. The best deh	Positif
Tempat yg bagus untuk refreshing.. sejuk, banyak pohon, banyak wahana air nys, seluncuran. Ada Mushollah, Kantin, Toko baju. Harga makanan relatif (harga restoran). Ada promo til	Positif
Tempatnya bagus dan bersih, banyak seluncuran yang memacu adrenalin jadi sebelum naik seluncuran tersebut usahakan kumpulmi dlu nyawata.	Positif
Tempat wisata yg sangat menarik, jalan bersama keluarga sangat recomended karena bisa puas bermain air mulai dari anak2 sampai dewasa, sangat banyak permainan..	Positif
Wahana water boom yang mengasikan	Positif
lengkap kolam renangnya, kolam ombaknya mantab, airminum wakaf buat saya sangat berkesan, harga sesuai dengan fasilitas.	Positif
Tempatnya sangat luass dan indah.	Positif
Tempatnya luas dan megah cocok untuk liburan keluarga bersama anak	Positif
ULASAN	LABEL
Tempat wahana air yang bagus namun fasilitas yang ada di tempat ini kurang terawat mulai dari jembatan kayu yang mulai lapuk dan lain sebagainya	Netral
Bagi keluarga yg ingin berlibur akhir tahun kunjungi bugis waterpark	Netral
Salah satu destinasi wisata hiburan yg terletak di kota makassar di kawasan bukit baru a ntang	Netral
Kalau bisa wahana dibuka semua dan ban karet yg sudah rusak di ganti dengan yg baru, serta kalau bisa ban karet di perbanyak.	Netral
Kantintnya gak terlalu recomended sih	Netral
Rajin-rajin cek promo jika ingin kesini lebih murah	Netral
Ada promo tiket masuk 50%	Netral
Kolam Renang Ini Berada Di Perbatasan Makassar - Maros	Netral
Selalu ingin datang kebugis water Park Makassar	Netral
anak-anak puas main soal harga jangan terlalu mahal supaya semua kalangan bisa menikmati dan lebih ramai pastinya	Netral
utk rombongan lebih hemat pakai voucher	Netral
di manax suri it perumahan bukit baru ga masuk kah	Netral
Promo selama bulan oktober 2019 hrg tiket 67rb	Netral
Menarik, cuman beberapa wahana dibatasi operasionalnya.	Netral
Water parknya keren tadi sayang agak mahal biayanya...	Netral
Mau liburan ke bugis water park saja. Asyik loh...	Netral
Buka di jam 8 donk	Netral
Bagus,cuma apabila ada info tdk di infokan	Netral
BagusNamun kurang banyak wahana	Netral
Bagus tapi harganya perlu di tinjau ulang	Netral
Belum ke sana baru liat fotox mantap	Netral
wahananya banyak tapi tidak terbuka secara bersamaan (jam tertentu).	Netral
Beberapa wahana waktu beroperasinya tidak serentak. Saya tidak tahu alasan ...	Netral
Tempatnya luass wahananya ada jam jamnya pastikan datang pagi yaa wkwkwk	Netral
Goooooodddd,cuma harus sharian ido mau main di semua wahana krn ada ketentuan jam untuk tiap wahananya	Netral
ULASAN	LABEL
kepuasan pelanggan tdk tercapai, tdk sesuai dengan promo..	Negatif
Wahana tertentu ada batasan waktunya , jd kita ga puas nikmati	Negatif
Tiket terlalu mahal, airnya kotor, wahana sedikit tidak sesuai dengan harga tiket	Negatif
Wahana permainan air bergantian, tidak bisa semuanya menyala. .	Negatif
Harga tiket masuk sangat mahal di hari libur dan gak boleh bawa makanan..	Negatif
Wahananya tidak full yang bisa dimainkan, ada waktu2 tertentu wahana baru beroperasi	Negatif
Wahana tertentu hanya beroperasi di jam tertentu tidak bisa digunakan sepenuhnya	Negatif
Hiburan dari penatnya kota Makassar. Mulai rindang, tp toiletnya baru	Negatif
Sebagian wahana dalam renovasi dan harga tiket yg sangat mahal.	Negatif
Tiketnya melonjak banget dari biasanya, dan anak" pun di hitung tiket dewasa. Ngecewain	Negatif
KARNA HARGA TIKETNYA MAHAL SEKALI, DI DALAM WAHANA BUGIS WATER PART SUDAH BAGUS DAN MENYENANGKAN	Negatif
Wahananya ngge kayak dulu, terutama wahana kolam ombaknya.	Negatif
Wahana di buka pakai waktu, air masih hijau	Negatif
Keren tp kayak kurg terawat pohon2x... jd kesank kayak dihutan...	Negatif
tiket masuk mahal, tp sesuaiilah dengan fasilitasnya	Negatif
Banyak gazebo yg ditutup jadi susah cari tempat istirahat	Negatif
Tidak ada toilet di luar	Negatif
Kurang bagus . Tidak kayak dulu . Wahananya tidak terbuka semua	Negatif
Airnya agak asin dan kurang bersih dan permainannya main bergantian	Negatif
Wahananya kurang memuaskan...pengunjung bahkan kurang Bebas...pnyanannya kurAng baik	Negatif
2 Sendal hilang... perlu diperbaiki keamanan barang ...	Negatif
Jln jelek pa lg klaw hujan banyak genangan air.	Negatif
Satpam staff dan karyawan nya gak becus dalam menjaga.	Negatif
Aduh waktu sy dtg wahananya omn 2 yg buka	Negatif
Harganya sudah naik...	Negatif

Lampiran 4. Hasil Data Ulasan Setelah Dilakukan Proses Cleaning

Cleaned_Text

Bagus buat rekreasi keluarga
Tempat yang bagus buat main air bulan oktober juga ada diskon cuma ribu tiap hari termasuk sabtu minggu dalam rangka hut kalla group yang mau diantar ayo ajakin aku
Harga masuk sesuai dengan fasilitas yang disediakan Secara umum tempatnya bersih banyak pilihan wahana bermain
Wahana air terbaik dikota ini
Sangat seru ada ombak tapi kolam dewasa sedang dalam perbaikan yang penting sangat seru wahana banyakmengagumkan dan menakutkan di larang membawa orang yang sakit apapun itu
kolamnya sangat terjaga kebersihannya salut sama pengelolanya
tempat yang menyenangkan penjaganya ramah biayanya terjangkau bersih cantik memanjakan mata
tempat nya bagus untuk rekreasi tempay hiburan untuk keluarga kita bisa mandi dan berenang ada permainan
selalu bagus dan menjadi rekomendasi untuk liburan keluarga
pemandangan yang sangat indah dengan banyaknya wahana dan lingkungan yang luas sekali

Lampiran 5. Hasil Data Ulasan Setelah Dilakukan Proses Case Folding

alhamdulillah senang banget suka dgn suasana alamnya banyak pepohonan paling suka naik ban karet sambil rebahan nikmati aliran air yg mlaju mengelilingi wahana semoga kedepan lebih baik sukses buat semuanya
mantap anak sampai usia tua senang sekali
sehari mah gak cukup kalo ke tempat ini luas beud tapi dijamin memuaskan harga no problem but happy
wahananya asik untuk berenang dan bermain air yang menjadi favorit adalah wahana
klo mau datang pagi pagi pasti asyik banyak wahana kereennn

Lampiran 6. Hasil Data Ulasan Setelah Dilakukan Proses Stopword

Stopword_Removed_Text

alhamdulillah senang banget suka dgn suasana alamnya pepohonan suka ban karet rebahan nikmati aliran air yg mlaju mengelilingi wahana semoga kedepan sukses
mantap anak usia tua senang
sehari mah gak kalo luas beud dijamin memuaskan harga no problem but happy
wahananya asik berenang bermain air favorit wahana
klo pagi pagi asyik wahana kereennn

Lampiran 8. Hasil Data Ulasan Setelah Dilakukan Proses Tokenizing

Stopword_Removed_Text	Tokens
<p>alhamdulillah senang banget suka dgn suasana alamnya pepohonan suka ban karet rebahan nikmat aliran air yg mlaju mengelilingi wahana semoga kedepan sukses</p>	<p>alhamdulillah, senang, banget, suka, dgn, suasana, alamnya, pepohonan, suka, ban, karet, rebahan, nikmat, aliran, air, yg, mlaju, mengelilingi, wahana, semoga, kedepan, sukses</p>
<p>mantap anak usia tua senang</p>	<p>mantap, anak, usia, tua, senang</p>

Lampiran 7. Hasil Data Ulasan Setelah Tahap Preprocessing

Sebelum	Sesudah
<p>Cukup mahal sih tiketnya .gzebonya juga, sy kesana pas peek season, gazebonya 200rb 🎡🎡🎡 dit4 lain kan gtu2 amat harganya, gazebonya malah ada yg gratis</p>	<p>mahal sih tiketnyagzebonya sy kesana pas peek season gazebonya rb dit gtu harganya gazebonya yg gratis</p>

Lampiran 9. Proses klasifikasi Naive Bayes

```
[12] from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y,test_size=0.1,random_state=42,)
```

```
[13] from sklearn.naive_bayes import MultinomialNB
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.metrics import accuracy_score
from sklearn.feature_extraction.text import TfidfVectorizer
```

```
[14] vectorizer = TfidfVectorizer()
x_train_vec = vectorizer.fit_transform(x_train)
x_test_vec = vectorizer.transform(x_test)
```

```
print(x_test)
```

```
326 wahana bermain air cocok keluarga sayang harga...
519 parkirana kendaraan berdebu dimusia kemarau pas...
580 wahana permainan air bergantian menyala
33 wahananya pilihanpokoknya cocok liburan keluar...
570 menyenangkan cocok liburan keluarga tpi makana...
...
174 wisata air menari terletak area perumahan buki...
23 pekan ramai pengunjung wahana yg disediakan pe...
554 sesuuu promosi marketingnya menawarkan paket d...
196 menyenangkan penjaganya ramah biayanya terjangkau...
405 nanya harga masuknya kesana
Name: ULASAN, Length: 73, dtype: object
```

```
[16] non_string_values = df[df['ULASAN'].apply(lambda x: not isinstance(x, str))]['ULASAN']
print("Non-string values in 'TEXT' column:")
print(non_string_values)
```

```
Non-string values in 'TEXT' column:
Series([], Name: ULASAN, dtype: object)
```

```
clf = MultinomialNB()
clf.fit(x_train_vec, y_train)
```

```
[18] y_pred = clf.predict(x_test_vec)
```

```
y_pred
```

```
array(['Netral', 'Netral', 'Positif', 'Positif', 'Netral', 'Netral',
       'Positif', 'Positif', 'Negatif', 'Negatif', 'Netral', 'Netral',
       'Netral', 'Positif', 'Negatif', 'Netral', 'Negatif', 'Netral',
       'Positif', 'Negatif', 'Positif', 'Positif', 'Negatif', 'Positif',
       'Positif', 'Positif', 'Negatif', 'Negatif', 'Netral', 'Positif',
       'Positif', 'Positif', 'Positif', 'Positif', 'Netral', 'Netral',
       'Negatif', 'Positif', 'Netral', 'Netral', 'Negatif', 'Positif',
       'Positif', 'Negatif', 'Positif', 'Negatif', 'Netral', 'Negatif',
       'Negatif', 'Netral', 'Positif', 'Positif', 'Positif', 'Netral',
       'Positif', 'Netral', 'Negatif', 'Positif', 'Negatif', 'Negatif',
       'Netral', 'Negatif', 'Positif', 'Negatif', 'Positif', 'Netral',
       'Positif', 'Positif', 'Netral', 'Positif', 'Negatif', 'Positif',
       'Negatif'], dtype=object)
```

Lampiran 11. Hasil Pengujian Naive Bayes Pada Data Uji

```

▶ # Hitung metrik evaluasi
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average='macro')
recall = recall_score(y_test, y_pred, average='macro')
f1 = f1_score(y_test, y_pred, average='macro')
classification_rep_test = classification_report(y_test, y_pred)

# Tampilkan hasil evaluasi
print(f'Accuracy: {accuracy:.2f}')
print(f'Precision: {precision:.2f}')
print(f'Recall: {recall:.2f}')
print(f'F1-Score: {f1:.2f}')
print("\nClassification Report:")
print(classification_rep_test)

```

```

↳ Accuracy: 0.77
Precision: 0.75
Recall: 0.75
F1-Score: 0.75

```

```

Classification Report:

```

	precision	recall	f1-score	support
Negatif	0.67	0.70	0.68	20
Netral	0.71	0.62	0.67	24
Positif	0.87	0.93	0.90	29
accuracy			0.77	73
macro avg	0.75	0.75	0.75	73
weighted avg	0.76	0.77	0.76	73

Lampiran 10. Proses Klasifikasi Regresi Logistik

```

[12] from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y,test_size=0.1,random_state=42,)

```

```

[13] from sklearn.linear_model import LogisticRegression
from sklearn.feature_extraction.text import CountVecorizer
from sklearn.metrics import accuracy_score
from sklearn.feature_extraction.text import TfidfVectorizer

```

```

▶ vectorizer = TfidfVectorizer()
x_train_vec = vectorizer.fit_transform(x_train)
x_test_vec = vectorizer.transform(x_test)

```

```

▶ print(x_test)
326 wahana bermain air cocok keluarga sayang harga...
519 parkirana kendaraan berdebu dimusim kemarau pas...
580 wahana permainan air bergantian menyala
33 wahananya pilihanpokoknya cocok liburan keluar...
570 menyenangkan cocok liburan keluarga tpi makana...
...
174 wisata air menari terletak area perumahan buki...
23 pekan ramai pengunjung wahana yg disediakan pe...
554 sesuuu promosi marketingnya menawarkan paket d...
196 menyenangkan penjaganya ramah biayanya terjangkau...
405 nanya harga masuknya kesana
Name: ULASAN, Length: 73, dtype: object

```

```

[16] non_string_values = df[df['ULASAN'].apply(lambda x: not isinstance(x, str))]['ULASAN']
print("Non-string values in 'ULASAN' column:")
print(non_string_values)

```

```

Non-string values in 'ULASAN' column:
Series([], Name: ULASAN, dtype: object)

```

```

logistic_reg = LogisticRegression()
logistic_reg.fit(x_train_vec, y_train)

```

```

LogisticRegression
LogisticRegression()

```

```

y_pred = logistic_reg.predict(x_test_vec)

```

```

[] y_pred

```

```

array(['Positif', 'Netral', 'Positif', 'Positif', 'Positif', 'Negatif',
       'Positif', 'Positif', 'Netral', 'Negatif', 'Netral', 'Netral',
       'Netral', 'Positif', 'Negatif', 'Netral', 'Netral', 'Netral',
       'Positif', 'Negatif', 'Positif', 'Positif', 'Netral', 'Positif',
       'Positif', 'Netral', 'Negatif', 'Negatif', 'Netral', 'Positif',
       'Positif', 'Positif', 'Positif', 'Positif', 'Negatif', 'Netral',
       'Negatif', 'Positif', 'Netral', 'Netral', 'Negatif', 'Positif',
       'Positif', 'Negatif', 'Positif', 'Negatif', 'Netral', 'Negatif',
       'Negatif', 'Netral', 'Positif', 'Negatif', 'Positif', 'Netral',
       'Netral', 'Netral', 'Negatif', 'Positif', 'Negatif', 'Negatif',
       'Netral', 'Negatif', 'Positif', 'Negatif', 'Positif', 'Netral',
       'Positif', 'Positif', 'Positif', 'Positif', 'Negatif', 'Positif',
       ...])

```

Lampiran 12. Hasil Pengujian Regresi Logistik Pada Data Uji

```

# Hitung metrik evaluasi
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average='macro')
recall = recall_score(y_test, y_pred, average='macro')
f1 = f1_score(y_test, y_pred, average='macro')
classification_rep_test = classification_report(y_test, y_pred)

```

```

# Tampilkan hasil evaluasi
print(f'Accuracy: {accuracy:.2f}')
print(f'Precision: {precision:.2f}')
print(f'Recall: {recall:.2f}')
print(f'F1-Score: {f1:.2f}')
print("\nClassification Report:")
print(classification_rep_test)

```

```

Accuracy: 0.82
Precision: 0.81
Recall: 0.81
F1-Score: 0.81

```

```

Classification Report:

```

	precision	recall	f1-score	support
Negatif	0.71	0.75	0.73	20
Netral	0.81	0.71	0.76	24
Positif	0.90	0.97	0.93	29
accuracy			0.82	73
macro avg	0.81	0.81	0.81	73
weighted avg	0.82	0.82	0.82	73

Lampiran 13. Source code

Naive Bayes

```
!pip install seaborn
!pip install scikit-learn seaborn

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
import sklearn
warnings.simplefilter("ignore")

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import MultinomialNB
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics import confusion_matrix,
classification_report

#Import manajemen dataset
df=pd.read_excel("STOPWORD.xlsx", sheet_name = 'Sheet1')

df

df.describe()

df

# Membersihkan nilai np.nan pada kolom 'text'
df['ULASAN'].fillna('', inplace=True) # Mengganti np.nan dengan
string kosong

x = df['ULASAN']
y = df['LABEL']

x.describe()
```

```

y
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test =
train_test_split(x,y,test_size=0.1,random_state=42,)

from sklearn.naive_bayes import MultinomialNB
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.metrics import accuracy_score
from sklearn.feature_extraction.text import TfidfVectorizer

vectorizer = TfidfVectorizer()
x_train_vec = vectorizer.fit_transform(x_train)
x_test_vec = vectorizer.transform(x_test)

print(x_test)

non_string_values = df[df['ULASAN'].apply(lambda x: not
isinstance(x, str))]['ULASAN']
print("Non-string values in 'TEXT' column:")
print(non_string_values)

clf = MultinomialNB()
clf.fit(x_train_vec, y_train)

y_pred = clf.predict(x_test_vec)

y_pred

from sklearn.metrics import accuracy_score, precision_score,
recall_score, f1_score
from sklearn.metrics import accuracy_score, confusion_matrix,
classification_report
#accuracy = accuracy_score(y_test, y_pred)
#print("Akurasi: {:.2f}%".format(accuracy * 100))

# Hitung metrik evaluasi
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average='macro')
recall = recall_score(y_test, y_pred, average='macro')
f1 = f1_score(y_test, y_pred, average='macro')

```



```

classification_rep_test = classification_report(y_test, y_pred)

# Tampilkan hasil evaluasi
print(f'Accuracy: {accuracy:.2f}')
print(f'Precision: {precision:.2f}')
print(f'Recall: {recall:.2f}')
print(f'F1-Score: {f1:.2f}')
print("\nClassification Report:")
print(classification_rep_test)

# Data baru dalam bentuk string
new_text = ["suasana asri luas wahana lumayan lengkap"]
new_text_vec = vectorizer.transform(new_text)

print(new_text_vec)

new_prediction = clf.predict(new_text_vec)

print(new_prediction)

for i, (text, true_label, predicted_label) in
enumerate(zip(x_test, y_test, y_pred)):
    print(f"Index: {i}")
    print(f"Text: {text}")
    print(f"True Label: {true_label}")
    print(f"Predicted Label: {predicted_label}")
    print("-----")

# Membuat DataFrame dari X_test, y_test, dan hasil prediksi
df_results = pd.DataFrame({'Text': x_test, 'True Label': y_test,
'Predicted Label': y_pred})

# Menyimpan DataFrame ke file Excel
df_results.to_excel('hasilprediksiNB.xlsx', index=False)

from sklearn.metrics import accuracy_score, confusion_matrix,
classification_report

# Evaluasi performa pada data pengujian
accuracy_test = accuracy_score(y_test, y_pred)
conf_matrix_test = confusion_matrix(y_test, y_pred)

```

```

# Tampilkan hasil evaluasi pada data pengujian
print("\nEvaluasi Model Naive Bayes pada Data Pengujian:")
print(f"Accuracy: {accuracy_test:.2f}")
print("\nConfusion Matrix Naive Bayes:")
print(conf_matrix_test)

#HEATMAP CONF MATRIX
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import confusion_matrix
from sklearn.metrics import ConfusionMatrixDisplay
from sklearn.metrics import ConfusionMatrixDisplay,
confusion_matrix

vectorizer = TfidfVectorizer()
x_train_vec = vectorizer.fit_transform(x_train)
x_test_vec = vectorizer.transform(x_test)

# Hitung confusion matrix
conf_matrix = confusion_matrix(y_test, y_pred, labels=['Positif',
'Netral', 'Negatif'])

# Plot confusion matrix
sns.set(font_scale=1.2)
disp = ConfusionMatrixDisplay(conf_matrix,
display_labels=['Positif', 'Netral', 'Negatif'])
disp.plot(cmap='PuBuGn', values_format='d', include_values=True)

plt.title('Confusion Matrix (Naive Bayes)')
plt.show()

# Membaca hasil prediksi dari file Excel
df = pd.read_excel('/content/hasilprediksiNB.xlsx') # Gantilah
'hasil_prediksi_naive_bayes.xlsx' dengan nama file yang sesuai

positive_count = (df['Predicted Label'] == 'Positif').sum()
neutral_count = (df['Predicted Label'] == 'Netral').sum()
negative_count = (df['Predicted Label'] == 'Negatif').sum()

# Hitung jumlah distribusi pada setiap kelas pada data uji

```

```

class_distribution = y_test.value_counts()

# Tentukan warna untuk setiap kelas
colors = ['green', 'gray', 'red']

# Visualisasi jumlah distribusi pada setiap kelas pada data uji
plt.figure(figsize=(8, 6))
sns.barplot(x=class_distribution.index,
            y=class_distribution.values, palette=colors)
plt.title('Distribusi Kelas pada Data Uji Naive Bayes')
plt.xlabel('Kelas')
plt.ylabel('Jumlah')
# Menampilkan angka distribusi pada setiap bar
for i in range(len(y_test.value_counts())):
    plt.text(i, y_test.value_counts()[i],
            y_test.value_counts()[i], ha = 'center', va = 'bottom',
            color='black', size=12)
plt.show()

```

Regresi Logistik

```

!pip install seaborn
!pip install scikit-learn seaborn

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
import sklearn
warnings.simplefilter("ignore")

import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.metrics import confusion_matrix

#Import manajemen dataset
df=pd.read_excel("STOPWORD.xlsx", sheet_name = 'Sheet1')

df
df.describe()

```

```

df

# Membersihkan nilai np.nan pada kolom 'text'
df['ULASAN'].fillna('', inplace=True) # Mengganti np.nan dengan
string kosong

x = df['ULASAN']
y = df['LABEL']

x.describe()

from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test =
train_test_split(x,y,test_size=0.1,random_state=42,)

from sklearn.linear_model import LogisticRegression
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.metrics import accuracy_score
from sklearn.feature_extraction.text import TfidfVectorizer

vectorizer = TfidfVectorizer()
x_train_vec = vectorizer.fit_transform(x_train)
x_test_vec = vectorizer.transform(x_test)

print(x_test)

non_string_values = df[df['ULASAN'].apply(lambda x: not
isinstance(x, str))]['ULASAN']
print("Non-string values in 'ULASAN' column:")
print(non_string_values)

logistic_reg = LogisticRegression()
logistic_reg.fit(x_train_vec, y_train)

y_pred = logistic_reg.predict(x_test_vec)

y_pred

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

```

```

from sklearn.metrics import accuracy_score, confusion_matrix,
classification_report
#accuracy = accuracy_score(y_test, y_pred)
#print("Akurasi: {:.2f}%".format(accuracy * 100))

# Hitung metrik evaluasi
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average='macro')
recall = recall_score(y_test, y_pred, average='macro')
f1 = f1_score(y_test, y_pred, average='macro')
classification_rep_test = classification_report(y_test, y_pred)

# Tampilkan hasil evaluasi
print(f'Accuracy: {accuracy:.2f}')
print(f'Precision: {precision:.2f}')
print(f'Recall: {recall:.2f}')
print(f'F1-Score: {f1:.2f}')
print("\nClassification Report:")
print(classification_rep_test)

# Data baru dalam bentuk string
new_text = ["suasana asri luas wahana lumayan lengkap"]
new_text_vec = vectorizer.transform(new_text)

print(new_text_vec)

new_prediction = logistic_reg.predict(new_text_vec)

print(new_prediction)

for i, (text, true_label, predicted_label) in
enumerate(zip(x_test, y_test, y_pred)):
    print(f"Index: {i}")
    print(f"Text: {text}")
    print(f"True Label: {true_label}")
    print(f"Predicted Label: {predicted_label}")
    print("-----")

# Membuat DataFrame dari X_test, y_test, dan hasil prediksi
df_results = pd.DataFrame({'Text': x_test, 'True Label': y_test,
'Predicted Label': y_pred})

```

```

# Menyimpan DataFrame ke file Excel
df_results.to_excel('hasilprediksiRL.xlsx', index=False)

from sklearn.metrics import accuracy_score, confusion_matrix,
classification_report

# Evaluasi performa pada data pengujian
accuracy_test = accuracy_score(y_test, y_pred)
conf_matrix_test = confusion_matrix(y_test, y_pred)

# Tampilkan hasil evaluasi pada data pengujian
print("\nEvaluasi Model Regresi Logistik pada Data Pengujian:")
print(f"Accuracy: {accuracy_test:.2f}")
print("\nConfusion Matrix Regresi Logistik:")
print(conf_matrix_test)

#HEATMAP CONF MATRIX
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import confusion_matrix
from sklearn.metrics import ConfusionMatrixDisplay
from sklearn.metrics import ConfusionMatrixDisplay,
confusion_matrix

vectorizer = TfidfVectorizer()
x_train_vec = vectorizer.fit_transform(x_train)
x_test_vec = vectorizer.transform(x_test)

# Hitung confusion matrix
conf_matrix = confusion_matrix(y_test, y_pred, labels=['Positif',
'Netral', 'Negatif'])

# Plot confusion matrix
sns.set(font_scale=1.2)
disp = ConfusionMatrixDisplay(conf_matrix,
display_labels=['Positif', 'Netral', 'Negatif'])
disp.plot(cmap='PuBuGn', values_format='d', include_values=True)

plt.title('Confusion Matrix (Regresi Logistik)')
plt.show()

```

```

# Membaca hasil prediksi dari file Excel
df = pd.read_excel('/content/hasilprediksiRL.xlsx') # Gantilah
'hasil_prediksi_naive_bayes.xlsx' dengan nama file yang sesuai

positive_count = (df['Predicted Label'] == 'Positif').sum()
neutral_count = (df['Predicted Label'] == 'Netral').sum()
negative_count = (df['Predicted Label'] == 'Negatif').sum()

# Hitung jumlah distribusi pada setiap kelas pada data uji
class_distribution = y_test.value_counts()

# Tentukan warna untuk setiap kelas
colors = ['green', 'gray', 'red']

# Visualisasi jumlah distribusi pada setiap kelas pada data uji
plt.figure(figsize=(8, 6))
sns.barplot(x=class_distribution.index,
            y=class_distribution.values, palette=colors)
plt.title('Distribusi Kelas pada Data Uji Regresi Logistik')
plt.xlabel('Kelas')
plt.ylabel('Jumlah')
# Menampilkan angka distribusi pada setiap bar
for i in range(len(y_test.value_counts())):
    plt.text(i, y_test.value_counts()[i],
            y_test.value_counts()[i], ha = 'center', va = 'bottom',
            color='black', size=12)
plt.show()

```

Lampiran 14. Surat Keterangan Bebas Plagiasi



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

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

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

SURAT KETERANGAN BEBAS PLAGIAT

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

Nama : Jumardin Ruslan
Nim : 105841104419
Program Studi : Teknik Informatika

Dengan nilai:

No	Bab	Nilai	Ambang Batas
1	Bab 1	3 %	10 %
2	Bab 2	19 %	25 %
3	Bab 3	7 %	10 %
4	Bab 4	8 %	10 %
5	Bab 5	0 %	5 %

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

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

Makassar, 21 Februari 2024
Mengetahui,

Kepala UPT- Perpustakaan dan Penerbitan,



Jl. Sultan Alauddin no 259 makassar 90222
Telepon (0411)866972,881 593,fax (0411)865 588
Website: www.library.unismuh.ac.id
E-mail : perpustakaan@unismuh.ac.id

Lampiran 15. Hasil Scan Plagiasi Per Bab



AB I Jumardin Ruslan 105841104419

ORIGINALITY REPORT

3 SIMILARITY INDEX	 LULUS	3% INTERNET SOURCES	1% PUBLICATIONS	1% STUDENT PAPERS
------------------------------	--	-------------------------------	---------------------------	-----------------------------

PRIMARY SOURCES

1	Submitted to Universitas Muhammadiyah Makassar Student Paper	1%
2	123dok.com Internet Source	1%
3	es.scribd.com Internet Source	1%

Exclude quotes Off
Exclude bibliography Off

Exclude matches Off





BAB II Jumardin Ruslan

105841104419

by Tahap Tutup

Submission date: 19-Feb-2024 03:33PM (UTC+0700)

Submission ID: 2298656071

File name: BAB_II_-_2024-02-19T164122.094.docx (537.21K)

Word count: 2824

Character count: 18383

AB II Jumardin Ruslan 105841104419

ORIGINALITY REPORT

19% LULUS 7%

SIMILARITY INDEX

INTERNET SOURCES

1%

PUBLICATIONS

17%

STUDENT PAPERS

PRIMARY SOURCES

Rank	Source	Percentage
1	Submitted to Universitas Muhammadiyah Makassar Student Paper	15%
2	repository.its.ac.id Internet Source	2%
3	ejurnal.its.ac.id Internet Source	2%
4	eprints.uny.ac.id Internet Source	<1%

Exclude quotes

Exclude matches

Exclude bibliography

BAB III Jumardin Ruslan

105841104419

by Tahap Tutup



Submission date: 20-Feb-2024 02:14PM (UTC+0700)

Submission ID: 2299467546

File name: BAB_III_-_2024-02-20T152426.725.docx (88.23K)

Word count: 1043

Character count: 7079

BAB III Jumardin Ruslan 105841104419

ORIGINALITY REPORT

7%



SIMILARITY INDEX

INTERNET SOURCES

2%

PUBLICATIONS

3%

STUDENT PAPERS

PRIMARY SOURCES

1

123dok.com

Internet Source

1%

2

Submitted to Universiti Teknologi Petronas

Student Paper

1%

3

id.scribd.com

Internet Source

1%

4

Submitted to UIN Sultan Syarif Kasim Riau

Student Paper

1%

5

Submitted to stipram

Student Paper

1%

6

journals.plos.org

Internet Source

1%

Exclude quotes Off

Exclude matches Off

Exclude bibliography Off

BAB IV Jumardin Ruslan

105841104419

by Tahap Tutup



Submission date: 20-Feb-2024 02:14PM (UTC+0700)

Submission ID: 2299467751

File name: BAB_IV_-_2024-02-20T152428.786.docx (1.23M)

Word count: 1816

Character count: 11582

AB IV Jumardin Ruslan 105841104419

ORIGINALITY REPORT

8%

SIMILARITY INDEX



4%

INTERNET SOURCES

6%

PUBLICATIONS

2%

STUDENT PAPERS

PRIMARY SOURCE



1	N. K. Kutha Ardana, Ruhiyat Ruhiyat, Nurfatimah Amany, Teofilus Kevin Irawan, Raymond Raymond, Rizalius Karunia, Syifa Fauzia. "Perbandingan Metode KNN, Naive Bayes, dan Regresi Logistik Binomial dalam Pengklasifikasian Status Ekonomi Negara", <i>Jambura Journal of Mathematics</i> , 2023 Publication	2%
2	ejournal.unesa.ac.id Internet Source	1%
3	Virginia Khoirunnisa, Sri Lestari. "IMPLEMENTASI KLASIFIKASI KEHAMILAN BERESIKO DENGAN METODE NAIVE BAYES PADA PUSKESMAS KELURAHAN MALAKA JAYA", <i>Jurnal Indonesia : Manajemen Informatika dan Komunikasi</i> , 2023 Publication	1%
4	repository.its.ac.id Internet Source	1%
5	Submitted to Universitas Islam Indonesia Student Paper	1%

BAB V Jumardin Ruslan

105841104419

by Tahap Tutup



Submission date: 20-Feb-2024 02:16PM (UTC+0700)

Submission ID: 2299468383

File name: BAB_V_-_2024-02-20T152429.823.docx (17.62K)

Word count: 614

Character count: 4008

BAB V Jumardin Ruslan 105841104419

ORIGINALITY REPORT

0%

SIMILARITY INDEX



0%

INTERNET SOURCES

0%

PUBLICATIONS

0%

STUDENT PAPERS

PRIMARY SOURCES

Exclude quotes

Exclude bibliography

Exclude matches

