

STUDY COMPLETION PROGRAM
INTERNSHIP TRACK

LAPORAN AKHIR INTERNSHIP

**ANALISIS SENTIMEN ULASAN PELANGGAN DAN
DAMPAKNYA TERHADAP LOYALITAS DI PT SANRAYA
ADI NATTAYA MENGGUNAKAN METODE NAIVE BAYES**



Oleh :

Nama : Sarah Ayu Rahmawati
NIM : 20210040084

FAKULTAS TEKNIK, KOMPUTER DAN DESAIN
PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS NUSA PUTRA
2024

LAPORAN AKHIR INTERNSHIP

ANALISIS SENTIMEN ULASAN PELANGGAN DAN DAMPAKNYA TERHADAP LOYALITAS DI PT SANRAYA ADI NATTAYA MENGUNAKAN METODE NAIVE BAYES

*Diajukan Untuk Memenuhi Salah Satu Syarat
Dalam Menempuh Seminar Hasil Internship pada Jalur Program Internship
Di Program Studi Teknik Informatika*



Oleh:

Nama : Sarah Ayu Rahmawati
NIM : 20210040084

**FAKULTAS TEKNIK, KOMPUTER DAN DESAIN
PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS NUSA PUTRA
2024**

LEMBAR PENGESAHAN
LAPORAN AKHIR INTERNSHIP

**JUDUL: ANALISIS SENTIMEN ULASAN PELANGGAN DAN
DAMPAKNYA TERHADAP LOYALITAS DI PT SANRAYA ADI
NATTAYA MENGGUNAKAN METODE NAIVE BAYES**

Disusun Oleh:

NAMA : Sarah Ayu Rahmawati

NIM : 20210040084

Laporan ini telah diseminarkan dihadapan penguji seminar *Internship* pada Program Internship
di Program Studi Teknik Informatika

Sukabumi, 2024

Ketua Penguji

Pembimbing Utama

.....

NIDN.

Somantri S. T., M.Kom

NIDN. 0419128801

KaProdi

Somantri S. T., M. Kom

NIDN. 0419128801

ABSTRAK

Perkembangan teknologi informasi dan komunikasi telah mengubah cara pelanggan berinteraksi dengan perusahaan, di mana ulasan dan umpan balik pelanggan menjadi sumber informasi yang sangat berharga. PT Sanraya Adi Nattaya, yang beroperasi di bidang experience menghadapi tantangan dalam memahami dan memanfaatkan ulasan pelanggan untuk meningkatkan kepuasan dan loyalitas. Penelitian ini bertujuan untuk mengeksplorasi peran sentiment analyst dalam mengubah ulasan menjadi loyalitas pelanggan serta strategi yang dapat diterapkan untuk meningkatkan kepuasan pelanggan.

Metode yang digunakan dalam penelitian ini meliputi pengumpulan data dari platform ulasan dari Goole PlayStore, diikuti dengan penerapan teknik analisis sentimen untuk mengidentifikasi opini yang terkandung dalam umpan balik pelanggan. Hasil analisis menunjukkan bahwa sentimen positif berkontribusi signifikan terhadap tingkat kepuasan dan loyalitas pelanggan, sementara sentimen negatif mengindikasikan area yang perlu diperbaiki

Berdasarkan temuan tersebut, penelitian ini merekomendasikan beberapa strategi untuk PT Sanraya Adi Nattaya, termasuk peningkatan respons terhadap umpan balik pelanggan, pengembangan produk yang lebih sesuai dengan harapan pelanggan, serta penerapan program loyalitas yang lebih menarik. Dengan memanfaatkan analisis sentimen secara efektif, perusahaan dapat meningkatkan pengalaman pelanggan, membangun loyalitas yang lebih kuat, dan mencapai keunggulan kompetitif di pasar

Kata kunci: *Analisis Sentimen, Naive Bayes, Loyalitas Pelanggan, Ulasan Pelanggan.*

DAFTAR ISI

HALAMAN JUDUL	ii
LEMBAR PENGESAHAN	ii
ABSTRAK.....	iii
DAFTAR ISI.....	iv
DAFTAR GAMBAR.....	v
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Tujuan Mengikuti Internship.....	2
1.3 Manfaat Mengikuti Internship.....	2
BAB II PELAKSANAAN INTERNSHIP	4
2.1 Profil Perusahaan Tempat Internship	4
2.2 Kegiatan Internship	6
BAB III HASIL DAN PEMBAHASAN	7
3.1 Metode Penelitian.....	7
3.2 Hasil Analisis Sentimen	12
BAB IV PENUTUP	21
DAFTAR PUSTAKA	22
LAMPIRAN : DOKUMENTASI INTERNSHIP.....	23

DAFTAR GAMBAR

Gambar 2. 1 Logo Perusahaan	5
Gambar 2. 2 Bagan Organisasi Perusahaan	
Gambar 3.1 1 Diagram Alur Analisis	
Gambar 3.1 2 Import Library	
Gambar 3.1 3 Import File.....	9
Gambar 3.1 4 Cleaning Data.....	9
Gambar 3.1 5 Tokenization.....	10
Gambar 3.1 6 Lemmatization and Removing Stopword	10
Gambar 3.1 7Ekspolarasi Data Analis	11
Gambar 3.1 8 Prediksi.....	
Gambar 3.2 1 Hasil Pemrosesan Data.....	13
Gambar 3.2 2 Hasil EDA	13
Gambar 3.2 3 Persentase Rating	14
Gambar 3.2 4 Positif Review Wold Cloud	15
Gambar 3.2 5Negatif Review Word Cloud.....	15
Gambar 3.2 6 Pembagian data set.....	16
Gambar 3.2 7 Confusion Matrix	
Gambar 3.2 8 Classification refort.....	18
Gambar 3.2 9 Hasil K-Ford Cross Validation	
Gambar 3.2 10 Prediksi.....	
Gambar 3.2 11 confusion matrix(y_test, y_pred)	19

BAB I PENDAHULUAN

1.1 Latar Belakang

Di era digital yang semakin berkembang, interaksi antara perusahaan dan pelanggan telah mengalami transformasi yang signifikan. Ulasan dan umpan balik dari pelanggan kini menjadi salah satu elemen kunci dalam membentuk citra dan reputasi suatu perusahaan. Menurut survei terbaru, lebih dari 80% konsumen mengandalkan ulasan online dalam pengambilan keputusan mereka, menjadikan ulasan sebagai sumber informasi yang sangat berpengaruh. Dalam konteks ini, PT Sanraya Adi Nattaya, yang beroperasi di bidang experiente, tidak dapat mengabaikan pentingnya memahami dan menganalisis sentimen yang terkandung dalam ulasan pelanggan

Sentiment analyst berperan penting dalam mengubah data ulasan menjadi wawasan yang dapat ditindaklanjuti. Dengan menggunakan teknik analisis sentimen, perusahaan dapat mengidentifikasi emosi, opini, dan persepsi pelanggan terhadap produk atau layanan yang ditawarkan. Proses ini tidak hanya membantu dalam memahami kepuasan pelanggan saat ini, tetapi juga memberikan gambaran yang lebih jelas tentang area yang perlu diperbaiki. Dengan demikian, analisis sentimen menjadi alat strategis yang dapat digunakan untuk meningkatkan pengalaman pelanggan secara keseluruhan.

Loyalitas pelanggan adalah tujuan akhir yang ingin dicapai oleh setiap perusahaan. Pelanggan yang loyal tidak hanya akan melakukan pembelian berulang, tetapi juga berpotensi menjadi duta merek yang merekomendasikan produk kepada orang lain. Oleh karena itu, penting bagi PT Sanraya Adi Nattaya untuk mengembangkan strategi yang efektif dalam mengubah ulasan pelanggan menjadi loyalitas. Dengan memanfaatkan analisis sentimen, perusahaan dapat merespons umpan balik pelanggan dengan lebih cepat dan tepat, serta menciptakan pengalaman yang lebih memuaskan tetapi juga menarik pelanggan baru di pasar yang semakin kompetitif.

1.2 Tujuan Mengikuti Internship

Adapun tujuan penulis dalam mengikuti kegiatan *Internship* ini meningkatkan kepuasan customer dengan melakukan analisis sentiment, selain itu tujuan dari intership ini antara lain:

- a. Membangun jalur relasi yang baik dalam dunia pekerjaan sehingga terciptanya nama yang baik bagi mahasiswa dan nama universitas sendiri.
- b. Untuk mempersiapkan diri seorang mahasiswa dalam menghadapi dunia kerja sesungguhnya sehingga mahasiswa bisa siap masuk kedalam dunia pekerjaan.
- c. Membangun kepercayaan diri, dalam hal ini kepercayaan diri mampu terbangun dimana program ini dijalankan, hal ini mahasiswa mampu memulai mencoba ditengah lingkungan yang sudah lebih dahulu.
- d. Meneliti hubungan antara sentimen pelanggan terhadap produk atau layanan dan tingkat loyalitas yang ditunjukkan oleh pelanggan.

1.3 Manfaat Mengikuti Internship

Internship merupakan proses untuk mempelajari praktek-praktek yang nyata di dalam dunia usaha, sehingga diharapkan dalam kegiatan internship ini memberikan manfaat sebagai berikut:

- a. Bagi Penulis
 - Penulis dapat mengaplikasikan ilmu serta teori-teori yang didapat selama kegiatan perkuliahan selama mengikuti kegiatan *intership*.
 - Dapat memberikan pengetahuan seputar dunia kerja yang sesungguhnya.
- b. Bagi Perusahaan
 - Dengan mengubah ulasan menjadi tindakan yang konkret, perusahaan dapat membangun loyalitas pelanggan yang lebih kuat, yang berpotensi meningkatkan penjualan dan pangsa pasar.
 - Penelitian ini dapat memberikan wawasan yang mendalam tentang kebutuhan dan harapan pelanggan, sehingga PT Sanraya Adi Nattaya dapat meningkatkan kualitas produk dan layanan yang ditawarkan.

c. Bagi Pihak Universitas

Menambah relasi atau hubungan yang baik dengan PT SANRAYA ADI NATTAYA untuk kedepan-nya dengan mengirimkan mahasiswanya ke perusahaan -perusahaan tersebut untuk mengikuti kegiatan *Internship*.

BAB II PELAKSANAAN INTERNSHIP

2.1 Profil Perusahaan Tempat Internship

Didirikan pada tahun 2016, PT. Sanraya Adi Nattaya, atau yang sering dikenal dengan GOERS merupakan sebuah perusahaan teknologi berbasis di Indonesia, yang awalnya berdiri sebagai manajemen tiket acara dan aplikasi platform acara. Pelanggan dapat menemukan sesuatu yang baru untuk dilakukan setiap hari, seperti konser festival, kelas, konferensi, acara gratis, dan banyak lagi. GOERS mendorong para pelanggannya untuk terhubung dengan menggabungkan 2 kreativitas dan teknologi. Dengan satu juta lebih pengguna dan 2.500 lebih mitra, GOERS memastikan para tim untuk menerapkan pikiran kreatif, demi menciptakan pengalaman yang mulus dan efisien bagi pelanggan.

Adanya GOERS sebagai platform digital diharapkan dapat membantu para pengguna untuk menjalani hidup mereka sepenuhnya sambil tetap mendukung bisnis dan komunitas lokal untuk bertumbuh dan berkembang. Untuk mewujudkannya, GOERS memiliki lima nilai inti. Nilai pertama adalah Growth, yaitu bersemangat untuk beradaptasi, mengejar pertumbuhan dan pembelajaran, serta bersemangat dan bertekad. Kedua adalah Ownership, yaitu membangun tim yang positif dan selalu mengedepankan kebersamaan. Ketiga adalah Excellence, yaitu memberikan hasil terbaik melalui pemikiran inovatif untuk memenangkan pasar. Keempat adalah Reliable, yaitu bagaimana GOERS mampu menjadi pemecah masalah. Kelima adalah Solid, yaitu menghargai satu sama lain melalui komunikasi terbuka.

Semakin bergerak dinamis, GOERS memiliki sebuah ekosistem yang diharapkan dapat memenuhi segala kebutuhan dari berbagai macam pihak. Ekosistem GOERS terbagi menjadi empat bagian. Pertama adalah GEM Solution, yaitu solusi terintegrasi untuk reservasi, pembayaran, serta penjualan dan kunjungan di destinasi wisata. Kedua adalah Goersapp. Goersapp merupakan sebuah marketplace yang menjual tiket destinasi wisata, acara dan direktori aktivitas, promo, serta wisata dengan kurasi berdasarkan preferensi pengguna berbasis aplikasi dan website. Selanjutnya adalah Nattaya. Nattaya hadir sebagai solusi manajemen penyelenggaraan MICE, lengkap dengan teknologi manajemen

kunjungan khusus untuk event dari skala kecil hingga besar. Terakhir adalah Sanraya.id, yaitu solusi dan konsultasi pembuatan aplikasi maupun website untuk kebutuhan perusahaan secara spesifik dalam rangka transformasi dan akselerasi digital proses bisnis.

2.1.1 Logo dan Visi Misi Perusahaan

Logo merupakan salah satu simbol suatu perusahaan yang merupakan ciri khas atau pembeda dari usaha atau perusahaan lainnya.



Gambar 2. 1 Logo Perusahaan

MISI Connecting People with Experience - perlu diwujudkan oleh setiap bisnis di sektor *experience*. Kami percaya bahwa kemitraan tidak hanya melanjutkan pertumbuhan dan pengembangan bagi perusahaan kami saja, tetapi juga para mitra kami. Maka dengan "berlari" berdampingan, bersama menuju pertumbuhan, keberlanjutan, proses bisnis yang praktis dan efisien, serta siap untuk mengatasi gangguan di masa depan. Ekosistem digital kami berorientasi pada pelanggan, menjadikannya prioritas utama kami. Kami membantu mendorong inovasi di seluruh industri: mulai dari hiburan, pariwisata hingga pendidikan, mari kita berkembang bersama!

2.1.2 Bagan Organisasi Perusahaan

PT. Sanraya Adi Nattaya terletak di Menara MTH, Jalan Letjen M.T. Haryono, RT.10/RW.9, Tebet Timur, Kecamatan Tebet, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12730. Per bulan Oktober 2022, GOERS memiliki 42 orang karyawan, yang terdiri dari 28 karyawan tetap dan 14 karyawan magang.



Gambar 2. 2 Bagan Organisasi Perusahaan

2.2 Kegiatan Internship

Pelaksanaan kegiatan internship ini dilaksanakan di PT SANRAYA ADI NATTAYA Pada intenship terdapat di PT SANRAYA ADI NATTAYA mahasiswa dikenalkan dengan berbagai pemecahan kasus yang menjadi fokus utama dalam sebuah perusahaan, yaitu customer. Kepuasan customer merupakan salah satu indikator kunci kesuksesan suatu perusahaan. PT Sanraya Adi Nattaya, sebagai platform yang menghubungkan orang dengan beragam experience dengan memberikan solusi teknologi untuk experience creators (seperti event organizer dan pemilik venue/destinasi wisata) dan memberikan opsi experience seru untuk para experience seekers melalui aplikasi dan website, harus memastikan bahwa web perusahaannya berfungsi optimal untuk mendukung kepuasan customer.

BAB III HASIL DAN PEMBAHASAN

3.1 Metode Penelitian

A. Metode Pengumpulan Data

Dalam penelitian ini, saya menganalisis 409 ulasan pengguna. Pengumpulan data dilakukan dengan mengakses ulasan yang diberikan oleh pengguna aplikasi PT Sanraya Adi Nattaya di Google Play Store. Proses pengumpulan data melibatkan langkah-langkah berikut

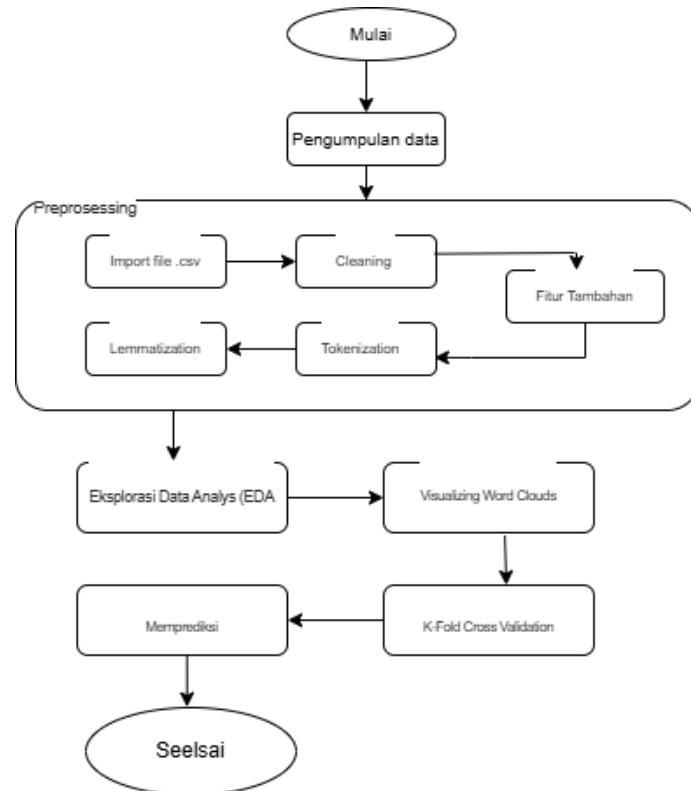
- 1) **Akses Platform:** Mengunjungi halaman aplikasi PT Sanraya Adi Nattaya di Google Play Store untuk mengumpulkan ulasan yang ditinggalkan oleh pengguna.
- 2) **Penyimpanan Data:** Menyimpan data yang telah dikumpulkan dalam format yang terstruktur, seperti spreadsheet atau database, untuk memudahkan analisis lebih lanjut
- 3) **Analisis Kualitas Data:** Melakukan pemeriksaan kualitas data untuk memastikan bahwa informasi yang dikumpulkan relevan dan dapat diandalkan sebelum digunakan dalam analisis

B. Metode Analisis

Analisis ini dilakukan dengan menggunakan metode Naive Bayes. Naive Bayes merupakan metode klasifikasi yang menggunakan pendekatan probabilistik (pembelajaran terbimbing). Pendekatan ini mengasumsikan bahwa data dihasilkan menggunakan model probabilistik (Musa'adah et al., 2018). Menurut Andrew McCallum dan Kamal Nigam (Musa'adah et al., 2018), klasifikasi Naive Bayesian adalah metode naif yang mengasumsikan bahwa semua atribut dalam data bersifat independen berdasarkan konteks kelas.

C. Tahapan Analisis Sentimen

Dalam melakukan analisis sentimen menggunakan *python*, setelah melakukan pengumpulan ada beberapa tahapan untuk melakukan analisis. Berikut menunjukkan alur analisis sentimen:



Gambar 3.1 1 Diagram Alur Analisis

1) Menyiapkan Libraries

Pada tahapan ini, import library python yang dibutuhkan seperti pandas, textblob, numpy, seaborn dan yang lainnya. Python telah sangat mempermudah kita dalam melakukan berbagai pekerjaan, salah satu alasannya adalah karena Python dilengkapi dengan berbagai library yang cukup powerful.

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
sns.set_style("whitegrid")

#set warning
import warnings
warnings.filterwarnings('ignore')

pd.pandas.set_option('display.max_columns', None)

```

Gambar 3.1 2 Import Library

2) Melakukan Pemrosesan Data

langkah selanjutnya adalah melakukan preprocessing data. Tujuan dari *preprocessing* data ini adalah untuk mengolah dan mempersiapkan data sebelum melanjutkan ke tahap analisis berikutnya.

- **Import terlebih dahulu file.csv**

```

filename = "UlasanGoersApp.csv"
df = pd.read_csv(filename, encoding = 'latin-1')
df.head()

```

Gambar 3.1 3 Import File

- **Cleaning the Text**

```

def clean_text(text):
    if isinstance(text, str):
        return re.sub('[^a-zA-Z]', ' ', text).lower()
    return ''

df['cleaned_text'] = df['Ulasan'].apply(clean_text)
df['label'] = df['Rating'].map({1.0: 0, 2.0: 0, 3.0: 0, 4.0: 1, 5.0: 1})

print(df[['Ulasan', 'cleaned_text', 'Rating', 'label']].head())

```

Gambar 3.1 4 Cleaning Data

Proses pembersihan mencakup penyesuaian format teks seperti mengubah huruf-huruf yang tidak kapital menjadi kapital untuk memastikan konsistensi. Selain itu, juga memberikan label pada data untuk mengklasifikasikan sentimen 1.0:0, 2.0:0, 3.0:0, 4.0:1, dan 5.0:1, di mana 0 menunjukkan sentimen negatif dan 1 menunjukkan sentimen positif.

- **Tokenization**




```
def tokenize_text(text):
    tokenized_text = text.split()
    return tokenized_text
df['tokens'] = df['cleaned_text'].apply(lambda x: tokenize_text(x))
df.head()
```

Gambar 3.1 5 Tokenization

Proses tokenization bertujuan untuk mengubah dokumen (Ulasan) yang awalnya kalimat utuh menjadi bagian yang lebih kecil, yaitu kata-kata individual.

- **Lemmatization and Removing Stopword**



```
import nltk
nltk.download('stopwords')
from nltk.corpus import stopwords
all_stopwords = stopwords.words('english')
all_stopwords.remove('not')
import nltk
nltk.download('wordnet')
nltk.download('omw-1.4')
def lemmatize_text(token_list):
    return " ".join([lemmatizer.lemmatize(token) for token in token_list if not token in set(all_stopwords)])

lemmatizer = nltk.stem.WordNetLemmatizer()
df['lemmatized_review'] = df['tokens'].apply(lambda x: lemmatize_text(x))
df.head()
```

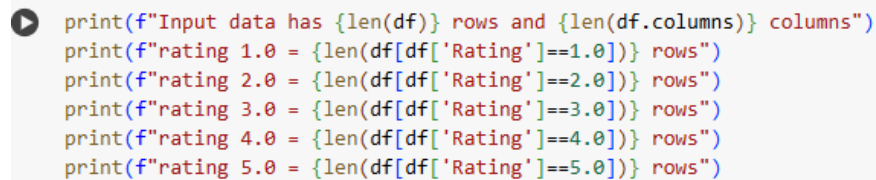
Gambar 3.1 6 Lemmatization and Removing Stopword

Lemmatization adalah teknik yang digunakan untuk mengubah kata-kata ke bentuk dasarnya atau lemma-nya. Proses ini tidak hanya mempertimbangkan bentuk kata, tetapi juga makna dan konteksnya. Sementara itu, proses Removing Stopwords bertujuan untuk menghapus kata-kata umum yang tidak memberikan makna signifikan dalam analisis, seperti "dan," "atau," "adalah," dan "yang."

3) Membuat Eksplorasi Data Analys (EDA)

Eksplorasi Data Analisis (Exploratory Data Analysis - EDA) mencakup proses kritis yang melibatkan uji investigasi awal terhadap sebuah dataset untuk mengidentifikasi pola, menemukan anomali, menguji hipotesis, dan memeriksa asumsi melalui penggunaan statistik ringkasan dan representasi grafis (visual).

- Bentuk dataset dan Perincian kelas



```
print(f"Input data has {len(df)} rows and {len(df.columns)} columns")
print(f"rating 1.0 = {len(df[df['Rating']==1.0])} rows")
print(f"rating 2.0 = {len(df[df['Rating']==2.0])} rows")
print(f"rating 3.0 = {len(df[df['Rating']==3.0])} rows")
print(f"rating 4.0 = {len(df[df['Rating']==4.0])} rows")
print(f"rating 5.0 = {len(df[df['Rating']==5.0])} rows")
```

Gambar 3.1 7Eksplorasi Data Analis

4) Visualizing Word Cload

Word Cloud adalah sebuah representasi visual yang terdiri dari kumpulan kata-kata, di mana ukuran dan ketebalan setiap kata mencerminkan frekuensi kemunculan atau tingkat kepentingan kata tersebut dalam suatu dokumen teks. Dalam konteks ini, semakin besar ukuran suatu kata, semakin sering kata tersebut muncul dalam teks yang dianalisis. Dengan demikian, Word Cloud tidak hanya berfungsi sebagai alat visualisasi, tetapi juga sebagai sarana untuk mengidentifikasi tema atau topik utama yang terkandung dalam kumpulan data teks.

5) K-Ford Cross Validation

K-Fold Cross Validation adalah sebuah teknik yang digunakan dalam machine learning untuk memisahkan dataset menjadi dua bagian, yaitu data training dan data testing, dengan tujuan untuk mengevaluasi kinerja model secara lebih akurat. Prosedur ini melibatkan pembagian dataset menjadi K subset yang sama besar, di mana setiap subset akan bergantian berfungsi sebagai data testing, sementara K-1 subset lainnya digunakan sebagai data training.

6) Memprediksi

Pada tahap akhir dalam proses pemodelan, kita melakukan prediksi dengan menggunakan dua pendekatan utama, yaitu CountVectorizer (Bag of Words) dan TF-IDF (Term Frequency-Inverse Document Frequency). CountVectorizer berfungsi untuk menghitung jumlah frekuensi kemunculan setiap kata dalam setiap dokumen, menghasilkan representasi numerik yang berupa bilangan cacah (bulat) tanpa pecahan. Meskipun metode ini sederhana dan mudah diterapkan, ia memiliki keterbatasan, terutama dalam konteks analisis yang lebih mendalam.

```
# Menggunakan TfidfVectorizer untuk mengubah teks menjadi representasi numerik
tfidf = TfidfVectorizer(max_df=0.5, min_df=2)
tfidf_train = tfidf.fit_transform(X_train['lemmatized_review'])
tfidf_test = tfidf.transform(X_test['lemmatized_review'])

# Melatih model SVC
classifier = SVC(kernel='linear', random_state=10)
classifier.fit(tfidf_train, y_train)
```

Gambar 3.1 8 Prediksi

3.2 Hasil Analisis Sentimen

1. Pemrosesan Data

Hasil pemrosesan data menunjukkan bahwa setelah melalui tahap pembersihan dan pra-pemrosesan, dataset yang terdiri dari 409 ulasan pengguna berhasil disiapkan untuk analisis lebih lanjut. Proses tokenization menghasilkan sejumlah kata individual yang relevan, sementara lemmatization memastikan bahwa setiap kata diubah ke bentuk dasarnya, sehingga mengurangi variasi yang tidak perlu. Selain itu, proses removing stopwords berhasil menghapus kata-kata umum yang tidak memberikan makna signifikan, sehingga fokus analisis dapat diarahkan pada kata-kata kunci yang lebih informatif.

	Ulasan	Rating	cleaned_text	label	tokens	lemmatized_review
0	Aku sangat seneng bgt bisa tau apk goers krna ...	5	aku sangat seneng bgt bisa tau apk goers krna ...	1	[aku, sangat, seneng, bgt, bisa, tau, apk, goe...	aku sangat seneng bgt bisa tau apk goer krna a...
1	Aplikasi goers lebih memudahkan kita untuk mem...	5	aplikasi goers lebih memudahkan kita untuk mem...	1	[aplikasi, goers, lebih, memudahkan, kita, unt...	aplikasi goer lebih memudahkan kita untuk memb...
2	Aplikasi goers lebih memudahkan kita untuk mem...	5	aplikasi goers lebih memudahkan kita untuk mem...	1	[aplikasi, goers, lebih, memudahkan, kita, unt...	aplikasi goer lebih memudahkan kita untuk memb...
3	Aplikasinya keren sekali, dari mulai tampilan ...	5	aplikasinya keren sekali dari mulai tampilan ...	1	[aplikasinya, keren, sekali, dari, mulai, tamp...	aplikasinya keren sekali dari mulai tampilan a...
4	Sangat mudah untuk membeli tiket event konser-...	5	sangat mudah untuk membeli tiket event konser ...	1	[sangat, mudah, untuk, membeli, tiket, event, ...	sangat mudah untuk membeli tiket event konser ...

Gambar 3.2 1 Hasil Pemrosesan Data

2. Exploratory Data Analysis – EDA

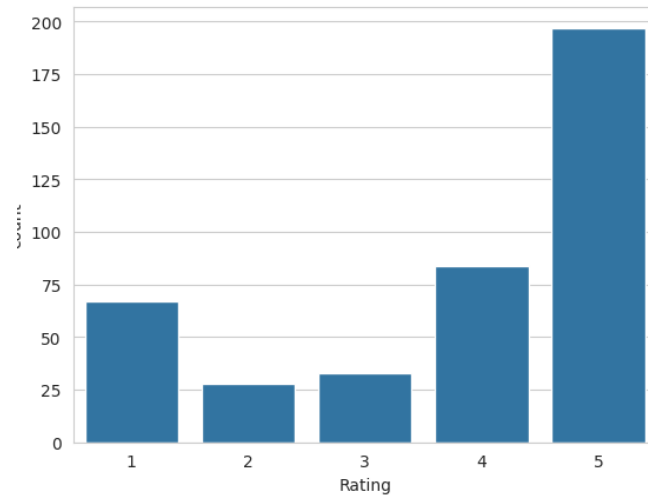
- Dari hasil Exploratory Data Analysis – EDA menunjukkan bahwa Sekitar 16,37% pengguna memberikan rating atau penilaian bintang 1, yang menunjukkan ketidakpuasan yang cukup tinggi terhadap aplikasi ini. 6,84% pengguna memberikan bintang 2, yang mengindikasikan bahwa meskipun ada beberapa aspek yang kurang memuaskan, mereka masih melihat potensi dalam aplikasi tersebut. Sebanyak 8,06% pengguna memberikan bintang 3, yang mencerminkan pandangan netral, di mana mereka merasa aplikasi ini tidak terlalu baik, tetapi juga tidak sepenuhnya buruk. 20,54% pengguna memberikan bintang 4, menunjukkan bahwa mereka memiliki pengalaman yang positif dan menghargai fitur-fitur tertentu dari aplikasi. Terakhir, sekitar 48,10% pengguna memberikan bintang 5, yang menandakan tingkat kepuasan yang tinggi dan rekomendasi kuat untuk aplikasi ini

```

Input data has 409 rows and 6 columns
rating 1.0 = 67 rows
rating 2.0 = 28 rows
rating 3.0 = 33 rows
rating 4.0 = 84 rows
rating 5.0 = 197 rows

```

Gambar 3.2 2 Hasil EDA



Gambar 3.2 3 Persentase Rating

3. Visualizing Word Cloud

Hasil visualisasi Word Cloud menunjukkan representasi grafis dari kata-kata yang paling sering muncul dalam ulasan pengguna. Dalam Word Cloud ini, kata-kata seperti 'bagus', 'layanan', dan 'cepat' terlihat lebih besar, menandakan frekuensi kemunculan yang tinggi dan menunjukkan aspek positif dari pengalaman pelanggan. Sebaliknya, kata-kata seperti 'buruk' dan 'lambat' juga muncul, meskipun dengan ukuran yang lebih kecil, yang mengindikasikan adanya sentimen negatif yang perlu diperhatikan.

Positive Reviews Word Cloud



Gambar 3.2 4 Positif Review Word Cloud

Negative Reviews Word Cloud



Gambar 3.2 5Negatif Review Word Cloud

ALGORITMA NAIVE BAYES

Sekitar 69.9% dari total data digunakan untuk pelatihan, sementara 30.1% digunakan untuk pengujian. Pembagian ini merupakan praktik umum dalam machine learning yang bertujuan untuk memastikan bahwa model dapat belajar dari data yang cukup besar sebelum diuji pada data yang tidak terlihat. Dengan cara ini, kita dapat mengevaluasi seberapa baik model dapat menggeneralisasi pengetahuan yang diperoleh dari data pelatihan ke data baru.

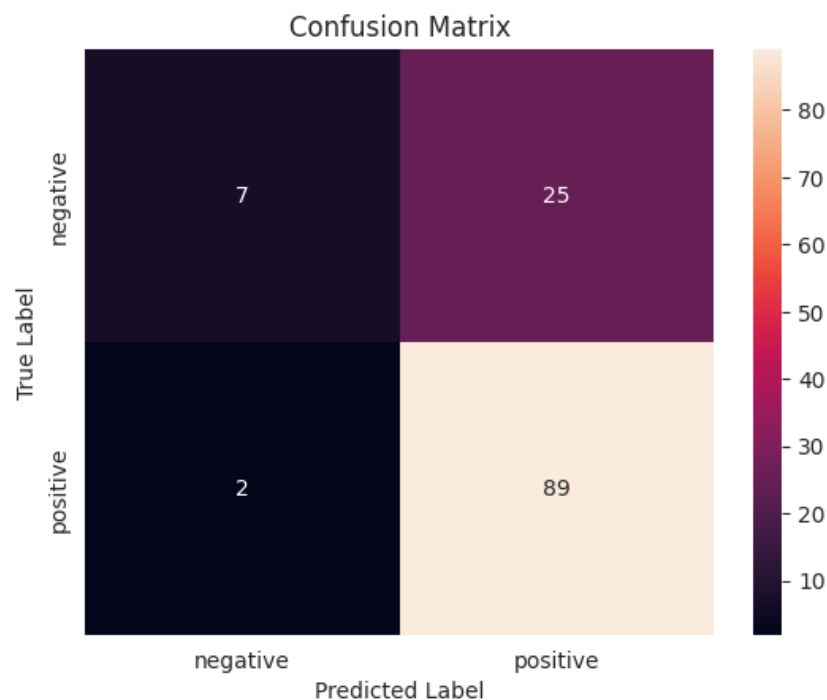
```
Shape X: (409, 3)
Shape y: (409,)
Shape X_train: (286, 3)
Shape X_test: (123, 3)
Shape y_train: (286,)
Shape y_test: (123,)
```

Gambar 3.2 6 Pembagian data set

- Shape X: (409, 3): Ini menunjukkan bahwa dataset X memiliki 409 sampel (data) dan masing-masing sampel memiliki 3 fitur (atribut). Fitur ini bisa berupa variabel numerik atau kategori yang digunakan untuk klasifikasi.
- Shape y: (409,): Ini menunjukkan bahwa ada 409 label (target) yang sesuai dengan setiap sampel di X. Label ini adalah nilai yang ingin diprediksi oleh model.
- Shape X_train: (286, 3): Dari total 409 sampel, 286 sampel digunakan untuk pelatihan model (training set). Ini adalah data yang digunakan untuk mengajarkan model tentang pola dalam data.
- Shape X_test: (123, 3): Sisa 123 sampel digunakan untuk pengujian model (test set). Data ini digunakan untuk mengevaluasi seberapa baik model dapat memprediksi label yang benar.
- Shape y_train: (286,): Ini adalah label yang sesuai dengan X_train. Ada 286 label yang digunakan untuk melatih model.
- Shape y_test: (123,): Ini adalah label yang sesuai dengan X_test. Ada 123 label yang digunakan untuk menguji model.

4. Mem-visualisasikan Confusion Matrix

Hasil visualisasi Confusion Matrix menunjukkan kinerja model dalam mengklasifikasikan sentimen ulasan pengguna. Confusion Matrix ini memberikan gambaran yang jelas tentang jumlah prediksi yang benar dan salah untuk setiap kelas sentimen. Dari hasil yang diperoleh, terlihat bahwa model berhasil mengklasifikasikan sentimen positif (rating 4.0 dan 5.0) dengan akurasi yang tinggi, sementara terdapat beberapa kesalahan klasifikasi pada sentimen negatif (rating 1.0 dan 2.0).



Gambar 3.2 7 Confusion Matrix

Kemudian untuk classification report dari model Naive Bayes dapat dilihat pada gambar berikut

	precision	recall	f1-score	support
0	0.78	0.22	0.34	32
1	0.78	0.98	0.87	91
accuracy			0.78	123
macro avg	0.78	0.60	0.60	123
weighted avg	0.78	0.78	0.73	123

Gambar 3.2 8 Classification report

5. K-Ford Cross Validation

Hasil dari K-Fold Cross Validation menunjukkan bahwa model yang diterapkan dalam analisis sentimen memiliki kinerja yang konsisten dan dapat diandalkan. Dengan membagi dataset menjadi K subset, melakukan evaluasi model secara menyeluruh, di mana setiap subset bergantian berfungsi sebagai data pengujian.

Naive Bayes

accuracy - 0.7626 ± 0.0524
precision - 0.7436 ± 0.0477
recall - 0.9895 ± 0.0211
f1 - 0.8479 ± 0.0286

SVM

accuracy - 0.8462 ± 0.0778
precision - 0.8469 ± 0.0748
recall - 0.9474 ± 0.0666
f1 - 0.8919 ± 0.0547

Logistic Regression

accuracy - 0.7974 ± 0.0615
precision - 0.7878 ± 0.0703
recall - 0.9684 ± 0.0537
f1 - 0.8653 ± 0.0361

KNN

accuracy - 0.6007 ± 0.0592
precision - 0.6749 ± 0.0391
recall - 0.7789 ± 0.0993
f1 - 0.7197 ± 0.0486

Random Forest

accuracy - 0.8425 ± 0.0457
precision - 0.8192 ± 0.0490
recall - 0.9842 ± 0.0337
f1 - 0.8937 ± 0.0324

Extra Trees

accuracy - 0.8565 ± 0.0488
precision - 0.8449 ± 0.0532
recall - 0.9684 ± 0.0482
f1 - 0.9004 ± 0.0302

Gambar 3.2 9 Hasil K-Ford Cross Validation

6. Memprediksi

```
Akurasi model SVC: 0.9106
Prediksi untuk data baru: [1 1 1 1]
```

Gambar 3.2 10 Prediksi

Model Support Vector Classifier (SVC) yang telah dilatih menunjukkan kinerja yang sangat baik dengan akurasi mencapai 91.06%. Ini berarti bahwa model ini berhasil memprediksi label yang benar untuk sekitar 91.06% dari total data yang diuji. Ketika diterapkan pada data baru, model ini memprediksi bahwa semua empat sampel yang diberikan termasuk dalam kelas 1, dengan hasil prediksi [1, 1, 1, 1]. Hal ini menunjukkan bahwa model memiliki kecenderungan untuk mengklasifikasikan data baru tersebut sebagai kelas 1, yang mungkin mencerminkan karakteristik fitur dari data baru yang lebih mirip dengan kelas tersebut.

Confusion Matrix:

```
[[25  7]
 [ 4 87]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.86	0.78	0.82	32
1	0.93	0.96	0.94	91
accuracy			0.91	123
macro avg	0.89	0.87	0.88	123
weighted avg	0.91	0.91	0.91	123

Gambar 3.2 11 confusion matrix(y_test, y_pred)

Interpretasi:

Kelas 0:

- True Negative (TN): 25 — Model memprediksi kelas 0 dan benar.
- False Positive (FP): 7 — Model memprediksi kelas 0 tetapi sebenarnya kelas 1.

Kelas 1:

- False Negative (FN): 4 — Model memprediksi kelas 0 tetapi sebenarnya kelas 1.
- True Positive (TP): 87 — Model memprediksi kelas 1 dan benar.

3.3 Dampak Analisis Sentimen terhadap Loyalitas Pelanggan

Dari hasil tersebut, analisis sentimen memainkan peran penting dalam membangun dan mempertahankan loyalitas pelanggan. Berikut beberapa dampak dari analisis yang memengaruhi loyalitas pelanggan:

- 1. Memahami sentimen pelanggan:** Analisis sentimen memungkinkan bisnis memahami bagaimana perasaan pelanggan terhadap produk dan layanan mereka.
- 2. Meningkatkan Pengalaman Pelanggan:** Dengan mengidentifikasi dan menyelesaikan masalah yang terungkap melalui analisis sentimen, bisnis dapat meningkatkan pengalaman pelanggan secara keseluruhan.
- 3. Meningkatkan loyalitas pelanggan:** Pelanggan yang merasa dihargai dan didengarkan cenderung lebih loyal. Analisis sentimen membantu mengidentifikasi masalah sebelum menjadi serius, memungkinkan bisnis mengambil langkah proaktif untuk mempertahankan pelanggan.
- 4. Membangun Hubungan yang Lebih Kuat:** Dengan merespons umpan balik pelanggan secara efektif, perusahaan dapat membangun hubungan yang lebih kuat dengan pelanggannya.. Hubungan positif ini menghasilkan loyalitas jangka panjang.
- 5. Meningkatkan Strategi Pemasaran:** Memahami emosi pelanggan memungkinkan perusahaan merancang kampanye pemasaran yang lebih relevan dan menarik sesuai harapan pelanggan serta meningkatkan keterlibatan dan loyalitas.

BAB IV PENUTUP

4.1 Kesimpulan

Penelitian ini memberikan wawasan yang berharga tentang persepsi pengguna terhadap aplikasi GoersApp di PT Sanraya Adi Nattaya. Dengan memahami faktor-faktor yang mempengaruhi kepuasan dan loyalitas pelanggan, perusahaan dapat merumuskan strategi yang lebih efektif untuk meningkatkan pengalaman pengguna dan mempertahankan pelanggan. Hasil analisis ini tidak hanya memberikan gambaran tentang apa yang dihargai oleh pengguna, tetapi juga mengungkapkan area yang perlu diperbaiki. Misalnya, jika banyak pengguna mengeluhkan masalah teknis atau kurangnya fitur tertentu, perusahaan dapat segera mengambil tindakan untuk mengatasi masalah tersebut. Dengan demikian, perusahaan tidak hanya dapat meningkatkan kepuasan pengguna saat ini, tetapi juga membangun loyalitas jangka panjang.

4.2 Saran

PT Sanraya Adi Nattaya sebaiknya fokus pada perbaikan kualitas aplikasi dengan mengaudit dan memperbaiki bug untuk meningkatkan performa. Menggunakan umpan balik pengguna untuk mengembangkan dan memperbarui fitur juga sangat penting. Selain itu, meningkatkan komunikasi dengan pengguna melalui respons cepat terhadap pertanyaan dan keluhan dapat memperkuat hubungan.

DAFTAR PUSTAKA

- Admin. (2022, February). *Analisis Sentimen (Sentiment Analysis) : Definisi, Tipe dan Cara Kerjanya*. Retrieved from lp2m.uma.ac.id: <https://lp2m.uma.ac.id/2022/02/21/analisis-sentimen-sentiment-analysis-definisi-tipe-dan-cara-kerjanya/>
- anonym. (2024, October). *How to use sentiment analysis to analyze customer opinion*. Retrieved from sproutsocial.com: <https://sproutsocial.com/insights/sentiment-analysis/>
- Hary Irvan Tarigan, Y., Manurung, Y. S., & Marpaung, W. (2019). Loyalitas pelanggan dan kualitas pelayanan pada pelanggan jasa transportasi online. *Philanthropy Journal of Psychology*, 3(1), 1-74.
- GOERS. (2021). *Web GOERS*. Retrieved from <https://www.goersapp.com/>
- Wati, R. (2023, April). *Pembobotan TF-IDF Menggunakan Naïve Bayes pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH*. Retrieved from <https://ojs.unikom.ac.id/index.php/jamika/article/view/9424>
- Widianto, M. H. (n.d.). *Algoritma Naive Bayes*. Retrieved from binus.ac.id: <https://binus.ac.id/bandung/2019/12/algoritma-naive-bayes/>
- wikipedia. (2024, September). *Sentiment analysis*. Retrieved from wikipedia.org: https://en.wikipedia.org/wiki/Sentiment_analysis
- Hary Irvan Tarigan, Y., Manurung, Y. S., & Marpaung, W. (2019). Loyalitas pelanggan dan kualitas pelayanan pada pelanggan jasa transportasi online. *Philanthropy Journal of Psychology*, 3(1), 1-74.
- Apriani, R., & Gustian, D. (2019). Analisis sentimen dengan Naïve Bayes terhadap komentar aplikasi Tokopedia. *Jurnal Rekayasa Teknologi Nusa Putra*, 6(1), 54-62.
- [Aponno, J. C. (2022). Penerapan algoritma sentiment analysis dan Naïve Bayes terhadap opini pengunjung di tempat wisata Pantai Pintu Kota, Kota Ambon. *Jurnal Teknik Informatika dan Sistem Informasi*, 9(4), 3180-3188

LAMPIRAN : DOKUMENTASI INTERNSHIP

