Journal of Computer and Information Systems Ampera
Vol. 7, No. 2, May 2026 e-ISSN: **2775-2496**
https://journal-computing.org/index.php/journal-cisa/index

ANALISIS SENTIMEN ULASAN APLIKASI OTA GOOGLE PLAYSTORE MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOUR

Hastri Cantya Danahiswari[1*], Eka Dyar Wahyuni[2], Amalia Anjani Afriyanti[3]
1,2,3 Sistem Informasi, Fakultas Ilmu Komputer, UPN "Veteran" Jawa Timur
E-mail: hastricantyad@gmail.com, ekawahyuni.si@upnjatim.ac.id, amalia_anjani.fik@upnjatim.ac.id

Abstrak

Penelitian ini bertujuan menganalisis sentimen ulasan pengguna aplikasi Online Travel Agent (OTA), yaitu Traveloka dan Tiket.com, yang diperoleh dari Google Play Store menggunakan algoritma K-Nearest Neighbour (KNN). Dataset penelitian terdiri atas 14.000 ulasan yang dikumpulkan melalui proses web scraping, kemudian dilakukan preprocessing dan pelabelan sentimen menggunakan RoBERTa dan VADER. Pemodelan dilakukan melalui beberapa skenario, yaitu data mentah, undersampling, dan oversampling, dengan pengujian nilai K dari 1 hingga 25. Hasil eksperimen menunjukkan bahwa skenario RoBERTa dengan oversampling pada K=1 menghasilkan akurasi tertinggi sebesar 97%, sedangkan K=6 memperoleh akurasi sebesar 92% dengan waktu komputasi 1,19 detik. Berdasarkan pertimbangan keseimbangan performa klasifikasi dan waktu komputasi, K=6 dipilih sebagai model final penelitian. Model tersebut selanjutnya diimplementasikan dalam sistem visualisasi berbasis web untuk melakukan prediksi sentimen ulasan, menampilkan distribusi sentimen dalam bentuk grafik, serta menghasilkan wordcloud. Pengujian sistem terhadap 50 ulasan menghasilkan 46 prediksi benar dan 4 prediksi salah, dengan akurasi sebesar 92%. Penelitian ini menunjukkan bahwa algoritma KNN dapat digunakan untuk analisis sentimen ulasan aplikasi OTA dan diimplementasikan dalam sistem visualisasi berbasis web.

Kata kunci: Analisis Sentimen, *Online Travel Agent, K-Nearest Neighbour*

1. PENDAHULUAN

Di era digital saat ini, berbagai macam aplikasi Online Travel Agent (OTA) di Google Play Store digunakan oleh masyarakat Indonesia untuk berbagai macam keperluan seperti berlibur atau travelling. Masing – masing aplikasi Online Travel Agent (OTA) memiliki produk dan fitur tersendiri yang membedakan antar aplikasi. Adapun, aplikasi OTA terpopuler di

This work is licensed under a Creative Commons Attribution 4.0 International License.

Indonesia, yaitu Traveloka dan Tiket.Com berdasarkan jumlah data pengunjung aplikasi tersebut.

Agar konsumen mendapatkan pelayanan terbaik terhadap performa aplikasinya maka, penting untuk meningkatkan pelayanan kepada pengguna saat memesan tiket pada aplikasi tersebut dengan fokus pada kebutuhan dan emosi pengguna ketika menggunakan aplikasi tersebut sebagai salah satu kunci keberhasilan sebuah aplikasi[1]. Bantuan teknologi informasi pada pemesanan tiket dapat meringkas yakni biaya dan waktu yang diperlukan saat pemesanan dengan tiket konvensional. Tiket online mengalami perkembangan dari tahun sebelumnya yakni kenaikan mencapai 10,4% menjadi $694.41 billion[2].

Selama ini sebuah predikat untuk aplikasi terbaik diberikan kepada aplikasi dengan jumlah unduhan dan pemberian bintang paling tinggi di Play Store[3]. Berdasarkan profil tiket.com dalam Google Play Store bahwa hingga akhir 2022 aplikasi tiket.com telah digunakan oleh lebih dari 10 juta pengguna dan mendapatkan rating 4.3 di Google Play Store[15]. Begitu juga dengan kompetitornya, Traveloka hingga akhir 2023 telah digunakan lebih dari 50 juta pengguna dan mendapatkan rating 4.8[16]. Data tersebut menunjukkan Tiket.com dan Traveloka merupakan aplikasi OTA yang cukup diminati. Namun, suatu aplikasi selalu memiliki kekurangan dan kelebihan masing-masing yang dapat menimbulkan berbagai respon dari pengguna aplikasi seperti kepuasan dan kekecewaan terhadap aplikasi tersebut[4].

Ulasan komentar terhadap aplikasi di Google Play menjadi salah satu tempat untuk melontarkan kepuasaan dan kekecewaan pengguna atau opini terhadap aplikasi tersebut. Maka dari itu, hal tersebut dapat dijadikan bahan analisis sentimen terhadap aplikasi tiket.com untuk membuktikan bahwa kolom ulasan komentar merupakan tempat para pengguna aplikasi dalam melontarkan kepuasan dan kekecewaan.

Pada aplikasi pemesanan tiket online seperti Traveloka dan tiket.com terdapat berbagai macam ulasan dari pengguna baik yang bersifat positif, negative, maupun netral. Adapun beberapa contoh penelitian mengenai analisis sentimen terhadap sebuah produk atau aplikasi yang dijadikan referensi dalam penelitian ini belum ada yang membahas mengenai analisis sentimen pada aplikasi android OTA (Online Travel Agent) [3][4] [5]. Maka dari itu, hasil dari penelitian ini diharapkan dapat membantu, mempermudah serta memberikan informasi tentang emosi yang terkandung dalam ulasan tentang aplikasi OTA yang ada pada situs Google Play Store.

Penelitian ini mengembangkan analisis sentimen ulasan aplikasi OTA dengan membandingkan metode pelabelan sentimen menggunakan RoBERTa dan VADER serta menerapkan teknik penyeimbangan dataset melalui undersampling dan oversampling. Selanjutnya, algoritma KNN digunakan untuk melakukan klasifikasi sentimen terhadap ulasan pengguna. Hasil klasifikasi diimplementasikan dalam sistem visualisasi berbasis web yang memungkinkan pengguna memperoleh informasi mengenai distribusi sentimen dan hasil analisis ulasan aplikasi OTA. Penelitian ini diharapkan dapat memberikan gambaran mengenai performa algoritma KNN dalam klasifikasi sentimen ulasan aplikasi OTA serta mendukung pemanfaatan ulasan pengguna sebagai bahan evaluasi terhadap kualitas layanan aplikasi.

2. METODOLOGI

Cross-Industry Standard Process for Data Mining (CRISP-DM) merupakan metodologi standar industri yang netral dalam proses data mining yang dikembangkan pada tahun 1996 oleh Daimler Chrysler di Eropa[6]. Pada metode ini terdiri dari 6 tahapan sebagai berikut.



Gambar 3 CRISP-DM

2.1 Business Understanding

Pemahaman bisnis merupakan tahapan awal dimana melakukan observasi terkait topik penelitian yaitu analisis sentimen pada ulasan komentar di Google Play Store dan menyusun rencana atau langkah untuk mencapai tujuan penelitian. Melakukan studi literatur terkait aplikasi Online Travel Agent dan menentukan ulasan aplikasi yang akan digunakan. Adapun, ulasan aplikasi Tiket.com dan Traveloka yang digunakan sebagai representasi aplikasi Online Travel Agent. Hal ini dipilih sesuai penelitian yang dilakukan oleh Katadata

pada tahun 2022 sektor aplikasi Online Travel Agent (OTA) 2 aplikasi teratas, yaitu Tiket.com dan Traveloka.

2.2 Data Understanding

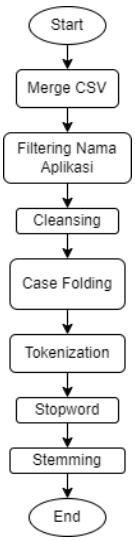
Setelah melakukan pemahaman bisnis Langkah selanjutnya, yaitu mengumpulkan ulasan komentar aplikasi di Google Play Store. Pada tahap ini dimulai tahap pengumpulan data ulasan komentar aplikasi tiket.com dimulai dari dengan total data sebanyak 7000 ulasan komentar Bahasa Indonesia dan pengguna dari Indonesia. Selanjutnya, dimulai tahap pengumpulan data ulasan komentar aplikasi Traveloka dengan total data sebanyak 7000 ulasan komentar Bahasa Indonesia dan pengguna dari Indonesia. Adapun, cara web scraping akan menggunakan tools Google Colab dengan Bahasa pemrograman Python dengan memanfaatkan link gits dari Google Play Store. Data yang telah di scrape akan di generate menjadi CSV (Comma Separated Value) yang dapat dibuka pada Microsoft Excel. Pada tahapan ini dilakukan pemahaman data dengan melakukan pengecekan pada kolom apa saja yang didapatkan dari hasil pengumpulan data (web scrapping) pada google play store. Selanjutnya, dilakukan pengecekan kolom yang telah didapatkan dari hasil web scrapping, yaitu UserName, Content, Score, reviewCreatedVersion, at (timestamp Review user), repliedAt (timestamp review developer), replyContent, appVersion.

2.3 Data Preparation

Setelah data sudah dikoleksi dan dilakukan pemahaman data, maka langkah selanjutnya tahap preprocessing. Tahap preprocessing adalah proses perubahan bentuk data yang terstruktur sembarang menjadi data yang terstruktur sesuai kebutuhan untuk proses dalam text mining[7]. Tahapan Preprocessing yang digunakan antara lain:

a. Merge CSV
   Setelah Scrapping didapatkan data berupa file csv untuk aplikasi Tiket.com dengan jumlah 7000 data dan ulasan aplikasi Traveloka sejumlah 7000, sehingga total jumlah data menjadi 14.000.
b. Filtering Nama Aplikasi
   Filtering ini dilakukan untuk memfilter nama aplikasi agar pada ulasan fokus pada kalimat sentimen dan tidak menimbulkan ambigu terutama pada kata “Tiket” karena nama aplikasinya “Tiket.com”.
c. Cleansing
   Cleansing merupakan tahap pembersihan teks ulasan dari elemen yang tidak relevan dalam proses analisis sentimen. Pada tahap ini, teks ulasan dibersihkan dari URL, tag HTML, karakter khusus, tanda baca, simbol, serta elemen lain yang tidak diperlukan. Selain itu, dilakukan pemeriksaan terhadap data kosong dan duplikasi ulasan untuk menjaga

kualitas dataset. Proses cleansing bertujuan menghasilkan teks ulasan yang lebih bersih dan terstruktur sehingga dapat diproses pada tahapan preprocessing berikutnya.



Gambar 4. Alur Preprocessing Data

d. Case Folding
Case folding dilakukan dengan mengubah semua karakter teks menjadi huruf kecil. Hal ini dilakukan untuk mengurangi kompleksitas dan perbedaan kapitalisasi teks.
e. Tokenization
Tokenizing dilakukan dengan pemisahan teks menjadi kata, frasa, simbol menjadi elemen yang berbeda biasanya proses pemisahan setiap kata dipisahkan dengan tanda koma (,).
f. Stopword
Stop Removal digunakan untuk menghilangkan kata umum yang tidak memiliki arti penting dan tidak digunakan dengan tujuan mengurangi jumlah kata yang disimpan oleh sistem[8].
g. Stemming
Stemming dilakukan dengan menyatukan atau mengubah varian bentuk kata dari umum menjadi kata dasar[9].

Tahapan – tahapan diatas merupakan tahapan penting sebelum labelling. Dimana langkah selanjutnya, akan dilakukan labelling menggunakan library vader dan Roberta. Labelling menggunakan library Roberta dilakukan langsung

dengan memanggil library. Hal ini berbeda dengan library vader karena, library vader tidak menyediakan library untuk Bahasa Indonesia, sehingga perlu diterjemahkan dengan menggunakan library translator. Setelah dilakukan penerjemahan maka, library vader akan dipanggil. Hasil dari library Roberta dan Vader tidak jauh berbeda dimana keduanya menghasilkan labelling positif dan negatif dari ulasan.

2.4 Pemodelan (Modelling)

Sebelum dilakukannya pemodelan, dari hasil data yang telah melewati preprocessing pada tahapan data preparation akan dilakukan pengecekan EDA. Hal ini diperlukan untuk mengetahui apakah terjadi imbalanced data. Apabila terjadinya imbalanced data, maka dilakukan skenario undersampling dan oversampling agar dataset menjadi seimbang[10]. Oversampling memiliki tujuan untuk meningkatkan sampel kelas minoritas sampai sama dengan kelas mayoritas lain dengan menduplikasi secara acak sampel kelas minoritas[11]. Sedangkan, random undersampling yaitu menghasilkan sub sampel acak dari instance kelas mayoritas. Undersampling merupakan metode sampling secara acak memilih sampel di kelas mayoritas dan menambahkannya ke kelas minoritas, membentuk sebuah dataset pelatihan baru[11]. Kemudian, langkah selanjutnya yaitu melakukan perhitungan nilai perbandingan antara nilai data yang sudah dilakukan preprocessing dan sebelum dilakukan preprocessing, terdiri dari beberapa tahapan yaitu pemecahan dataset menggunakan split data yang terdiri dari data training dan data testing dengan nilai perbandingan 75:25 dengan menerapkan algoritma k-nearest neighbors sebagai nilai perbandingan mana hasil nilai akurasi tertinggi dengan pengecekan nilai k yang beragam.

2.5 Evaluasi (Evaluation)

Pada tahapan ini peneliti melakukan evaluasi terhadap hasil perhitungan algoritma k-nearest neighbors yang dilakukan terhadap ulasan komentar yang sudah melewati proses data preparation dan data preprocessing dengan menggunakan confusion matrix yang digunakan untuk menghitung nilai akurasi dari proses klasifikasi[14]. Evaluasi menggunakan confusion matrix dilakukan untuk mengetahui perkiraan yang benar dan salah dari hasil klasifikasi. Evaluasi digunakan untuk mengetahui keakuratan model yang sudah dibuat. Hasil akurasi tersebut didapatkan dengan membandingkan data yang sudah di beri label benar atau salah pada model dengan hasil prediksi label. Tahap pengujian dilakukan dengan menggunakan data uji. Evaluasi model klasifikasi dianalisis dengan membandingkan nilai akurasi secara keseluruhan dan waktu yang diperlukan untuk melakukan training model. Lalu, dilakukannya pengecekan apakah dengan nilai k yang telah dimasukkan pada

pemodelan, dataset mengalami overfit atau underfit dengan melihat grafik evaluasi K-Nearest Neighbour.

2.6 Deployment (Visualisasi Web)

Setelah hasil dari tingkat akurasi kedua metode klasifikasi didapat maka langkah selanjutnya yaitu visualisasi[12]. Proses visualisasi disini berarti proses pengkodean dan pembuatan tampilan implementasi antar muka agar para pengguna dapat memahami hasil dari klasifikasi ini. Hasil dari klasifikasi data dengan kedua metode tersebut akan divisualisasikan ke dalam bentuk grafik yang mudah dipahami dan dari klasifikasi juga ditampilkan kata yang menjadi pemicu sentimen positif, negative, dan netral dalam beropini yang akan diimplementasikan dalam suatu aplikasi berbasis web menggunakan Bahasa PHP. Fitur dari visualisasi yaitu website bisa memprediksi label dari ulasan yang dimasukkan ke dalam inputan dan menampilkan line graphic berdasarkan tren tiap bulan dengan hasil label ulasan.

3 HASIL DAN PEMBAHASAN

3.1 Data Understanding

Data yang diambil merupakan ulasan terbaru dengan masing – masing aplikasi diambil 7000 data dengan kriteria ulasan oleh akun dengan negara Indonesia dan bahasa Indonesia. Kolom yang diambil terdiri dari username, content, score, reviewCreatedVersion, at, replyContent, repliedAt, dan appVersion.

3.2 Data Preparation

a. Merge CSV

Setelah Scrapping didapatkan data berupa file csv untuk aplikasi Traveloka dengan jumlah 7000 data dan ulasan aplikasi Traveloka sejumlah 7000. Sehingga, dataset yang digunakan ada 14000.

b. Filtering Kata Traveloka dan Tiket.com

Filtering ini dilakukan untuk memfilter kata “Traveloka”, “Traveloka”, “TRAVELOKA”, “Tiket.com”, “tiket.Com”, “tiket.com”, “TIKET.COM”, “Tiketcom”, “tiketcom” agar pada ulasan fokus pada kalimat sentimen dan tidak menimbulkan ambigu terutama pada kata “Tiket” karena nama aplikasinya “Tiket.com”.

c. Cleansing

- Menghilangkan URL situs lain atau URL gambar dan video Untuk menghilangkan URL situs lain, gambar, dan video dari teks.
- Menghilangkan tanda baca dan angka.
- Menghilangkan spasi di depan dan di belakang untuk kalimat yang repetisi.

d. Tokenization
Pertama, library yang diperlukan seperti pandas, nltk, dan TweetTokenizer dari NLTK diimport. Selanjutnya, Google Drive di-mount untuk mengakses file CSV yang akan diproses. File CSV dibaca menggunakan pandas dan disimpan dalam DataFrame yang disebut df.

e. Stopword
Stop Removal digunakan untuk menghilangkan kata umum yang tidak memiliki arti penting dan tidak digunakan dengan tujuan mengurangi jumlah kata yang disimpan oleh sistem

f. Stemming
Stemming dilakukan dengan menyatukan atau mengubah varian bentuk kata dari umum menjadi kata dasar

h. Labelling
Labelling dilakukan dengan menggunakan library Roberta dan Vader. Untuk Roberta dapat digunakan langsung untuk dataset berbahasa Indonesia. Sedangkan, untuk Vader perlu ditranslate terlebih dahulu karena Vader hanya dapat digunakan untuk dataset berbahasa Inggris.

Berikut merupakan perbandingan hasil labelling library Roberta dan Vader dilakukan untuk validasi dan ditemukan hasilnya tidak jauh berbeda.

Berikut merupakan tabel perbandingan labelling:

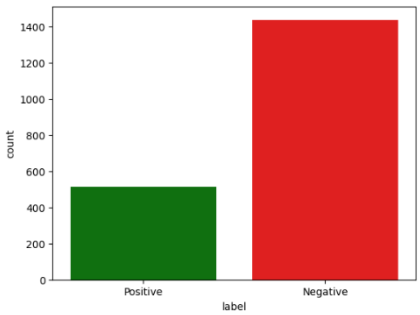
Tabel 1 Perbandingan Hasil Labelling Library Roberta dan Vader

No.	Ulasan	Labelling Vader	Labelling Roberta	Labelling Manual
1.	Sangat bantu diskon harga pantastis	Positive	Positive	Positive
2.	Moga hasil kecewa mesan orang tua	Negative	Negative	Negative
3.	sangat kecewa beli tiket dgn paylater reschedule tglnya dr juli juli ibu sakit keras indonesia bel paylater lg kena tagih dobel tiket yg baru ga masalah bayar yg sdh tiket khan sdh rubah sdngkan biaya reschedule rb orang tp say tetap harus bayar paylater	Positive	Negative	Negative

	yg sdh rubah jadwal khan aneh jadi asa beli tiket jt kalu sya byar dobel sejak malam sampe skrg no tanggap			
4.	lmyn gampang ringkes utk aplikasi nya	Negative	Positive	Positive
5.	bagus cepat layan nya tiket cari	Negative	Positive	Positive

Dari tabel diatas dapat dilihat bahwa hasil dari library roberta dan vader tidak jauh berbeda dan dilakukan untuk validasi.

i. EDA (Explatory Data Analysis)
Pada penelitian ini akan dilakukan beberapa skenario data ulasan karena setelah dilakukan preprocessing dan dilakukan pengecekan jumlah perbandingan data positive dan negative didapat hasil berupa:



Gambar 1 Perbandingan data label positive dan negative

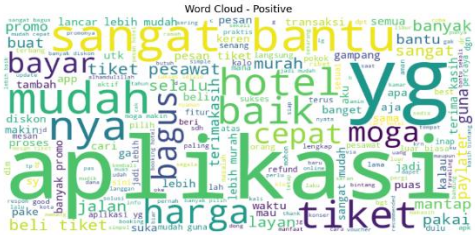
Pada grafik tersebut dapat terlihat bahwa data ulasan yang telah melewati preprocessing data mengalami imbalanced dataset (tidak seimbang). Maka dari itu, perlu dilakukan penyeimbangan data seperti, undersampling dan oversampling.
Adapun dari data tersebut dilakukan pengecekan menggunakan word cloud untuk mengetahui kata yang sering muncul pada ulasan OTA.

Journal of Computer and Information Systems Ampera

Vol. 7, No. 2, May 2026 e-ISSN: **2775-2496**

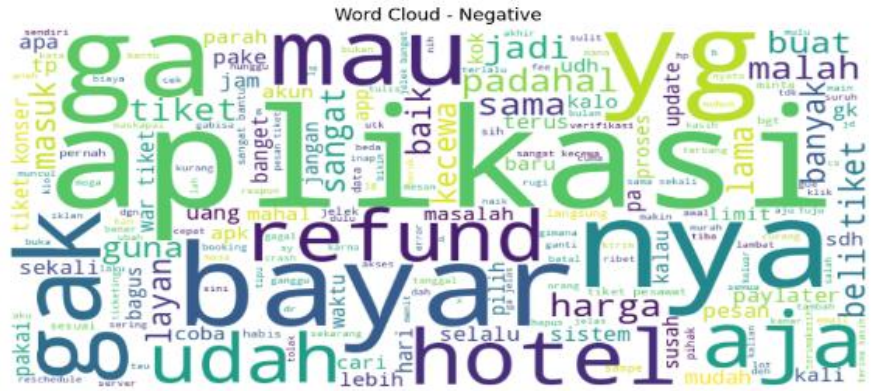
https://journal-computing.org/index.php/journal-cisa/index





Gambar 2 Word Cloud Ulasan OTA Gambar 3 Word Cloud

PositivUlasan OTA



Gambar 4. Word Cloud Negative Ulasan OTA

3.2 Pemodelan (Modelling)

a. Hold Out

Langkah selanjutnya setelah melakukan skenario untuk balancing data, yaitu hold out. Hold Out dilakukan untuk menguji model yang akan digunakan. Pada metode ini, dataset yang dimiliki akan dibagi menjadi dua bagian yaitu data latih (training) dan data uji (testing). Terdapat berbagai perbandingan seperti perbandingan 70:30, 75:25, dan 80:20. Pada penelitian ini digunakan perbandingan 75:25. Berikut merupakan contoh source code hold out:

```
# Assuming your CSV has a 'text' column for input text and a 'label' column for labels
X = data['Content']
y = data['label']

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25, random_state=42)

# TF-IDF vectorization
tfidf_vectorizer = TfidfVectorizer(max_features=1000)  # You can adjust max_features as needed
X_train_tfidf = tfidf_vectorizer.fit_transform(X_train)
X_test_tfidf = tfidf_vectorizer.transform(X_test)
```

Gambar 5 Source Code Hold Out

b. Pemodelan Dengan Algoritma K-Nearest Neighbour

Langkah selanjutnya setelah melakukan hold out untuk menguji data latih, yaitu pemodelan dengan algoritma K-Nearest Neighbour. Pada metode ini, dilakukan pengujian pemodelan KNN nilai k=1 hingga k=25 dengan menggunakan skenario data mentah, data labelling Vader, dan data labelling Roberta. Berikut merupakan hasil pengujian nilai akurasi dengan algoritma KNN.

No.	Nilai K	Roberta Under Sampling	Vader Under Sampling	Data Mentah (Imbalanced)	Roberta Over Sampling	Vader Over Sampling
1.	1	0.64	0.78	0.77	0.97	0.77
2.	2	0.60	0.65	0.74	0.87	0.74
3.	3	0.50	0.63	0.75	0.94	0.75
4.	4	0.49	0.74	0.75	0.91	0.73
5.	5	0.49	0.71	0.75	0.82	0.72
6.	6	0.49	0.75	0.74	0.92	0.71
7.	7	0.48	0.70	0.74	0.72	0.71
8.	8	0.69	0.75	0.74	0.72	0.71
9.	9	0.65	0.72	0.74	0.49	0.71
10.	10	0.74	0.76	0.74	0.80	0.71
11.	11	0.72	0.74	0.74	0.79	0.71
12.	12	0.75	0.78	0.74	0.80	0.71
13.	13	0.74	0.76	0.74	0.76	0.71
14.	14	0.77	0.76	0.75	0.79	0.71
15.	15	0.75	0.76	0.74	0.76	0.73
16.	16	0.77	0.77	0.75	0.81	0.72
17.	17	0.77	0.76	0.74	0.79	0.74
18.	18	0.79	0.75	0.65	0.80	0.73
19.	19	0.77	0.75	0.77	0.78	0.74
20.	20	0.78	0.75	0.73	0.80	0.74

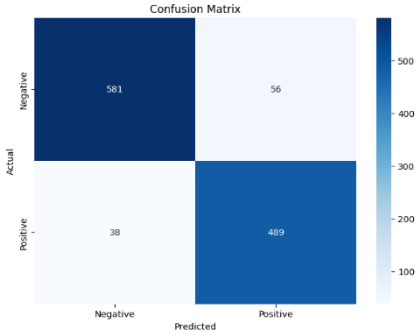
21.	21	0.78	0.75	0.74	0.78	0.76
22.	22	0.79	0.77	0.75	0.80	0.75
23.	23	0.79	0.75	0.74	0.80	0.77
24.	24	0.79	0.76	0.74	0.82	0.76
25.	25	0.80	0.75	0.78	0.80	0.78

Tabel 2 Hasil Pengujian dan Perbandingan Skenario Ulasan

Dari hasil pengujian dan perbandingan diatas, dapat dilihat bahwa nilai akurasi terbaik diantara skenario balancing data, yaitu Roberta oversampling dengan nilai k=1 dan akurasi tertinggi, yaitu 97%. Namun, setelah dilakukan evaluasi pada grafik menunjukkan nilai k=6 dengan akurasi 92% yang optimal pada studi kasus. Terlihat bahwa nilai k=6 tidak mengalami underfit maupun overfit dan menemukan titik temu dari grafik testing dan training. Selain itu, pada process time, nilai k=6 mendapat nilai tercepat dibandingkan dengan nilai k lainnya. Sedangkan, pada tabel lainnya, nilai akurasi yang berwarna kuning merupakan akurasi tertinggi dari masing - masing scenario.

3.4 Evaluasi

Langkah selanjutnya, yaitu confusion matrix. Confusion Matrix atau juga bisa disebut matriks klasifikasi merupakan metode perhitungan yang biasa digunakan dalam supervised learning[13].
Dilakukan perhitungan confusion matrix dengan hasil sebagai berikut:

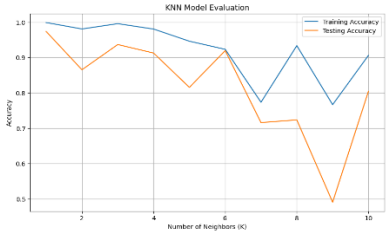


Gambar 7 Confusion Matrix

Dengan jumlah data 4653 dengan data testing yang digunakan 25%, sehingga hasil perhitungan confusion matrix didapatkan nilai:

a. TP (True Positive) dimana prediksi benar positif dengan jumlah 581
b. FP (False Positive) dimana prediksi salah positif dengan jumlah 56
c. FN (False Negative) dimana prediksi salah negative dengan jumlah 38
d. TN (True Negative) dimana prediksi benar negative dengan jumlah 489

Dengan memasukkan range nilai k untuk menampilkan nilai k neighbours untuk menghitung train accuracy dengan test accuracy. Lalu, pembentukan grafik training accuracy dan testing accuracy, sehingga dapat mengetahui nilai k=6 adalah yang menemukan titik temu antara training accuracy dan testing accuracy.
Berikut merupakan hasil berbentuk grafik evaluasi nilai k=6 pada model KNN yang telah dibuat:



Gambar 8 Grafik Evaluasi Model KNN

Pada grafik tersebut dapat terlihat bahwa nilai k=6 akurasi training dan akurasi testingnya menemukan titik Tengah dimana nilai akurasinya = 92%. Adapun, pada grafik tersebut dapat terlihat bahwa nilai k=1 data trainnya memiliki akurasi = 100% namun, testing akurasinya memiliki akurasi = 97%. Maka dari itu, nilai k yang digunakan pada penelitian ini, yaitu k=6.
Untuk informasi lebih rinci mengenai report hasil klasifikasi dari setiap skenario dapat dilihat dari tabel berikut.

Percobaan nilai K		Akurasi	Waktu Proses	Presisi	Recall	F1-Score	Support
k=1	Negative	97%	1.97 s	96%	100%	98%	637
	Positive			99%	95%	97%	527
k=5	Negative	82%	1.42 s	93%	72%	81%	637
	Positive			73%	93%	82%	527
k=6	Negative	92%	1.19 s	94%	91%	93%	637
	Positive			90%	93%	91%	527
k=7	Negative	72%	1.94 s	93%	52%	67%	637
	Positive			62%	95%	75%	527

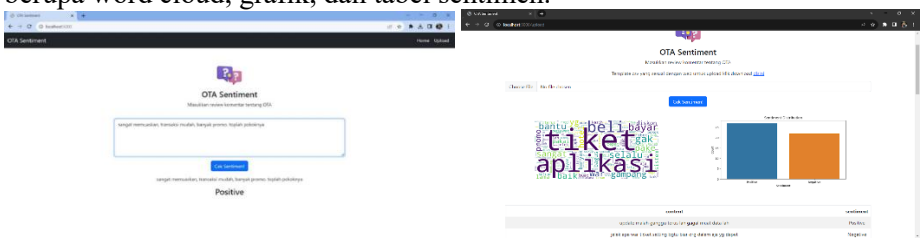
Tabel 3 Hasil Percobaan Nilai K Skenario Oversampling

Dari tabel berikut, telah dilakukan percobaan nilai k = 1, 5, 6, dan 7. Adapun, hasil yang didapatkan, yaitu k = 1 dengan nilai akurasi tertinggi 97% dan waktu proses 1.97 detik. Lalu, k=5 dengan nilai akurasi 82% dan waktu proses 1.42 detik. Selanjutnya k=6 dengan nilai akurasi 92% dan waktu proses tercepat 1.19 detik. Dan k=7 dengan nilai akurasi 72% dan waktu proses 1.94 detik. Dari hasil tersebut, nilai k memiliki akurasi tertinggi namun menunjukkan similarity dengan 1 instance lain dan memiliki waktu proses yang paling lama, yaitu 1.97 detik. Lalu, akurasi terbaik lainnya yaitu k=6 dengan nilai akurasi 92% dengan waktu proses tercepat, yaitu 1.19 detik. Nilai k=5 dan 7 menunjukkan akurasi tidak lebih baik dari k=6. Maka dari itu, pada penelitian ini digunakan nilai k=6

sebagai nilai k yang memiliki akurasi yang tinggi dan waktu proses yang tercepat.

3.5 Deployment (Web)

Langkah selanjutnya, yaitu deployment web. Tahapan ini merupakan pengaplikasian dari analisis sentiment dalam bentuk web. Pada web ini terdapat beberapa fitur, seperti input ulasan per kalimat, input ulasan menggunakan csv sesuai template, menampilkan hasil sentimen, menampilkan hasil sentimen berupa word cloud, grafik, dan tabel sentimen.



Gambar 9 Tampilan Web

4. KESIMPULAN DAN SARAN

Berikut adalah ringkasan tahapan membangun model analisis sentimen ulasan aplikasi online travel agent di google play store:

1. Tahapan membangun model analisis sentimen adalah studi literatur, analisis kebutuhan, perancangan model yang meliputi pengumpulan data, penyaringan data, text preprocessing, EDA, pelabelan data, klasifikasi model, pembagian data hold out, evaluasi model klasifikasi, dan visualisasi.
2. Dalam pengembangan model klasifikasi menggunakan algoritma KNN (K-Nearest Neighbour) menggunakan berbagai macam skenario. Skenario yang dibuat terdiri dari penggunaan library Vader dan Roberta untuk pelabelan data. Lalu, terdapat pengecekan EDA dimana data yang didapatkan imbalanced, sehingga perlu dilakukan skenario pengecekan hasil akurasi undersampling dan oversampling. Adapun, pengecekan skenario nilai tetangga k=1 hingga k =25 untuk kedua skenario tersebut. Hasil penelitian menunjukkan bahwa akurasi tertinggi diperoleh oleh model klasifikasi menggunakan library Roberta dengan metode oversampling dengan model KNN nilai k=6, yaitu 92% dan process time 1,14 detik. Model library Roberta oversampling k=6 juga memberikan nilai F-1 Score yang baik, yaitu

untuk kelas negatif 93% dan untuk kelas positif, 91%. Evaluasi confusion matrix menunjukan tingkat kebenaran prediksi yang tinggi, yaitu 94% untuk kelas negatif dan 90% untuk kelas positif. Oleh karena itu, model klasifikasi menggunakan menggunakan library Roberta dengan metode oversampling nilai k=6 dipilih untuk diterapkan pada proses visualisasi.

3. Proses visualisasi ulasan aplikasi OTA dilakukan dengan menyajikan hasil sentimen, grafik untuk setiap kelas sentimen, wordcloud pada dataset yang diupload, dan tabel kelas sentimen. Pada proses visualisasi dibuat dengan menggunakan html (web). Dengan mengimport model yang telah dibuat berupa berkas (.pkl) lalu mengimport flask dan dimasukkan ke dalam html.

Berikut beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya:

1. Memperluas kemampuan sistem untuk mengatasi ulasan dengan kata-kata singkatan dan kesalahan pengejaan (typo), sehingga dapat meningkatkan akurasi dalam mengolah data yang tidak standar.
2. Menambahkan kata tidak baku seperti “gak” saat preprocessing, sehingga dapat meningkatkan akurasi dalam pemberian label sentiment.
3. Menambah kelas sentimen netral ketika library sudah menyediakannya di kemudian hari.
4. Menyempurnakan fitur yang dikeluhkan oleh pengguna pada aplikasi OTA dan menjaga fitur yang telah baik oleh developer aplikasi OTA. Berdasarkan data penelitian, kata sentiment negatif ulasan yang sering muncul, yaitu refund, bayar, war, dan limit paylater. Maka dari itu, pihak developer perlu melakukan improvisasi terkait fitur tersebut. Sedangkan, kata sentiment positif ulasan yang sering muncul, yaitu kemudahan, kelancaran, kecepatan, promo, dan transaksi, sehingga fitur – fitur terkait perlu dijaga agar pengguna puas dan terbantu saat menggunakan aplikasi OTA.

Dengan mengimplementasikan saran-saran di atas, diharapkan penelitian ini dapat memberikan kontribusi yang lebih besar dalam analisis sentimen pada ulasan aplikasi OTA baik di Google Play Store atau App store, serta meningkatkan kualitas dan keakuratan model klasifikasi yang digunakan.

6. DAFTAR RUJUKAN

[1] Maharani, L. (2022). Evaluasi User Experience pada Aplikasi Pemesanan Tiket Pesawat Traveloka dan Tiket.com Menggunakan User Experience Questionnaire (UEQ). *Universitas Islam Negeri*

Syarif Hidayatullah Jakarta. https://repository.uinjkt.ac.id/dspace/handle/123456789/65499

[2] Emarketer. (2018). *Global Digital Travel Sales 2018 Ease of Mobile Payment Options Boost Online Bookings*.

[3] Erfina, A., Basryah, E. S., Saepulrohman, A., & Lestari, D. (2020). Analisis Sentimen Aplikasi Pembelajaran Online di Play Store pada Masa Pandemi Covid-19 Menggunakan Algoritma Support Vector Machine. Seminar Nasional Informatika (SEMNASIF), 1(1), 145-152.

[4] Giovani, A. P., Ardiansyah, A., Haryanti, T., Kurniawati, L., & Gata, W. (2020). Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. Jurnal Teknoinfo, 14(2), 115. https://doi.org/10.33365/jti.v14i2.679

[5] Ernawati, S., & Wati, R. (2018). Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Review Agen Travel. *Jurnal Khatulistiwa Informatika*, *6*(1), 64–69.

[6] M. Bramer, *Principles of Data Mining*, Springer, 2007.

[7] Takdirillah, R. (2020). Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan. Edumatic : Jurnal Pendidikan Informatika, 4(1), 37–46. https://doi.org/10.29408/edumatic.v4i1.2081

[8] Sugianto, C. A., & Apandi, T. H. (2018). Pengaruh Tokenisasi Kata N-Grams Spam SMS Menggunakan Support Vector Machine [Preprint]. INA-Rxiv. https://doi.org/10.31227/osf.io/vjc7k

[9] Kannan & Gurusamy. (2015). Preprocessing Techniques for Text Mining. https://www.researchgate.net/publication/273127322_Preprocessing_Techniques_for_Text_Mining

[10] Cici, H., & Hwan, S. Y. (2023). *Mitigating Class Imbalance in Sentiment Analysis through GPT-3-Generated Synthetic Sentences.* https://doi.org/10.3390/app13179766

[11] He, H., Zhang, W., & Zhang, S. (2018). *A novel ensemble method for credit scoring: adaption of different imbalance ratios. Expert Systems With Applications*. doi:10.1016/j.eswa.2018.01.012

[12] Ilmawan, L. B., & Winarko, E. (2015). Aplikasi Mobile untuk Analisis Sentimen pada Google Play. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *9*(1), 53. https://doi.org/10.22146/ijccs.6640

[13] Manning, C, Raghavan, P, dan Schutze, H. 2009. Introduction to Information Retrieval. Cambridge: Cambridge University Press. Dapat diunduh dari: https://nlp.stanford.edu/IR-book /pdf/irbookon linereading.pdf.

[14] Obviously.ai. (2022). Cara Mengetahui Apakah Model Machine Learning Anda Memiliki Performa Baik. https://www.obviously.ai/post/machine-learning model-performance
[15] Tiket.com. (2023). Tiket.com – Tentang Kami, https://www.tiket.com/info/about us.
[16] Traveloka. (2021). Traveloka – Tentang Kami. https://www.traveloka.com/id id/about-us.