

Implementasi Algoritma Naïve Bayes Terhadap Review Aplikasi KFCKU

¹Angga Aditya Permana, ²Rohmat Taufiq, ³Muhammad Ibnu Wijaya

¹Prodi Informatika, Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara,
Jl. Scientia Boulevard, Gading Serpong, Kel. Curug Sangereng, Kec. Kelapa Dua, Kab.
Tangerang, Banten

^{2,3}Prodi Teknik Informatika, Fakultas Teknik, Univeersitas Muhammadiyah Tangerang,
Jl. Perintis Kemerdekaan 1 No 33, Kota Tangerang, Banten
e-mail: angga.permana@umn.ac.id

Receive: 29 November 2023

Accepted: 27 Desember 2023

Abstract

Sentiment analysis or sentiment analysis in Indonesian is a technique or method used to identify how feelings are expressed through text and how these feelings can be classified as positive or negative. One fast food company that often receives sentiment analysis is KFC (Kentucky Fried Chicken). To view sentiment analysis, this research took data sources from the Google Play Store. In this research, the Naïve Bayes method was also used which aims to classify data and increase the level of accuracy of the classification method. This research aims to analyze consumer sentiment towards KFC fast food restaurants using a Naïve Bayes approach based on data from the Google Play Store platform. Data is taken from conversations and reviews of application users related to KFC over a certain period of time. The Naïve Bayes method is used to classify consumer sentiment into three categories, namely positive, negative and neutral. The results of the sentiment analysis are then evaluated to assess the level of consumer satisfaction and perception towards KFC. This research can provide valuable information for KFC management to improve the quality of their services and products based on consumer feedback.

Keywords: KFCKU, Naïve Bayes, Sentiment Analysis.

Abstrak

Analisis sentimen atau *Sentimen analysis* dalam bahasa Indonesia adalah teknik atau metode yang digunakan untuk identifikasi bagaimana perasaan diungkapkan melalui teks dan bagaimana Perasaan ini dapat diklasifikasikan sebagai positif atau negatif. Salah satu perusahaan makanan cepat saji yang sering mendapatkan sentimen analisis yaitu KFC (*Kentucky Fried Chicken*). Untuk melihat sentimen analisis penelitian ini mengambil sumber data dari Google Play Store. Pada penelitian ini juga digunakan metode *Naïve Bayes* yang bertujuan untuk mengklasifikasikan data dan meningkatkan tingkat akurasi dari metode klasifikasi. Penelitian ini bertujuan untuk menganalisis sentimen konsumen terhadap restoran cepat saji KFC menggunakan pendekatan *Naïve Bayes* berbasis data dari platform Google Play Store. Data diambil dari percakapan dan ulasan pengguna Aplikasi terkait KFCKU dalam kurun waktu tertentu. Metode *Naïve Bayes* digunakan untuk mengklasifikasikan sentimen konsumen menjadi tiga kategori, yaitu positif, negatif, dan netral. Hasil analisis sentimen kemudian dievaluasi untuk menilai tingkat kepuasan dan persepsi konsumen terhadap KFC. Penelitian ini dapat memberikan informasi berharga bagi manajemen KFC untuk meningkatkan kualitas pelayanan dan produk mereka berdasarkan umpan balik konsumen.

Kata Kunci: KFCKU, Naïve Bayes, Sentimen Analisis.

PENDAHULUAN

Perkembangan wisata kuliner telah menjadi daya tarik utama bagi wisatawan yang menjelajahi berbagai destinasi, baik di dalam maupun di luar negeri. Memanjakan lidah dengan kekayaan cita rasa lokal menjadi pengalaman tak terlupakan, yang mampu menggambarkan keunikan dan identitas suatu daerah. Saat ini, dengan kemajuan teknologi dan tren digital, eksplorasi kuliner semakin terbuka melalui aplikasi dan situs ulasan, yang memainkan peran penting dalam membimbing para pengunjung menuju pengalaman kuliner yang memuaskan. (Santoso & Hartono, 2022)

Dalam konteks ini, restoran KFC telah menjadi salah satu destinasi kuliner yang populer di kalangan masyarakat. Sebagai bagian dari industri kuliner yang kompetitif, KFC tidak hanya dihadapkan pada persaingan dengan restoran sejenis, tetapi juga harus memastikan kepuasan pelanggan untuk mempertahankan reputasi dan daya tariknya. Dalam menjawab tantangan ini, penelitian ini mengambil pendekatan inovatif dengan mengimplementasikan algoritma Naïve Bayes untuk menganalisis sentiment (Andrianto & Adinugroho, 2018; PERMANA & NOVIYANTO, 2023; Rosadi et al., 2021) dari ulasan pelanggan terhadap aplikasi KFCKU.

Pentingnya ulasan pelanggan dalam mempengaruhi minat beli dan citra restoran membuat pemahaman mendalam terhadap sentimen konsumen menjadi krusial. Berbagai metode analisis sentimen telah diterapkan, namun penelitian ini memilih Naïve Bayes Classifier karena kemampuannya dalam melakukan klasifikasi dengan jumlah data training yang relatif kecil. (Damayunita et al., 2022; Metode et al., 2023; Rusli et al., 2020; Sinaga, 2017) dan hal ini memberikan landasan kuat bagi pemilihan metode ini dalam kerangka penelitian ini.

Tujuan utama penelitian ini adalah untuk memahami sentimen yang terkandung dalam ulasan pelanggan terhadap aplikasi KFCKU. Fokus utama mencakup segala aspek yang berkaitan dengan makanan, minuman, dan pengalaman keseluruhan di restoran KFC. Analisis sentimen yang dilakukan konsumen dan masyarakat pecinta kuliner akan memberikan wawasan berharga untuk meningkatkan kualitas layanan, serta memberikan masukan konstruktif kepada KFC untuk terus memperbaiki dan mengembangkan aplikasinya.

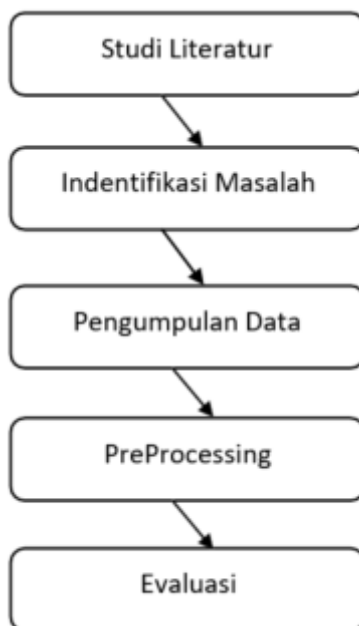
Dengan demikian, melalui implementasi algoritma Naïve Bayes pada analisis sentimen, penelitian ini tidak hanya memberikan kontribusi terhadap pemahaman sentimen pelanggan terhadap aplikasi KFCKU, tetapi juga menjadi langkah inovatif dalam pengembangan metode analisis sentimen di industri kuliner.

METODE PENELITIAN

Metode yang digunakan pada penelitian ini adalah dengan pengambilan data sampel yang berjumlah sebanyak 1520 data yang diperoleh melalui situs jejaring sosial pada akun Google Play Store @KFCKU yang bertujuan untuk mengetahui sentimen analisis positif dan negatif dari pengguna aplikasi @KFCKU mengenai kualitas pelayanan dan tingkat kepuasan konsumen pada perusahaan makanan cepat saji KFC ini.

Tahapan Penelitian

Pada penelitian dibutuhkan perencanaan dan langkah-langkah agar penelitian dapat berjalan dengan baik. Tahapan penelitian yang dilakukan dapat dilihat seperti gambar di bawah berikut:



Gambar 1. Metode Penelitian

A. Studi Literatur

Pada penelitian ini melakukan pencarian pada referensi teori yang paling relevan menggunakan analisis sentimen untuk mengetahui gambaran umum mengenai penelitian yang akan dilakukan. referensi ini dapat dicari dari jurnal, buku, artikel laporan penelitian, dan situs-situs pada internet. Hasil dari studi literatur ini ialah terkoleksinya referensi yang relevan dengan rumusan masalah.

B. Identifikasi Masalah

Identifikasi Masalah dilakukan untuk menganalisis sentiment yang terdapat pada aplikasi KFCKU di ulasan Google play store.

C. Pengumpulan Data

Data yang di gunakan untuk penelitian ini adalah data yang di ambil/ dikumpulkan dari ulasan aplikasi KFCKU yang terpadat pada google play store.

D. PreProcessing

Pre-processing dilakukan untuk menghasilkan data berkualitas sehingga dapat menghasilkan keputusan yang baik. Pada proses ini data yang sudah didapat akan dilakukan tahapan preprocessing untuk membersihkan data serta menghilangkan karakter yang mengganggu. Gangguan tersebut dapat mengakibatkan pembobotan tidak berjalan dengan tepat, ketika pembobotan tidak berjalan tepat maka data akan sulit untuk diklasifikasikan. Pada penelitian ini terdapat beberapa tahapan dalam text preprocessing sebagai berikut: (Fairuz et al., 2021; Oliveira et al., 2022)

- o Pelabelan Data

Pelabela data merupakan proses mengidentifikasi data mentah (gambar, file teks, video, dll). Serta menambahkan satu atau beberapa label yang bermakna dan informatif untuk memberikan konteks agar model machine learning dapat

belajar darinya.

- Handling Missing Valuee
Handliny missing values adalah proses mengatasi nilai-nilai yang hilang dalam dataset. Nilai yang hilang dapat muncul karena berbagai alasan, seperti ketidakhadiran data, kesalahan pengumpulan data, atau kegagalan sistem.
- Processing
Processing adalah bahasa pemrograman open source dan lingkungan bagi orang-orang yang ingin membuat gambar, animasi, dan interaksi.
- Case Folding
Case Folding merupakan tahapan dalam text processing yang di lakukan untuk menyeragamkan karakter pada data dan merubah text dalam dokumen menjadi huruf kecil (lowercase).
- Stopword
Stopword adalah proses filtering, pemilihan kata-kata penting dari hasil token yaitu kata-kata apa saja yang digunakan untuk mewakili dokumen.
- Tokenizing
Tokenizing yaitu tahap melakukan pemotongan string input berdasarkan tiap kata yang menyusunnya dan digunakan untuk membuang tanda baca.
- Steaming
Steaming data adalah data yang terus menerus dihasilkan oleh sumber yang berbeda. Data tersebut harus diproses secara bertahap menggunakan teknik pemrosesan aliran tanpa memiliki akses utama.
- Spliting Data
Spliting Data adalah metode membagi data menjadi dua bagia atau lebih yang membentuk subhimpunan data.
- Pembobotan
Pembobotan data adalah proses pemberian nilai atau bobot pada setiap observasi dalam dataser berdasarkan karakteristik tertentu.

E. Evaluasi

Pada tahap ini dilakukan evaluasi klasifikasi berdasarkan hasil pengujian sentimen positif dan negatif dan mengukur akuratnya hasil yang diperoleh menggunakan metode Naïve Bayes. Pada penelitian ini akan dievaluasi sesuai dari nilai.

HASIL DAN PEMBAHASAN

1. Pelabelan Data

Pelabela data merupakan proses mengidentifikasi data mentah (gambar, file teks, video, dll). Source Code dan Hasil Ulasan :

```
def pelabelan(score):
    if score < 3:
        return "Negatif"
    elif score == 4:
        return "Positif"
    elif score == 5:
        return "Positif"
my_df['Label'] = my_df['score'].apply(pelabelan)
my_df.head(50)
```

content	score	Label
134	Ngga ada masalah sama app nya, tapi tolong don...	1 Negatif
206	Uang saya dari pembatalan pesanan pembayaran Q...	1 Negatif
466	Tempat duduk di KFC ramainya parah penuh yg d...	1 Negatif
1086	LOGINNYA SUSAH BANGET	1 Negatif
96	Lebih cepet ga perlu antri 🙄	5 Positif
904	Malas pke app @_adid2 perbahana	2 Negatif
48	Klo orderan kurang komplan kemna? D apk menu ...	2 Negatif
89	Banyak masalahnya..QR nya tdk bisa terbuka	5 Positif
200	Baru makan baru donlot bener kata yg lain ...	1 Negatif
25	Untuk pemesanan burger tolong tambahkan opsi b...	3 None
23	Tidak ada masalah slma penggunaan..harusnya ada...	5 Positif
93	layanan buat customer gak lama mantap	5 Positif
417	Tidak pernah nkras mau dihar	1 Negatif

Gambar 2. Pelabelan Data

2. Handling Missing Value

Handling missing values adalah proses mengatasi nilai-nilai yang hilang dalam dataset. Source Code dan Hasil Ulasan :

```
my_df.dropna(subset=['Label'], inplace = True)
my_df.isnull().sum()
my_df.head(50)
```

content	score	Label
0	Ngga ada masalah sama app nya, tapi tolong don...	1 Negatif
1	Uang saya dari pembatalan pesanan pembayaran Q...	1 Negatif
2	Tempat duduk di KFC ramainya parah penuh yg d...	1 Negatif
3	LOGINNYA SUSAH BANGET	1 Negatif
4	Lebih cepet ga perlu antri 🙄	5 Positif
5	Malas pke app @_adid2 perbahana	2 Negatif
6	Klo orderan kurang komplan kemna? D apk menu ...	2 Negatif
7	Banyak masalahnya..QR nya tdk bisa terbuka	5 Positif

Gambar 3. Handling Missing Values

3. Processing

Source Code dan Hasil Ulasan :

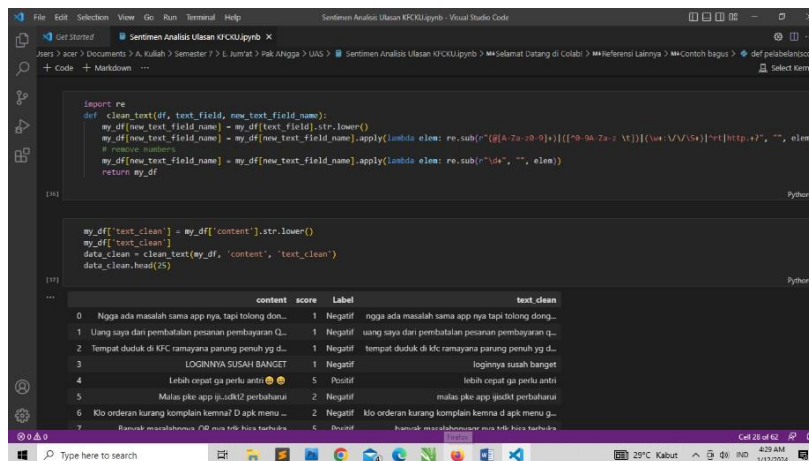
```
import pandas as pd
df = pd.read_csv('content/KFC02dataprocessing.csv')
df.head(75)
```

content	score	Label
0	Ngga ada masalah sama app nya, tapi tolong don...	1 Negatif
1	Uang saya dari pembatalan pesanan pembayaran Q...	1 Negatif
2	Tempat duduk di KFC ramainya parah penuh yg d...	1 Negatif
3	LOGINNYA SUSAH BANGET	1 Negatif
4	Lebih cepet ga perlu antri 🙄	5 Positif
...
70	Tidak berprkembangan	1 Negatif
71	Parah sih, 2x order by aplikasi 2x di batalkan...	1 Negatif
72	Harusnya ada opsi bayar di counter secara tunai...	4 Positif
73	Aplikasi yang baik untuk pesan tanpa antri	5 Positif
74	Di perbaiki dik mau cancel ga bisa mau larak p...	1 Negatif

Gambar 4. Tahap Processing

4. Case Folding

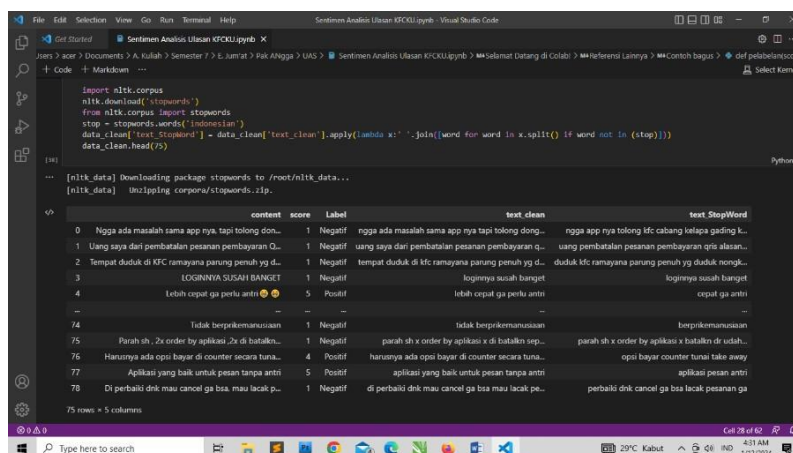
Case Folding merupakan tahapan dalam text processing yang di lakukan untuk menyeragamkan karakter pada data dan merubah text dalam dokumen menjadi huruf kecil (lowercase). Source Code dan Hasil Ulasan :



Gambar 5. Hasil Case Folding

5. Stopword

Stopword adalah proses filtering, pemilihan kata-kata penting dari hasil token yaitu kata-kata apa saja yang digunakan untuk mewakili dokumen. Source Code dan Hasil Ulasan :



Gambar 6. Stopword

6. Tokenizing

Tokenizing yaitu tahap melakukan pemotongan string input berdasarkan setiap kata yang menyusunnya dan digunakan untuk membuang tanda baca. Source Code dan Hasil Ulasan :

```
import nltk
nltk.download('punkt')
from nltk.tokenize import sent_tokenize, word_tokenize
data_clean['text_tokens'] = data_clean['text_stopword'].apply(lambda x: word_tokenize(x))
data_clean.head()
```

	content	score	Label	text_clean	test_StopWord	test_tokens
0	Nggak ada masalah sama app nya, tapi tolong don...	1	Negatif	nggak ada masalah sama app nya tapi tolong dong...	nggak app nya tolong kfc cabang kelapa gading k...	[nggak app nya, tolong, kfc, cabang, kelapa, gading, ...]
1	Uang saya dari pembatalan pesanan pembayaran Q...	1	Negatif	uang saya dari pembatalan pesanan pembayaran q...	uang pembatalan pesanan pembayaran q...	[uang, pembatalan, pesanan, pembayaran, q...
2	Tempat duduk di KFC rasanya memang enak yg d...	1	Negatif	tempat duduk di kfc rasanya memang enak yg d...	duduk kfc rasanya memang enak yg duduk nong...	[duduk, kfc, rasanya, memang, enak, yg, duduk, ...]
3	LOGINYA SUSAH BANGET	1	Negatif	loginya susah banget	loginya susah banget	[loginya, susah, banget]
4	Lebih cepat ga perlu antri	5	Positif	lebih cepat ga perlu antri	cepat ga antri	[cepat, ga, antri]

```
pip install Sastrawi
```

```
Collecting Sastrawi
  Downloading Sastrawi-1.0.1-py2.py3-none-any.whl (209 kB)
```

Gambar 7. Tokenizing

7. Steaming

Steaming data adalah data yang terus menerus dihasilkan oleh sumber yang berbeda. Source Code dan Hasil Ulasan :

```
pip install Sastrawi
```

```
Collecting Sastrawi
  Downloading Sastrawi-1.0.1-py2.py3-none-any.whl (209 kB)
    0/209.7 kB | eta -:--:--
    92.2/209.7 kB | 2.7 MB/s | eta 0:00:01
    209.7/209.7 kB | 3.0 MB/s | eta 0:00:00

Installing collected packages: Sastrawi
Successfully installed Sastrawi-1.0.1
```

```
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
factory = StemmerFactory()
stemmer = factory.create_stemmer()
```

```
#-----STEMMING-----
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
# import splitter

# create stemmer
factory = StemmerFactory()
stemmer = factory.create_stemmer()
```

Gambar 8. Steaming

8. Splitting Data

Splitting Data adalah metode membagi data menjadi dua bagian atau lebih yang membentuk subhimpunan data. Source Code dan Hasil Ulasan :

```
#Mungkin kita importkan library re, kemudian kita lakukan preproses
import re

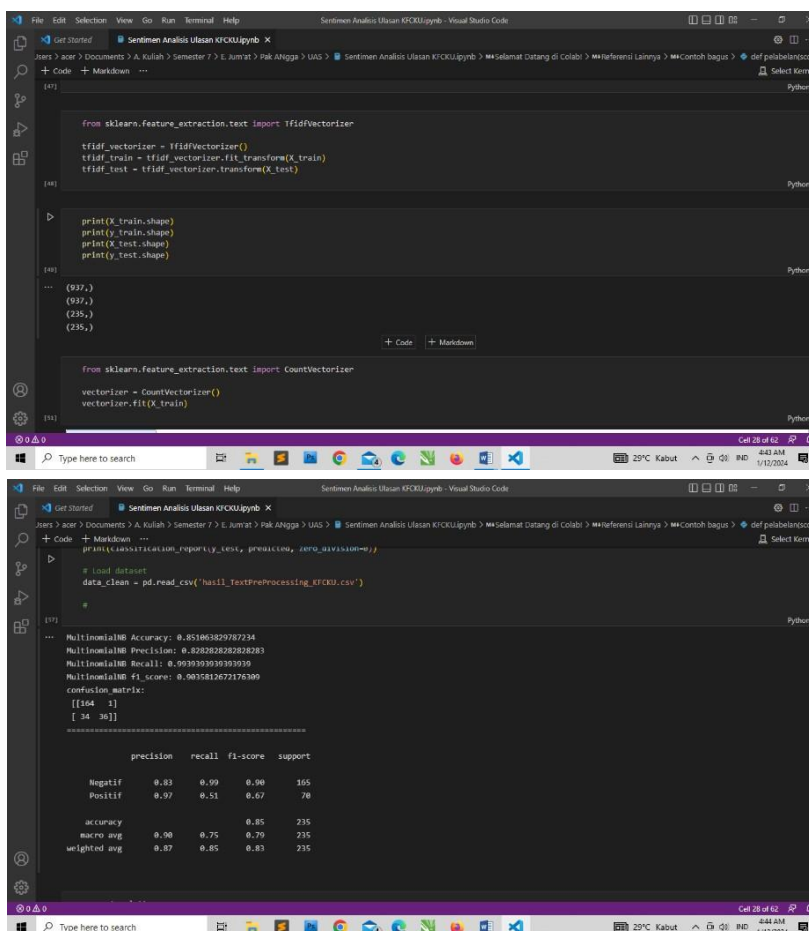
def preproses(text):
    text = re.sub('[0-9]+', '', text)
    emoticons = re.findall('(?:::|:)|--)(?:::|:)|[DIP]',
    text)
    text = re.sub('[0-9]+', '', text.lower()) +
    ''.join(emoticons).replace('-', '')
    return text
```

```
#membagi data menjadi data training dan testing dengan test_size = 0.20 dan random state nya 0
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(data_clean['content'], data_clean['Label'],
    test_size = 0.20,
    random_state = 0)
```

Gambar 9. Splitting Data

9. Pembobotan

Pembobotan data adalah proses pemberian nilai atau bobot pada setiap observasi dalam dataset berdasarkan karakteristik tertentu. Source Code dan Hasil Ulasan :



```
from sklearn.feature_extraction.text import TfidfVectorizer
tfidf_vectorizer = TfidfVectorizer()
tfidf_train = tfidf_vectorizer.fit_transform(X_train)
tfidf_test = tfidf_vectorizer.transform(X_test)

print(X_train.shape)
print(y_train.shape)
print(X_test.shape)
print(y_test.shape)

(937,)
(937,)
(235,)
(235,)

from sklearn.feature_extraction.text import CountVectorizer
vectorizer = CountVectorizer()
vectorizer.fit(X_train)

print(classification_report(y_test, predicted, zero_division=0))

# Load dataset
data_clean = pd.read_csv('hasil_textPreProcessing_KFCU.csv')

#
MultinomialNB Accuracy: 0.851963829787234
MultinomialNB Precision: 0.8282828282828283
MultinomialNB Recall: 0.9939393939393939
MultinomialNB f1_score: 0.903812672176389
confusion matrix:
[[104  1]
 [ 34 36]]

=====
              precision    recall  f1-score   support

Negatif    0.83     0.99     0.90     165
Positif    0.97     0.51     0.67     78

accuracy    0.85     0.85     0.85     235
macro avg   0.90     0.75     0.79     235
weighted avg 0.87     0.65     0.63     235
```

Gambar 10. Pembobotan

SIMPULAN

Dari hasil penelitian ini, ditemukan variasi tingkat akurasi sebesar 85%, presisi sebesar 82%, recall sebesar 99%, dan F1 Score sebesar 90%, dimaksudkan untuk mencapai keseimbangan antara data sentimen analisis positif dan negatif. Penambahan fitur stemming bertujuan untuk mereduksi kata-kata berimbuhan menjadi kata dasar, memperkaya analisis sentimen dengan informasi yang lebih fokus dan relevan.

Pendekatan Naive Bayes terbukti sebagai metode yang sangat efektif dalam menganalisis sentimen konsumen terhadap restoran cepat saji KFC. Metode ini berhasil mengklasifikasikan sentimen konsumen menjadi kategori positif, negatif, dan netral dengan akurasi yang memuaskan. Melalui analisis sentimen berbasis data Google Play Store, KFC dapat memperoleh umpan balik konsumen secara real-time dan menyeluruh. Inisiatif ini membantu perusahaan memahami pandangan dan persepsi konsumen

terhadap layanan dan produk mereka, serta memberikan kemampuan untuk mengidentifikasi area di mana peningkatan kualitas layanan dibutuhkan.

Dengan informasi yang diperoleh dari analisis sentimen, KFC dapat menjadi lebih responsif terhadap kebutuhan dan harapan pelanggan. Kemampuan untuk merespons dengan cepat terhadap masukan negatif memungkinkan perusahaan untuk mengimplementasikan perubahan yang diperlukan guna memperbaiki pengalaman pelanggan. Secara keseluruhan, penelitian ini memberikan landasan yang kuat bagi KFC untuk terus meningkatkan dan memperbaiki layanannya, menciptakan lingkungan bisnis yang lebih adaptif dan berorientasi pada kepuasan pelanggan.

UCAPAN TERIMAKASIH

Terimakasih Kepada Universitas Multimedia Nusantara dan Universitas Muhammadiyah Tangerang yang telah memberikan support dalam melakukan penelitian ini, sehingga penelitian ini telah selesai dilaksanakan.

DAFTAR PUSTAKA

- Andrianto, B., & Adinugroho, S. (2018). Analisis Sentimen Konten Radikal Melalui Dokumen Twitter Menggunakan Metode Backpropagation. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(12), 7380–7385.
- Damayunita, A., Fuadi, R. S., & Juliane, C. (2022). Comparative Analysis of Naive Bayes, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) Algorithms for Classification of Heart Disease Patients. *Jurnal Online Informatika*, 7(2), 219–225. <https://doi.org/10.15575/join.v7i2.919>
- Fairuz, A. L., Ramadhani, R. D., & Tanjung, N. A. F. (2021). Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter. *Journal of Dinda : Data Science, Information Technology, and Data Analytics*, 1(1), 42–51. <https://doi.org/10.20895/dinda.v1i1.180>
- Metode, D., Bayes, N., & Nbc, C. (2023). *Sentimen analisis opini masyarakat terhadap kasus pelecehan seksual di kpi pada social media twitter dengan metode naïve bayes classifier (nbc) 1*. 12(01), 122–133.
- Oliveira, A. S., Renda, A. I., Correia, M. B., & Antonio, N. (2022). Hotel customer segmentation and sentiment analysis through online reviews: An analysis of selected European markets. *Tourism and Management Studies*, 18(1), 29–40. <https://doi.org/10.18089/tms.2022.180103>
- PERMANA, A. A., & NOVIYANTO, W. A. (2023). Comparison of the Accuracy of the Lexicon-Based and Naive Bayes Classifier Methods To Public Opinions About Removing Masks on Social Media Twitter. *Journal of Theoretical and Applied Information Technology*, 101(3), 1174–1183.
- Rosadi, A., Gustiana, D., Informatika, M., Sti, S. J., No, J. B. R. I., Dalam, R., Baru, K., Selatan, J., Kata, P., Baku, T., & Bernegasi, M. K. (2021). *Analisis Sentimen Berdasarkan Opini Pengguna pada Media Twitter Terhadap BPJS Menggunakan Metode Lexicon Based dan Naïve Bayes Classifier Twitter Text Mining*. 20, 39–52.
- Rusli, A., Suryadibrata, A., Nusantara, S. B., & Young, J. C. (2020). A Comparison of Traditional Machine Learning Approaches for Supervised Feedback Classification in Bahasa Indonesia. *IJNMT (International Journal of New Media Technology)*, 7(1), 28–32. <https://doi.org/10.31937/ijnmt.v1i1.1485>
- Santoso, J. T., & Hartono, B. (2022). *DSS (Decision Support Systems) Sistem Pendukung Keputusan*. 466.

Sinaga, A. (2017). Implementasi Sentiment Analysis untuk Menentukan Tingkat Popularitas Tujuan Wisata. *Prosiding Seminar Nasional Teknologi Dan Rekayasa Informasi Tahun 2017, November, 24–25.*