

Ekstraksi Ulasan Pelanggan Restoran dalam Menggali Aspek Layanan pada Google Review dan Tripadvisor sebagai Faktor Peningkatan Kualitas

Dwija Wisnu Brata¹, Welly Purnomo², Achmad Nofandi³

¹FILKOM Universitas Brawijaya

²FILKOM Universitas Brawijaya

³FILKOM Universitas Brawijaya

¹wisnubrata@ub.ac.id, ²wepe@ub.ac.id, ³achmadnofandi@ub.ac.id

ABSTRAK. Ulasan pelanggan merupakan pandangan dan komentar yang diberikan oleh para pelanggan setelah mereka mengalami produk atau layanan. Ulasan pelanggan dapat menjadi sumber informasi berharga bagi pemilik, tetapi ada permasalahan tersendiri terkait ulasan yang menumpuk dan tidak sempatnya dalam membaca maksud isi ulasan, serta mencermati aspek apa yang mendapat sorotan pelanggan. Penelitian dilakukan menggunakan *deep learning*, pengumpulan data menggunakan teknik *web scrapping* dengan selenium tool berbasis python, diperoleh data sebanyak 2065 ulasan terdiri dari 1955 data google review dan 110 data tripadvisor. Aspek yang dibahas aspek kualitas pelayanan, kualitas makanan, lingkungan, dan harga dengan menggunakan algoritma *Artificial Neural Network (ANN)* dengan pembobotan kata menggunakan TF-IDF. Implementasi ketidakseimbangan dataset, diterapkan teknik random undersampling. Penyetelan hyperparameter dilakukan melalui fungsi GridsearchCV dari pustaka scikit-learn. Hasil pengujian model dievaluasi menggunakan *confusion matrix*, menghasilkan nilai akurasi 89%. Selanjutnya, dilakukan proses perankingan ulasan negatif untuk mengidentifikasi ulasan negatif yang paling sering diberikan pelanggan dan aspek yang menyertainya.

Kata Kunci: *Aspek Layanan, Sentimen Layanan Restoran, Web Scrapping*

ABSTRACT. Customer reviews are views and comments given by customers after they experience a product or service. Customer reviews can be a valuable source of information for owners, but there are certain problems related to reviews piling up and not having time to read the meaning of the contents of the reviews, as well as paying attention to what aspects are highlighted by customers. The research was conducted using deep learning, data collection using web scrapping techniques with the Python-based Selenium tool, data obtained for 2065 reviews consisting of 1955 Google review data and 110 TripAdvisor data. The aspects discussed are service quality, food quality, environment and price using the Artificial Neural Network (ANN) algorithm with word weighting using TF-IDF. Implementation of dataset imbalance, random undersampling technique applied. Hyperparameter tuning was done via the GridsearchCV function from the scikit-learn library. The model testing results were evaluated using a confusion matrix, producing an accuracy value of 89%. Next, a negative review ranking process was carried out to identify the negative reviews most frequently given by customers and the aspects that accompany them.

Keywords: *Service Aspects, Restaurant Service Sentiment, Web Scrapping*

PENDAHULUAN

Dalam menjalankan kegiatan bisnisnya, seorang pemilik restoran tidak hanya memikirkan strategi untuk menjadikan produk menjadi yang terbaik, tetapi juga memikirkan strategi untuk menjaga hubungan yang baik dengan pelanggan dari restoran. Salah satu isu strategis dalam lingkup Customer Relationship Management adalah Customer relationships itu sendiri dimana hal ini berkaitan dengan cara perusahaan memanfaatkan komunikasi umpan balik (feedback) dari pelanggan dalam proses bisnisnya (Frow P.E. et al., 2009). Salah satu strategi yang dapat dilakukan untuk menjaga hubungan baik dengan pelanggan adalah dengan memperhatikan pendapat, saran, masukan dan kritikan dari pelanggan untuk menyusun strategi demi kemajuan restoran. Pendapat dan saran tidak hanya berbentuk lisan saja, tetapi pada era digital, hal tersebut juga dapat disampaikan pada media online (baik forum, ataupun web terkait wisata), tidak menutup kemungkinan penyampaian tulisan pada media online atau sosial terkadang bernada positif (komen positif) ada juga yang bernada

negative (komen negative). Fenomena seperti ini menjadi hal yang menarik bagi peneliti untuk menganalisis terkait sentiment negative dan positif sebuah ulasan. Proses penggalian informasi terkait suatu entitas dan secara otomatis mengidentifikasi subjektivitas dari entitas tersebut dengan tujuan menentukan apakah teks tersebut menyampaikan opini positif, negatif, atau netral juga dilakukan oleh Dang et al., (2020). Analisis sentimen tanggapan masyarakat Indonesia terhadap pandemi Covid-19 pada media sosial twitter juga dilakukan oleh Pamungkas & Kharisudinaa (2021) dengan membandingkan tiga metode klasifikasi dan diperoleh algoritma SVM menghasilkan akurasi terbaik. Nilai akurasi SVM sebesar 90,01% , naïve bayes sebesar 79,2% dan algoritma K-NN sebesar 62,1%. Rekomendasi strategi bisnis juga dapat diperoleh berdasarkan analisis ulasan seperti yang dilakukan oleh Michelle Reene (2019) tentang “Improving Restaurants' Business Performance Using Yelp Data Sets through Sentiment Analysis”. Priska Liliani (2020) menyebutkan kualitas makanan dan kualitas pelayanan secara signifikan mempengaruhi kepuasan pelanggan. Sehingga, ulasan dapat dimanfaatkan untuk mendapatkan wawasan baru dalam penyusunan strategi bisnis. Penyusunan sebuah strategi bisnis diawali dengan diperolehnya aspek ulasan yang berkaitan dengan layanan, Sehingga strategi bisnis yang dihasilkan sesuai dengan data di lapangan. Besarnya jumlah ulasan yang mencapai ribuan, maka diperlukan cara efektif dalam mengolah ulasan. Hal ini menjadi pertimbangan dalam mendapatkan peningkatan kualitas restoran yang berasal dari data ulasan online.

Dari permasalahan yang telah dijelaskan pada paragraph sebelumnya, konsep yang akan dilakukan adalah dengan analisis sentimen ulasan pelanggan Restoran Warung Wareg pada aspek-aspek yang menjadi kunci pada proses bisnis restoran. Kontribusi pada penelitian ini, terdapatnya nilai aspek layanan yang diperoleh dari hasil analisis ulasan sentiment yang dilakukan, dimana terdapat aspek yang ingin dicapai berupa aspek kualitas pelayanan, kualitas makanan, lingkungan, dan harga. Aspek yang dibentuk juga akan mengaitkan dengan ulasan yang terkait, sehingga akan lebih mudah dalam memahami maksud dan memetakan strategi apa yang akan direncanakan terhadap info tersebut.

METODE

Langkah-langkah yang dilakukan penulis dalam melakukan penelitian digambarkan dalam bagan berikut:



Gambar 1. Diagram Alir Penelitian

Pengumpulan Data

Melakukan pengumpulan data yang dibutuhkan untuk melakukan penelitian ini. Data yang dibutuhkan berupa ulasan pelanggan restoran Warung Wareg yang terletak di Jl. Raya Dieng No.9, Sidomulyo, Kec. Batu, Kota Batu, Jawa Timur 65317.

Situs Google Review dan Tripadvisor menyediakan data ulasan pengguna yang digunakan untuk analisis sentimen. Bahasa pemrograman Python dan teknologi driver web Selenium digunakan untuk mengumpulkan data. Web scrapping adalah metode pengambilan data yang digunakan untuk mendapatkan informasi yang diperlukan dari halaman web Google Review dan Tripadvisor. Nama pengguna, konten ulasan, dan tanggal ulasan adalah tiga jenis data yang diambil. Sebanyak 110 titik data dari Tripadvisor dan 1955 data ulasan mentah dari Google Reviews diperoleh. menghasilkan

total 2065 ulasan yang dikumpulkan sebagai data ulasan. Selanjutnya data akan disimpan berupa dataset dengan format csv.

Labeling Data

Memberi label klasifikasi pada teks ulasan secara manual untuk menentukan apakah teks ulasan termasuk kedalam sentimen positif atau negatif. Selain itu data yang sudah diberi label kemudian dikelompokkan sesuai dengan kategori aspek yang sudah ditentukan yaitu aspek pelayanan, makanan, lingkungan, dan harga.

Text Processing

Text preprocessing bertujuan menyiapkan data mentah hasil dari tahapan pengumpulan data yang akan dilakukan klasifikasi. Text preprocessing berisi beberapa tahapan proses antara lain melakukan case folding untuk merubah seluruh huruf menjadi huruf kecil dan menghilangkan karakter selain huruf abjad, stop-word removal berupa penghapusan kata – kata yang tidak memiliki makna khusus pada kalimat, stemming bertujuan mengubah kata menjadi bentuk dasarnya dengan cara menghapus imbuhan dan akhiran pada kata, tokenizing bertujuan memecah kalimat menjadi kata-kata, dan character normalization berupa mengubah kata - kata yang tidak baku menjadi kata baku. Tahapan ini akan menggunakan library NLP Bahasa Indonesia seperti sastrawi

Classification Model

Melakukan klasifikasi teks ulasan dengan algoritma Artificial Neural Network pada dataset yang sudah dilakukan preprocessing sebelumnya. Klasifikasi juga dilakukan pada setiap kelompok aspek yang sudah ditentukan. Klasifikasi dilakukan dengan membagi dataset menjadi dua bagian yaitu dataset training dan dataset testing. Klasifikasi dilakukan dengan menggunakan library machine learning yaitu scikit-learn dengan bahasa pemrograman python.

Analisis dan Pengujian Hasil

Melakukan pengujian terhadap hasil klasifikasi sentimen ulasan untuk mengetahui tingkat akurasi dari pemodelan yang telah dilakukan sebelumnya. Pengujian dilakukan untuk mengetahui tingkat performa model klasifikasi yang telah dibuat dalam melakukan prediksi klasifikasi sentimen. Nilai yang dihasilkan dari proses pengujian yaitu Accuracy, Recall, Precision, dan f1-Score. Analisis mendalam akan dilakukan terhadap hasil klasifikasi untuk mengetahui informasi baru yang belum diketahui sebelumnya.

Pencarian Ulasan Terbanyak

Melakukan pencarian ulasan terbanyak dari keseluruhan teks ulasan pelanggan Warung Wareg. Pencarian dilakukan pada ulasan dengan sentiment positif dan negatif.

HASIL DAN PEMBAHASAN

Text Pre-Processing

Pada Kode Program 1 dapat dilihat proses preprocessing dengan random under sampler pada penelitian ini.

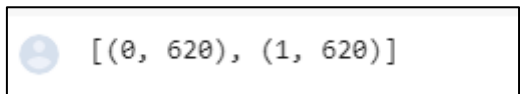
```
Random Under Sampling
from collections import Counter
from imblearn.under_sampling import RandomUnderSampler

x = df['stop2'].values.reshape(-1, 1)
y = df["Label"].values
```

```
undersampler = RandomUnderSampler(random_state=0)
X_undersampled, y_undersampled = undersampler.fit_resample(x, y)
print(sorted(Counter(y_undersampled).items()))
```

Kode Program 1 Random Under Sampling

Dengan menggunakan kode program tersebut, dataset akan mengalami proses undersampling menggunakan metode `RandomUnderSampler`. Hal ini akan menghasilkan dataset baru (`X_undersampled`, `y_undersampled`) yang memiliki jumlah sampel yang seimbang antara kelas mayoritas dan kelas minoritas. Kode terakhir mencetak jumlah sampel untuk setiap kelas setelah proses undersampling dilakukan. *Output* dari kode program tersebut dapat dilihat pada Gambar 3.



Gambar 3 Output Random Under Sampling

Pembagian Dataset

Pada Kode Program 2 dapat dilihat proses pembagian dataset pada penelitian ini.

```
Pembagian Dataset
from fast_ml.model_development import train_valid_test_split

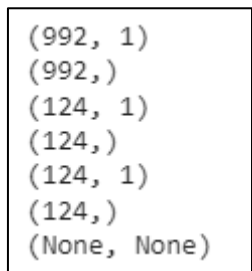
X_train, y_train, X_valid, y_valid, X_test, y_test =
train_valid_test_split(test, target = 'label',

train_size=0.8, valid_size=0.1, test_size=0.1)

# Get the shape of all the datasets
print(X_train.shape), print(y_train.shape)
print(X_valid.shape), print(y_valid.shape)
print(X_test.shape), print(y_test.shape)
```

Kode Program 1 Pembagian Dataset

Dengan menggunakan kode di atas, dataset test akan dibagi menjadi tiga subset: train, validasi, dan test, dengan proporsi masing-masing yang telah ditentukan. Hasilnya akan disimpan dalam enam variabel yang sesuai. Selanjutnya, kode akan mencetak bentuk (jumlah baris dan kolom) dari setiap subset untuk memastikan pemisahan data telah dilakukan dengan benar. *Output* dari kode program tersebut dapat dilihat pada Gambar 4.



Gambar 4 Hasil Split Ratio 80:20

Proses Tokenisasi dan Padding Teks

Pada Kode Program 3 dapat dilihat proses tokenisasi dan padding pada teks pada penelitian ini.

```
Tokenisasi dan Padding Teks

# Parameters
vocab_size = 10000
max_length = 75
trunc_type='post'
oov_tok = "<OOV>"

# Initialize the Tokenizer class
tokenizer = Tokenizer(num_words = vocab_size, oov_token=oov_tok)

# Generate the word index dictionary for the training sentences
tokenizer.fit_on_texts(X_train['text'])
word_index = tokenizer.word_index

# Generate and pad the training sequences
sequences = tokenizer.texts_to_sequences(X_train['text'])
padded = pad_sequences(sequences,maxlen=max_length,
truncating=trunc_type)

# Generate and pad the test sequences
testing_sequences = tokenizer.texts_to_sequences(X_test['text'])
testing_padded = pad_sequences(testing_sequences,maxlen=max_length)
```

Kode Program 2 Tokenisasi dan Padding Teks

Dalam keseluruhan kode di atas, dilakukan proses tokenisasi dan padding pada teks dari dataset `X_train` dan `X_test` menggunakan objek `Tokenizer` dan fungsi `pad_sequences`. Tujuannya adalah untuk mengubah teks menjadi sequence yang memiliki panjang yang seragam agar dapat diolah lebih lanjut dalam pemodelan atau analisis teks.

Membangun Model ANN

Pada Kode Program 4 dapat dilihat proses pembangunan model ANN pada penelitian ini.

```
Membangun Model ANN
```

```
# Parameters
embedding_dim = 32
dense_dim = 128
tf.keras.backend.clear_session()
model_ann = tf.keras.Sequential([
    tf.keras.layers.Embedding(vocab_size,
embedding_dim, input_length=max_length),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(dense_dim, activation='relu'),
    tf.keras.layers.Dense(1, activation='sigmoid')
])

# Set the training parameters
model_ann.compile(loss='binary_crossentropy',
optimizer='adam',metrics=['accuracy'])

# Print the model summary
model_ann.summary()
```

Kode Program 3 Membangun Model ANN

Dalam keseluruhan kode di atas, sebuah model neural network (ANN) dibangun menggunakan API keras. Model ini terdiri dari beberapa layer yang ditentukan secara berurutan. Setelah itu, model dikompilasi dengan fungsi loss, optimizer, dan metrik evaluasi yang sesuai. Terakhir, ringkasan model dicetak untuk melihat struktur dan jumlah parameter yang digunakan dalam model tersebut. Hasil dari model ANN yang dibangun dapat dilihat pada Gambar 5.

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 75, 32)	320000
flatten (Flatten)	(None, 2400)	0
dense (Dense)	(None, 128)	307328
dense_1 (Dense)	(None, 1)	129
Total params: 627,457		
Trainable params: 627,457		
Non-trainable params: 0		

Gambar 5 Hasil Model ANN

Confusion Matrix

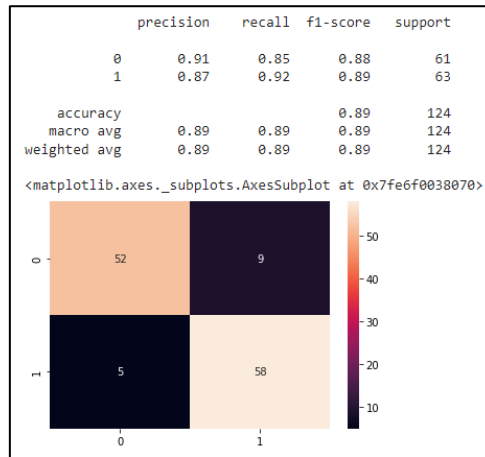
Pada Kode Program 5 dapat dilihat proses *generate confusion matrix* pada penelitian ini.

```
Confusion Matrix

from sklearn.metrics import classification_report, confusion_matrix
print(classification_report(actual_labels, predicted_rounded))
visualize_cm = metrics.confusion_matrix(actual_labels,
predicted_rounded)
sns.heatmap(visualize_cm, annot=True)
```

Kode Program 4 Confusion Matrix

Dengan menggunakan kode di atas, dapat dicetak laporan klasifikasi yang memberikan informasi detail tentang performa model dalam melakukan klasifikasi. Selain itu, divisualisasikan *confusion matrix* dengan menggunakan heatmap untuk mendapatkan pemahaman visual tentang seberapa baik model melakukan klasifikasi pada setiap kelas. Hasil dari *confusion matrix* dapat dilihat pada Gambar 6.



Gambar 6 Confusion Matrix

Analisis Hasil Confusion Matrix

Hasil klasifikasi akan diuji dengan menggunakan 1200 data ulasan yang telah dipisahkan menjadi 2 polaritas sentimen, yaitu sentimen negatif dan positif, dan 4 kategori aspek, yaitu pelayanan, makanan, lingkungan, dan harga. Setelah tahap pelatihan, pengujian mencoba mengevaluasi kinerja model klasifikasi. Pada tahap pelatihan, dataset ulasan dibagi menjadi dua bagian: data pelatihan (90% dari dataset) dan data uji (10% dari dataset). Kemudian, ada 120 data ulasan dan 120 data uji. Setelah itu, teknik Artificial Neural Networks (ANN) akan digunakan untuk melatih data pelatihan.

Analisis Actual dan Predicted Confusion Matrix

Modul confusion matrix digunakan untuk menguji klasifikasi sentimen. Nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) akan ditampilkan dalam Confusion Matrix. Nilai ini dihasilkan dengan membandingkan hasil prediksi dengan data uji yang telah disiapkan sebelumnya. Tabel di bawah ini menunjukkan kesimpulan dari confusion matrix penelitian ini.

Tabel 1 Confusion Matrix Table

Actual	Prediksi	
	0 (Negatif)	1 (Positif)
0 (Negatif)	58	5
1 (Positif)	9	52

Perhitungan Sentimen Positif : $Presisi = \frac{52}{52+5} = 0,912$

$$Recall = \frac{52}{52+9} = 0,852$$

$$Accuracy = \frac{52+58}{52+5+58+9} = 0,887$$

Analisis dari hasil sentimen positif didapat bahwa presisi sebesar 0,912 menunjukkan bahwa dari semua contoh yang diprediksi oleh model sebagai positif, sekitar 91,2% adalah benar-benar positif. Hal ini menunjukkan bahwa ketika model melabeli sebuah contoh sebagai positif, biasanya model tersebut benar. Ketepatan yang tinggi sangat penting ketika biaya positif palsu (salah melabeli sesuatu sebagai positif) relatif tinggi. Berikutnya, nilai recall sebesar 0,852 menunjukkan bahwa dari semua contoh positif yang sebenarnya, model berhasil mengidentifikasi dengan benar sekitar 85,2% dari mereka. Recall yang lebih tinggi menunjukkan model yang mahir dalam menangkap contoh positif, yang sangat penting ketika kehilangan contoh positif dapat memiliki konsekuensi yang signifikan. Nilai accuracy sebesar 0,887 yang menunjukkan bahwa model mengklasifikasikan 88,7% contoh dengan benar. Meskipun akurasi adalah metrik yang banyak digunakan, ini mungkin bukan pilihan terbaik ketika berhadapan dengan set data yang tidak seimbang atau ketika biaya positif palsu dan negatif palsu tidak sama.

Berdasarkan metrik ini, dapat disimpulkan bahwa model tersebut berkinerja baik dalam hal prediksi positif (presisi) dan menangkap contoh positif yang sebenarnya (recall). Akurasi keseluruhan juga cukup tinggi, menunjukkan bahwa model ini mahir dalam membuat prediksi yang benar.

Perhitungan Sentimen Negatif : $Presisi = \frac{58}{58+9} = 0,865$

$$Recall = \frac{58}{58+5} = 0,920$$

$$Accuracy = \frac{58+52}{58+9+52+5} = 0,887$$

Hasil perhitungan presisi, recall, accuracy dapat dianalisis bahwa hasil presisi 0,865 menunjukkan 86,5% dari contoh yang diprediksi sebagai positif benar-benar positif dan juga menjelaskan bahwa ketika model menetapkan sebuah contoh sebagai positif, model tersebut benar sekitar 86,5% dari waktu. Sedangkan recall menghasilkan 0,920 menandakan bahwa model mengidentifikasi dengan benar sekitar 92% dari semua contoh positif yang sebenarnya (sensivitas). Pada nilai akurasi sebesar 0,887 menunjukkan bahwa model membuat prediksi yang akurat untuk sekitar 88,7% kasus.

Berdasarkan metrik-metrik ini, terbukti bahwa model berkinerja baik dalam hal prediksi positif (presisi) dan menangkap contoh positif yang sebenarnya (recall). Akurasi keseluruhan juga cukup tinggi, menunjukkan bahwa model ini mahir dalam membuat prediksi yang akurat secara umum.

Pencarian Ulasan Terbanyak

Dilakukan pencarian ulasan paling dominan pada kelas sentimen positif dan negatif. Tujuan pencarian ini adalah untuk mengidentifikasi frasa kata yang paling umum muncul dalam seluruh dokumen ulasan pelanggan Warung Wareg. Kode Program 6 menunjukkan Kode untuk Pencarian Ulasan Terbanyak. Langkah pertama yang diambil melibatkan perhitungan kemunculan setiap kata dalam dataset (Tf), sebagaimana terlihat dalam baris kode 1-9. Setelah itu, langkah berikutnya melibatkan penghapusan kata yang sama dalam setiap ulasan untuk mencegah ketidakseimbangan perhitungan skor pada ulasan, sebagaimana dijelaskan dalam baris kode 14-17. Pada akhirnya, dilakukan perhitungan skor dari kalimat-kalimat dalam ulasan, seperti yang tercermin dalam baris kode 26-37.

Menurut penelitian Ferlin, Bachtiar, dan Rusydi, rumus untuk menentukan skor kalimat ulasan diterapkan dengan cara menjumlahkan skor semua kata dalam kalimat ulasan, kemudian membaginya dengan jumlah kata (Bachtiar et al., 2020).

```

1 def split(text):
2     return text.split(" ")
3
4 df_negatif['test']=df_negatif['stop2'].apply(split)
5
6 freq=FreqDist(sum(df_negatif['test'], []))
7 freq1 = FreqDist(sum(df_negatif['test'], [])).most_common(5000)
8 freq_df=pd.DataFrame(freq1, columns=['word', 'score'])
9 freq_df
10
11 print(freq_df.sort_values(by=['score'], ascending = False))
12
13 #menghapus kata yang sama dalam setiap kalimat ulasan
14 def RemoveSameWord(text):
15     str = set(text.split())
16     return ''.join(dict.fromkeys(text.split()))
17 df_negatif['test'] = df_negatif['stop2'].apply(lambda x: RemoveSameWord(x))
18
19 #menghitung score kalimat
20 def sentenceScore(text):
21     score = 0
22     text = text.split()
23     for t in text:
24         for i in range(0,len(freq_df)):
25             if t == freq_df.iloc[i][0]:
26                 score += freq_df.iloc[i][1]
27                 break;
28     score = score/len(text)
29     return score
30 df_negatif['score']=df_negatif['test'].apply(lambda
31 x:sentenceScore(x))
32 # Export CSV hasil pencarian ulasan terbanyak untuk dilakukan
33 root cause
34 df_negatif.to_csv('ulasan-terbanyak-negatif.csv', encoding='utf-
35 8')
```

Kode Program 6 Kode Pencarian Ulasan Terbanyak

Analisis Klasifikasi ANN

Tahap klasifikasi dilakukan pada dataset lengkap ulasan pelanggan dari Warung Wareg, yang terdiri dari total 3839 ulasan yang diterima selama tahap perincian ulasan. Data yang digunakan untuk ulasan ini diperoleh dari sumber-sumber terkemuka seperti Google Review dan Tripadvisor. Ulasan mencakup empat dimensi utama, khususnya aspek kualitas makanan, harga, lingkungan, dan layanan pelanggan. Setiap komponen dikaitkan dengan dua polaritas sentimen, khususnya negatif dan positif. Setelah tahap prapemrosesan selesai, data ulasan akan dikategorikan ke dalam ulasan yang berbeda.

Hasil Klasifikasi Aspek Pelayanan

Analisis ini berfokus pada komponen layanan Warung Wareg, seperti yang ditunjukkan dalam ulasan pelanggan. Secara khusus, studi ini meneliti kualitas layanan yang diberikan oleh personil, efisiensi penyampaian layanan, dan efektivitas sistem reservasi tempat. Jumlah total ulasan untuk komponen layanan berjumlah 530 ulasan.

Tabel 2 Jumlah Klasifikasi Aspek Pelayanan

Sentimen	Jumlah Ulasan Pelanggan
Positif	295
Negatif	235
Total	530

Dengan selisih 60 ulasan, dapat dikatakan bahwa elemen pelayanan memiliki kuantitas evaluasi yang seimbang. Hal ini menunjukkan bahwa untuk mengurangi jumlah ulasan negatif, tingkat layanan yang ditawarkan oleh personil, kecepatan layanan, dan kemampuan untuk memesan tempat harus ditingkatkan.

Hasil Klasifikasi Aspek Lingkungan

Ulasan mengenai aspek lingkungan, seperti lokasi, suasana, kebersihan, tata letak, dan fasilitas, dapat ditemukan di ulasan pelanggan Warung Wareg. Data ulasan mengenai aspek lingkungan berjumlah 1323 ulasan.

Tabel 3 Jumlah Klasifikasi Aspek Lingkungan

Sentimen	Jumlah Ulasan Pelanggan
Positif	1226
Negatif	97
Total	1323

Terlihat jelas bahwa mayoritas penilaian sentimen yang menguntungkan adalah untuk komponen lingkungan. Hal ini menunjukkan bahwa mayoritas pelanggan Restoran Warung Wareg menyukai, puas, dan memenuhi persyaratan untuk pertimbangan lingkungan. Perbedaan yang signifikan dalam volume data juga menunjukkan bahwa bagi pelanggan restoran, faktor lingkungan adalah yang paling penting dan berkesan.

Hasil Klasifikasi Aspek Makanan

Evaluasi pelanggan terhadap sajian makanan Warung Wareg dikategorikan berdasarkan rasa, presentasi, kualitas, jumlah, keragaman menu, dan aksesibilitas. Terdapat 1.550 ulasan mengenai aspek-aspek yang berhubungan dengan makanan.

Tabel 4 Jumlah Klasifikasi Aspek Makanan

Sentimen	Jumlah Ulasan Pelanggan
Positif	1354
Negatif	196
Total	1550

Terbukti bahwa ulasan yang berhubungan dengan makanan sangat positif. Hal ini menunjukkan bahwa mayoritas pengunjung di Warug Wareg menikmati, puas, dan merasa sesuai dengan kriteria terkait makanan yang berkaitan dengan rasa, presentasi, kualitas, jumlah, keragaman menu, dan ketersediaan makanan & minuman. Namun, tidak mungkin untuk mengabaikan jumlah ulasan yang tidak menguntungkan.

Hasil Klasifikasi Aspek Harga

Ulasan pelanggan tentang harga menu makanan dan minuman Warung Wareg dan perbandingan kesesuaian harga dengan kualitas dan kuantitas menu menjadi dasar ulasan aspek harga. Jumlah ulasan untuk bagian makanan adalah 436 ulasan.

Tabel 5 Jumlah Klasifikasi Aspek Harga

Sentimen	Jumlah Ulasan Pelanggan
Positif	344
Negatif	92
Total	436

Jika dilihat maka 78.9% pelanggan yang memberikan ulasan pada aspek harga sudah merasa aspek harga memuaskan. Namun 21.1% pelanggan masih merasa harga yang diberikan terlalu mahal bagi mereka.

Analisis Prediksi Hasil Sentimen

Analisis kesalahan hasil prediksi sentimen dapat dilihat pada tabel 5.6 dibawah ini. Terlihat bahwa model melakukan kesalahan prediksi baik berupa False Negatif maupun False Positif. Penyebab kesalahan prediksi oleh model terjadi karena beberapa faktor. Salah satu contohnya pada ulasan nomor 1 pada tabel 5.6 memiliki label asli positif namun kata “antri” dan “lama” yang terdapat pada ulasan tersebut membuat model memprediksi ulasan tersebut bersentimen negatif karena kata “antre” dan “lama” banyak terdapat pada data latih dengan sentimen negatif.

Pada ulasan nomor 4 maksud dari ulasan tersebut adalah untuk memberi saran agar memperbanyak menu, namun kata perbanyak nantinya akan masuk ke proses stemming sehingga akan menjadi kata “banyak” dimana kata ini sendiri banyak terdapat pada data latih kelas sentimen positif sehingga model memprediksi ulasan ini sebagai ulasan positif.

Tabel 6 Perbedaan Label Awal dan Prediksi Pada Aspek

No	Ulasan	Aspek	Label Awal	Label Prediksi
1	Walaupun pas dtg antri dulu tp antrian lumayan cepet.. Pas duduk pesen, ga lama makanannya datang.. Cepet bgt pelayanannya pas weekend mantap	Pelayanan	1	0
2	ada inovasi reservasi sehingga tidak menunggu terlalu lama sehingga kepuasan pengunjung semakin meningkat		1	0
3	cepat enggak kecewa		1	0
4	Perbanyak menu dan	Makanan	0	1
5	Enakkk bangetttt ga ad yang mengecewakan Kesini hr minggu dan waiting list tapi waktu tungguanya cepet, ga berasa		1	0
	Enak dan ga rugi kesini Menyaja wajib pesen gurami Mentor			
	Pengen balik sini lg buat gurami mentor Terus pesen cah kangkung jg enak Ayam gorengnya jg enak, empuk & bumbunya meresap			
6	ikan gurami manis tidak bau jeruk manis enak		1	0
7	tempat not cozy but nice...	Lingkungan	0	1
8	Tempat makan dengan nuansa yg alami. Banyak kolam ikan dan pohon yang menambah kesejukan. Bila weekend lebih ramai, jadi awas jangan sampai lama di antri. Lokasi di depan jalan dari pertigaan selecta. Memiliki parkirana yang luas.		1	0
9	Kedatangan saya untuk yg kedua kalinya, dan setiap kali selalu tidak mengecewakan, baik dari segi kebersihan maupun		1	0
10	harga enggak terlalu mahal	Harga	1	0
11	harga sesuai standar enggak terlalu mahal		1	0

12	Harganya murah tetapi sangat ramai tergantung mau nunggu apa enggak..	1	0
----	---	---	---

Hasil Pencarian Frekuensi Ulasan Terbanyak

Distribusi ulasan konsumen sangat penting dalam menyediakan informasi untuk mengambil keputusan. Pada sentimen baik dan negatif, pencarian dilakukan untuk ulasan yang paling umum. Teks ulasan yang paling sering muncul dan skor ulasan dikembalikan oleh pencarian. Tabel 7 dibawah menunjukkan 10 ulasan teratas yang ada pada kelas sentimen positif untuk seluruh aspek.

Tabel 7 Ulasan Terbanyak Sentimen Positif

No	Aspek	Ulasan	Skor
1	Makanan	tempat makan enak	1194
2	Lingkungan	makan tempat enak	1194
3	Lingkungan	tempat enak	1188
4	Lingkungan	enak tempat	1188
5	Makanan	enak makan	1188
6	Makanan	makan enak	1188
7	Makanan	enak	1170
8	Lingkungan	tempat makan enak nyaman	1021
9	Makanan	tempat nyaman makan enak	1021
10	Lingkungan	enak tempat makan keluarga	972

Tabel 8 dibawah menunjukkan 10 ulasan teratas yang ada pada kelas sentimen positif untuk seluruh aspek.

Tabel 8 Ulasan Terbanyak Sentimen Negatif

No	Aspek	Ulasan	Skor
1	PELAYANAN	pelayanan kurang	118
2	MAKANAN	makan biasa	113
3	PELAYANAN	pelayanan lama	112
4	MAKANAN	menu makan kurang banyak	105
5	MAKANAN	makan lumayan	105
6	PELAYANAN	lama	102
7	MAKANAN	makan standar	101
8	PELAYANAN	pesan makan keluar lama	99
9	LINGKUNGAN	makan antre tempat	98
10	MAKANAN	makan kurang variatif	98

KESIMPULAN DAN SARAN

Berdasarkan hasil dan pembahasan penelitian maka dapat disimpulkan bahwa Analisis ulasan pelanggan dapat dilakukan menggunakan metode analisis sentimen dengan bahasa pemrograman python yang terdiri dari beberapa tahapan seperti pengumpulan data, pelabelan data, text preprocessing, term weighting, pemodelan klasifikasi, analisis aspek dan pengujian hasil, pencarian ulasan terbanyak.

Klasifikasi sentimen yang dilakukan menggunakan metode Artificial Neural Network (ANN) dengan Term Frequency-Inverse Document Frequency (TFIDF) sebagai pembobotan kata menghasilkan nilai precision senilai 0.87, recall 0.92 dan f1-score 0.89 pada kelas sentimen negatif. Sedangkan pada kelas sentimen positif menghasilkan nilai precision senilai 0.91, recall 0.85, dan f1-score 0.88. Penentuan performa model menggunakan acuan nilai accuracy karena dataset yang digunakan seimbang selain itu tujuan dari klasifikasi untuk memprediksi kedua kelas secara akurat. Nilai dari accuracy sebesar 0.89 sehingga model dapat memprediksi kedua kelas dengan baik.

Rekomendasi bisnis dilakukan dengan mengacu pada akar permasalahan yang sudah didapatkan. Rekomendasi didapatkan dari hasil diskusi dengan pihak stakeholder untuk mendapatkan rekomendasi yang sesuai dengan akar permasalahan yang ada. Pada aspek pelayanan menghasilkan rekomendasi untuk melakukan penerapan SOP untuk pelayanan oleh pegawai dan penambahan pegawai pada saat high season. Rekomendasi untuk aspek makanan berupa pemilihan bahan mentah dengan kualitas terbaik, menetapkan SOP bagi koki, dan menambah varian menu. Pada aspek lingkungan menghasilkan rekomendasi untuk menambah meja atau memperluas tempat jika memungkinkan, menambah fasilitas wifi, dan penerapan teknis petugas kebersihan. Pada aspek harga menghasilkan rekomendasi untuk menggunakan bahan baku dengan kualitas terbaik.

Saran yang dapat penulis berikan untuk penelitian selanjutnya dengan tema serupa adalah menggunakan metode scraping selain selenium karena implementasi scroll pada selenium memakan waktu dan tidak semua review terambil. Web driver juga membutuhkan resource ram yang besar saat menampilkan banyak jumlah ulasan. Cara alternatif dapat melakukan scraping menggunakan resource dari cloud computing. Selain itu dapat menggunakan Apify yang menyediakan fungsi scraping secara otomatis. Menggunakan data ulasan yang terbaru baik dari sumber yang sama maupun dari sumber lain yang berbasis teks. Proses pelabelan secara manual memakan waktu yang lama dan bersifat subyektif. Proses pelabelan dapat memanfaatkan algoritma seperti feature extraction dan clustering agar dapat dilakukan secara otomatis.

DAFTAR RUJUKAN

- Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(1), 147. <https://doi.org/10.25126/jtiik.0813944>.
- Ching, M. R. D., & De Dios Bulos, R. (2019). Improving Restaurants' Business Performance Using Yelp Data Sets through Sentiment Analysis. *Proceedings of the 2019 3rd International Conference on E-Commerce, E-Business and E-Government - ICEEG 2019*. <https://doi.org/10.1145/3340017.3340018>.
- Dang, N. C., Moreno, M. N., & De La Prieta, F. (2020). Sentiment Analysis Based on Deep Learning: A Comparative Study. *Electronics*, 9(3), 483. <https://doi.org/10.3390/electronics9030483>.
- Djufri, M. (2020). PENERAPAN TEKNIK WEB SCRAPING UNTUK PENGGALIAN POTENSI PAJAK (Studi Kasus pada Online Market Place Tokopedia, Shopee dan Bukalapak) (Vol. 13).
- Gifari, O. I., Adha, M., Rifky Hendrawan, I., Freddy, F., & Durrand, S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine. *JIFOTECH (JOURNAL OF INFORMATION TECHNOLOGY)*, 2(1).
- Gunawan, D., Riana, D., Ardiansyah, D., Akbar, F., & Alfarizi, S. (2020). Komparasi Algoritma Support Vector Machine Dan Naïve Bayes Dengan Algoritma Genetika Pada Analisis Sentimen Calon Gubernur Jabar 2018-2023. *Jurnal Teknik Komputer AMIK BSI*. <https://doi.org/10.31294/jtk.v6i1.6866>.

- Imady, P. Al, Setianingsih, C., & Ruriawan, M. F. (2023). Deteksi Threat Dan Vulnerability Pada Unggahan Twitter Menggunakan Algoritma Naïve Bayes. *E-Proceeding of Engineering*, 10(1), 304–312.
- Liliani, P. (2020). Analisis Pengaruh Kualitas Makanan Dan Kualitas Pelayanan Terhadap kepuasan pelanggan Dan Dampaknya Pada behavioral intention restoran top Yammie. *Jurnal Bina Manajemen*, 8(2), 18–48. <https://doi.org/10.52859/jbm.v8i2.85>
- Maisat, Z., Darmawan, E., & Fauzan Dianta, A. (2023). Implementasi Optimasi Hyperparameter GridSearchCV Pada Sistem Prediksi Serangan Jantung Menggunakan SVM. *Jurnal Ilmiah Sistem Informasi Unipdu*, 13(1), 8–15. <https://doi.org/10.26594/teknologi.v13i1.3098>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. In *Jurnal Sains Komputer & Informatika (J-SAKTI (Vol. 5, Issue 2))*.
- Nugraha, A. E., Rizal, S., & Pratiwi, N. K. C. (2022). Klasifikasi Penyakit Pada Tanaman Singkong Menggunakan Arsitektur VGGNET Berbasis Deep Learning.
- Pamungkas, F., & Kharisudin, I. (2021). Analisis Sentimen dengan SVM, NAIVE BAYES dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter. *PRISMA, Prosiding Seminar Nasional Matematika*, 4, 628-634
- Samsir, Wiguna, C., Verawardina, U., Edi, F., & Watrianthos, R. (2021). Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(1), 157–163. <https://doi.org/10.30865/mib.v5i1.2604>
- Saputra, C. B., Muzakir, A., & Udariansyah, D. (2020). ANALISIS SENTIMEN MASYARAKAT TERHADAP #2019GANTIPRESIDEN BERDASARKAN OPINI DARI TWITTER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER. *Bina Darma Conference on Computer Science*.
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>
- Sartini. (2020). ANALISIS SENTIMEN TWITTER BAHASA INDONESIA MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK.
- Satriajati, S., Bagus Panuntun, S., & Pramana, S. (2020). IMPLEMENTASI WEB SCRAPING DALAM PENGUMPULAN BERITA KRIMINAL PADA MASA PANDEMI COVID-19 Studi Kasus: Situs Berita detik.com.
- Setiawan, A., Santoso, L. W., & Adipranata, R. (2020). Klasifikasi Artikel Berita Bahasa Indonesia Dengan Naive Bayes Classifier.
- Suthaharan, S. (2016). Support Vector Machine (pp. 207–235). https://doi.org/10.1007/978-1-4899-7641-3_9.