



Klasifikasi Kepuasan Pelanggan Terhadap Platform E-Commerce dengan Metode K-Nearest Neighbor (K-NN)

Tio Novian Halim¹, Ryan Martin², Rista Kesia Br Hutasoit³, Siti Aisyah⁴

^{1,2,3,4}Universitas Prima Indonesia, Indonesia

Email: tionovianhalim@gmail.com¹, ryanmartin090301@gmail.com², rikeshhutasoit@gmail.com³

Abstract

This study aims to determine customer satisfaction, towards e-commerce platforms. E-commerce is a platform that can make various types of payments and purchases online from a smartphone. This research is a quantitative approach using survey techniques, namely by distributing questionnaires. The data used uses primary data in the form of a questionnaire obtained through google form. Sampling using Convenience Sampling technique so that a research sample of 1001 respondents was obtained. Data processing in this study using the K-Nearest Neighbor (K-NN) method, the results showed that customer satisfaction had a significant effect on e-commerce both partially. based on training and data processing calculations of the K-NN model, the two best e-commerce were obtained according to the respondents' assessment, namely Shopee and TokoPedia, with a percentage of Shopee 82.98% of 779 total respondents and TokoPedia 80.30% of 133 total respondents.

Keywords: E-Commerce, KNN, EDA, Pyht, Classification

Abstrak

Penelitian ini bertujuan untuk mengetahui kepuasan pelanggan, terhadap platform e-commerce. E-commerce adalah sebuah platform yang dapat melakukan beragam jenis pembayaran dan pembelian secara online dari smartphone. Penelitian ini merupakan pendekatan kuantitatif dengan menggunakan teknik survey dengan penyebaran kuesioner. Data yang digunakan menggunakan data primer berupa kuesioner yang diperoleh melalui google form. Pengambilan sampel menggunakan teknik Convenience Sampling sehingga diperoleh sampel penelitian sebanyak 1001 responden. Pengolahan data dalam penelitian ini menggunakan metode K-Nearest Neighbor (K-NN), Hasil penelitian menunjukkan bahwa kepuasan pelanggan berpengaruh signifikan terhadap e-commerce baik secara parsial. berdasarkan pelatihan dan perhitungan pengolahan data model K-NN didapatkan dua e-commerce terbaik sesuai dengan penilaian responden yaitu Shopee dan TokoPedia, dengan persentase Shopee 82.98% dari 779 jumlah responden dan TokoPedia 80.30% dari 133 jumlah responden.

Kata kunci: E-Commerce, KNN, EDA, Python, klasifikasi

1. PENDAHULUAN

Perkembangan teknologi saat ini ter-khususnya internet berdampak pada banyak bidang kehidupan, termasuk perdagangan. Perdagangan tidak hanya mengubah pemasaran produk, tetapi juga proses jual beli. Proses (kegiatan) perdagangan ini umumnya banyak dikenal sebagai perdagangan *electronic* atau singkatnya *E-Commerce* [1]. *E-Commerce* adalah sekumpulan aplikasi, teknologi, dan proses bisnis dinamis yang menghubungkan perusahaan maupun konsumen, serta masyarakat melalui transaksi elektronik dan pertukaran barang, layanan, dan informasi melalui elektronik. Dalam perkembangan saat ini semakin banyak platform *E-Commerce* yang mengubah gaya hidup masyarakat, sehingga kami

tertarik dengan Klasifikasi kepuasan pelanggan terhadap platform *E - Commerce* ini seperti Lazada, Shopee, Tokopedia, Bukalapat, Blibli, dan lain sebagainya [2].

Semakin meningkatnya perkembangan *E - Commerce* yang ada di Indonesia saat ini terbilang sangatlah pesat. Hal itupun dapat dibuktikan dari banyaknya minat masyarakat dalam berbelanja online yang dapat kita lihat dari riset terbaru google dan termasuk dalam laporan *E-conomy SEA 2018* menunjukkan bshwasanya ekonomi digital Indonesia mengalami peningkatan hingga 500 % di tahun 2018 [3]. Ditahun 2019 - 2021 dimana berbelanja online meningkat drastis dan hal ini di sebabkan oleh *COVID-19* khusus nya di Indonesia. Hal ini dibuktikan oleh riset google satatistik data boks Sebanyak 88,1% pengguna internet yang ada di Indonesia memakai layanan *E - Commerce* guna membeli produk tertentu dalam kurun waktu beberapa bulan terakhir. Persentase tersebut merupakan yang tertinggi di dunia dalam hasil survei *We Are Social* pada April 2021 [4].

Data mining adalah proses mengumpulkan dan mengelolah data yang dimana bertujuan untuk mengekstrak informasi penting pada data. Proses pengumpulan dan ekstraksi informasi tersebut dapat dilakukan menggunakan perangkat lunak dengan bantuan perhitungan statistika, matematika, ataupun teknologi *Artificial Intelligence* [5]. Algoritma K-NN ialah algoritma untuk melakukan Klasifikasi terhadap objek dengan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. KNN dalam data mining bekerja berdasarkan jarak terpendek dari query instance ke training data. Salah satu cara untuk menghitung jarak dekat atau jauhnya tetangga menggunakan metode *euclidian distance* [6]. Algoritma ialah suatu upaya dengan urutan operasi yang disusun secara logis dan sistematis untuk menyelesaikan suatu masalah untuk menghasilkan suatu output tertentu [7]. Algoritma K-NN adalah metode untuk mengklasifikasikan fitur berdasarkan data pelatihan yang paling dekat dengan fitur tersebut. Data pelatihan diproyeksi ke ruang multidimensi di mana setiap dimensi mewakili fitur data [8]. K-NN atau k-tetangga terdekat adalah salah satu metode yang menerapkan algoritma *supervised*, algoritma tesebut dibagi menjadi 2 jenis yaitu *supervised learning* dan *unsupervised learning*. *Supervised learning* ialah suatu metode yang mengklasifikasikan dimana setiap pengamat memiliki variabel prediktor yang berkaitan dengan variabel respon dengan tujuan guna mendapatkan pola baru, sedangkan tujuan dari *unsupervised learning* yaitu untuk mendapatkan pola dalam sebuah data [9]. Beberapa penelitian yang pernah dilakukan dengan menggunakan K-NN salah satunya Klasifikasi terkait kepuasan konsumen [10].

2. METODOLOGI PENELITIAN

2.1. Jenis Penelitian

Adapun jenis penelitian yang dilakukan merupakan penelitian data yang diperoleh berdasarkan responeden dikalangan mahasiswa/i Universitas Prima Indonesia fakultas sains dan teknologi yang disebarkan peneliti agar medapatkan hasil klasifikasi dari bebera platfrom *E-Commerce* di Indonesia. Mencari jawaban atas suatu permasalahan melalui pengumpulan fakta - fakta. Pada tahap penelitian ini, peneliti menggunakan jenis penelitian survey dengan penyebaran kuisisioner yang dilakukan secara *online* dan data tersebut akan diolah menggunakan metode

K-NN. Tempat penelitian ini dilaksanakan di Universitas Prima Indonesia yang melibatkan langsung mahasiswa/i Universitas Prima Indonesia pada Februari 2023.

2.2. Kerangka Penelitian

Agar Penelitian ini dapat berjalan dengan benar serta selesai tepat pada waktunya maka terdapat kerangka kerja penelitian. Adapun kerangka kerja pada penelitian ini ialah sebagai berikut ini:



Gambar 1. Flowchart Kerangka Penelitian

2.2.1. Pengumpulan Data

Penulis mengumpulkan data dari kuisisioner dengan teknik survey dimana kuisisioner tersebut disebarakan kepada mahasiswa/mahasiswi Universitas Prima Indonesia. Kuisisioner ini juga sebagai tempat berbagi ide, inspirasi, dan informasi guna mengetahui informasi dari beberapa platform *E-Commerce* di Indonesia.

Pengumpulan data ialah mengumpulkan data penelitian dari sumber data (subyek sampel penelitian). Metode pengumpulan data ini dibutuhkan karena digunakan sebagai dasar untuk mengumpulkan bahan penelitian [11].

1	Timestamp	Nama	Umur	Jenis Kelamin	Domisili	Pekerjaan	Apakah anda mengetahui E-commerce mana yang	Apakah anda menggo	Bagaimana kualitas bara	Bagaimana respon!	
2	03/02/2023 15:13:26	nista kesiya hutasoit	15 - 25 tahun	Perempuan	Medan	Mahasiswa	Iya	Shopee	Event menarik	Sangat baik	Sangat baik
3	03/02/2023 15:27:57	ABDUL WAHID HASIBU	15 - 25 tahun	Laki-laki	Medan	pelajar	Iya	Shopee	Event menarik	Sangat baik	Sangat baik
4	03/02/2023 15:30:56	ABEDNEGO MARULI HC	15 - 25 tahun	Laki-laki	Binjai	Mahasiswa	Iya	Shopee	Event menarik	Sangat baik	Sangat baik
5	03/02/2023 15:33:21	ABRAHAM MANALU	15 - 25 tahun	Laki-laki	Medan	Mahasiswa	Iya	Shopee	Event menarik	Sangat baik	Sangat baik
6	03/02/2023 15:34:30	ANDI SANJAYA SIMANU	15 - 25 tahun	Laki-laki	Binjai	Mahasiswa	Iya	Shopee	Toko yang terjangkau	Baik	Baik
7	03/02/2023 15:35:19	ADE IRMA SURYANI	15 - 25 tahun	Perempuan	Tebing Tinggi	Mahasiswa	Iya	Shopee	Toko yang terjangkau	Sangat baik	Sangat baik
8	03/02/2023 15:36:43	ANDIKA	15 - 25 tahun	Laki-laki	Medan	Mahasiswa	Iya	Shopee	Event menarik	Baik	Sangat baik
9	03/02/2023 15:37:28	ADE LUTHFI RAMADHA	15 - 25 tahun	Perempuan	Medan	Mahasiswa	Iya	Shopee	Toko yang terjangkau	Sangat baik	Sangat baik
10	03/02/2023 15:39:30	ADE SYAHPITRI	15 - 25 tahun	Perempuan	Medan	Mahasiswa	Iya	Shopee	Toko yang terjangkau	Sangat baik	Sangat baik

Gambar 2. Dataset 10 Teratas pada kuisisioner pengambilan data

2.2.2. Analisis Data

Penulis menganalisis data yang telah terkumpul dan data yang diperoleh dari hasil kuisisioner dikalangan mahasiswa/i Universitas Prima Indonesia. Analisis data adalah upaya secara sistematis mengambil bahan untuk meningkatkan pemahaman peneliti tentang studi kasus dan menyajikan hasil kerja. Kita juga harus menganalisa untuk menemukan maknanya [12].

2.2.3. Penyusunan Data

Setelah analisis data selesai penulis akan Menyusun atau mengumpulkan data yang akan diolah lagi ke tahap selanjutnya. Penyusunan data yaitu menyusun data kedalam susunan yang teratur agar bisa dibaca, dilihat, serta mudah dipahami oleh pembaca itu sendiri. Penyusunan data terdiri juga cara mendeteksi adanya kemungkinan kesalahan, serta ketidak konsistenan dan ketidak teraturan maupun ketidak tepatan data yang telah dikumpulkan [13].

2.3. Pengolahan Data Menggunakan Metode KNN

Pengolahan data yang sudah diperoleh akan diolah menggunakan metode K-NN untuk melakukan klasifikasi data [14]. *Python* sendiri merupakan bahasa pemrograman yang *freeware* atau perangkat bebas dalam arti sebenarnya, tidak ada batasan dalam penyalinan atau mendistribusikannya. Lengkap dengan *source code*, *debugger* dan *profiler*, antar muka yang terkandung di dalamnya untuk pelayanan antarmuka, fungsi sistem, *GUI* (antarmuka pengguna)[15].

3. HASIL DAN PEMBAHASAN

Pada bab ini membahas tentang tahapan penelitian mulai dari Persiapan data, Preprocessing data, Pembentukan Model, Melatih data dan melakukan Prediksi serta Memvisualisasikan data untuk mendapatkan hasil prediksi yang baik. Peneliti menggunakan bahasa pemrograman Python melalui Google Collab dengan beberapa library seperti Numpy, Pandas dan Keras dari Tensorflow.

3.1. Data Preprocessing

Langkah pertama, peneliti melakukan persiapan untuk mengolah dataset berupa angka yang akan divisualisasikan dalam bentuk grafik. Import Library merupakan proses memasukkan dan menggunakan fungsi, metode, atau kelas dalam Pustaka atau model eksternal ke dalam program yang sedang dibangun.

```
0d # Import Library yang diperlukan
import pandas as pd
import re
import string
import math
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score
```

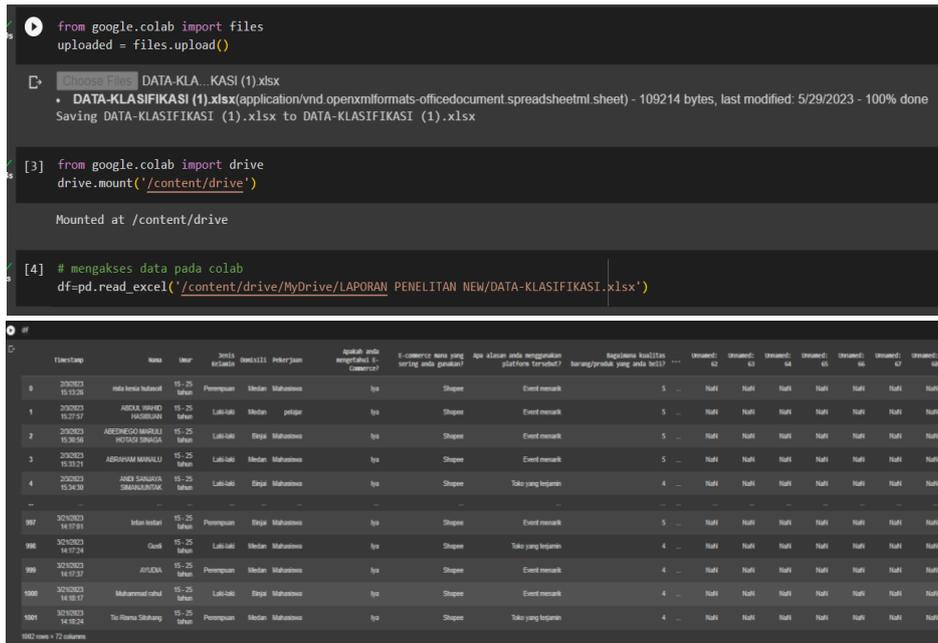
Gambar 3. Kode Python Untuk Import Library

Dalam tahap ini, peneliti menggunakan Platform Google Collab untuk menjalankan Program Python dan mengimport library yang digunakan. Code berikut merupakan beberapa baris perintah untuk mengimpor modul atau pustaka yang diperlukan dalam pemrosesan data dan pembuatan model dengan

menggunakan bahasa pemrograman Python dan beberapa pustaka populer seperti *Pandas*, *Math*, *Numpy*, *Matplotlib*, *seaborn* dan *Scikit-learn*.

3.1.1. Import dan menampilkan Dataset

Setelah dataset diimport dalam bentuk excel atau csv lalu kemudian dimasukkan kedalam google collab kemudian tampilkan dataset akan muncul.

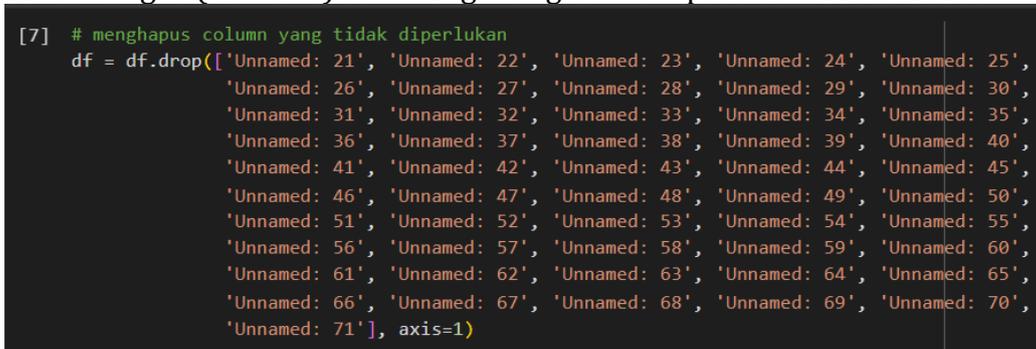


Gambar 4. Tampilan Dataset dan import data di Google Colab

Kode tersebut merupakan bagian dari program Python yang berjalan di Google Colaboratory. Kode ini digunakan untuk mengimpor dataset dari komputer lokal ke Google. Modul "files" diimpor dari pustaka "google.colab". Modul ini digunakan untuk mengunggah file dari komputer lokal ke Google Colab.

3.1.2. Menghapus Kolom Pada Dataset

Pada bagian ini beberapa kolom akan dihapus agar data yang ditampilkan berbentuk angka (numeric) dan menghilangkan NaN pada dataset tersebut.



Gambar 5. Potongan Kode Python untuk menghapus kolom yang tidak diperlukan

3.1.3. Tampilan Dataset yang Sudah di Drop

Kode `df` digunakan untuk menampilkan data farme pada dataset yang telah di drop, dari data yang diatas bisa dilihat dataset tersebut kolom nya menjadi 21 dari 71 kolom dan kolom yang digunakan diubah dalam berbentuk skala int.

Tipe	Platform	E-commerce mana yang sering anda gunakan?	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi	Regulasi			
0	20/02/2023 15:13:26	ada bessa halim	15-25 tahun	Pemampuan	Medan	Mahasiswa	Ya	Shopee	Event menarik	5	5	5	5	5	5	5	5	5	5	5	4	Ada	
1	20/02/2023 15:27:57	ADIA WAHID HADIBUN	15-25 tahun	Laki-laki	Medan	perajar	Ya	Shopee	Event menarik	5	5	4	4	5	5	5	5	5	5	5	5	5	Ada
2	20/02/2023 15:38:58	ABENEDDO HARAHUSIN SINGKA	15-25 tahun	Laki-laki	Siang	Mahasiswa	Ya	Shopee	Event menarik	5	5	5	5	5	5	5	5	5	5	5	5	4	Ada
3	20/02/2023 15:33:21	ABRIHAM MANKU	15-25 tahun	Laki-laki	Medan	Mahasiswa	Ya	Shopee	Event menarik	5	5	5	5	5	5	5	5	5	5	5	5	4	Ada
4	20/02/2023 15:34:30	AKA SHAKIA SIMANJUNTAK	15-25 tahun	Laki-laki	Siang	Mahasiswa	Ya	Shopee	Toko yang bagus	4	5	3	4	4	4	4	4	4	4	4	4	4	Ada
997	32/02/2023 14:17:26	Isan Isdan	15-25 tahun	Pemampuan	Siang	Mahasiswa	Ya	Shopee	Event menarik	5	5	5	5	4	5	5	5	5	5	5	5	5	Ada
998	32/02/2023 14:17:24	Gusti	15-25 tahun	Laki-laki	Medan	Mahasiswa	Ya	Shopee	Toko yang bagus	4	5	4	4	4	4	4	4	4	4	4	4	5	Ada
999	32/02/2023 14:17:27	ARTUDA	15-25 tahun	Pemampuan	Medan	Mahasiswa	Ya	Shopee	Event menarik	4	5	3	4	4	4	4	4	4	4	4	4	4	Ada
1000	32/02/2023 14:18:17	Muhammad alha	15-25 tahun	Laki-laki	Siang	Mahasiswa	Ya	Shopee	Event menarik	4	5	5	5	5	5	5	5	5	5	5	5	5	Ada
1001	32/02/2023 14:18:24	Ta Rama Dibrang	15-25 tahun	Pemampuan	Medan	Mahasiswa	Ya	Shopee	Toko yang bagus	4	5	4	4	4	4	4	4	4	4	4	4	4	Ada

Gambar 6. Potongan Kode Python untuk menampilkan data frame yang telah didrop

3.1.4. Menampilkan Grafik E-commerce yang Paling Sering digunakan

Codingan tersebut menggunakan library matplotlib dan seaborn untuk membuat sebuah grafik statistik (count plot) dari kolom data 'E-commerce yang Paling Sering digunakan?' pada DataFrame 'df'. Dengan menggunakan `seaborn(sns.countplot(data=df, x='E-commerce mana yang sering anda gunakan?'))` dengan sumbu -x mewakili platform e-commerce dan sumbu -y mewakili jumlah responden.

Selanjutnya, `plt.xlabel('E-commerce')` dan `plt.ylabel('Jumlah Responden')` digunakan untuk menetapkan label pada sumbu x dan sumbu y, sedangkan `plt.title('Jumlah E-commerce yang sering digunakan')` digunakan untuk menetapkan judul grafik.



Gambar 7. Grafik E-commerce yang Paling Sering digunakan

3.1.5. Memilih Kolom Dataset yang Akan Diolah

Dalam mengolah data menggunakan KNN, `col_names = df.columns.tolist()` adalah kode perintah dimana kita mengambil daftar nama kolom dari DataFrame 'df' yang diperlukan menggunakan metode `columns` yang mengembalikan Index objek, dan kemudian menggunakan metode `tolist()` untuk mengonversinya menjadi list.

```
[ ] # Memilih kolom yang diperlukan
col_names = df.columns.tolist()
selected_columns = ['Bagaimana kualitas barang/produk yang anda beli?',
                   'Bagaimana respon/kenyamanan chat to seller?',
                   'Bagaimana tampilan dari platform tersebut?',
                   'Bagaimana performa dari platform tersebut?',
                   'Bagaimana privasi atau keamanan akun tersebut?',
                   'Bagaimana pengemasan barang/produk tersebut?',
                   'Bagaimana harga dari produk platform tersebut?',
                   'Apakah anda puas menggunakan platform tersebut?']

selected_dataset = df[[col for col in col_names if col in selected_columns]]
```

Gambar 8. Potongan Kode Python untuk Menentukan Kolom yang di perlukan

3.1.6. Memisahkan Fitur (features) dan Label (target) Dalam Dataset dan Normalisasi Data

Pada kode `X = selected_dataset.iloc[:, :-1].values` kita mengambil semua baris ':' dan hanya kolom terakhir '-1' dari `selected_dataset`. Dan values untuk mengonversi hasil seleksi menjadi bentuk array. Dan kolom pada bagian Y kita memilih kolom terakhir dari 'selected_dataset' sebagai label atau target yang ingin diprediksi.

Langkah ini dilakukan untuk memisahkan fitur dan label karena pada metode K-Nearest Neighbors (KNN), fitur-fitur dataset digunakan untuk menghitung jarak dan memprediksi label berdasarkan tetangga terdekat. Bertujuan untuk memanipulasi dan memproses data dengan lebih mudah, serta mempersiapkannya untuk proses pelatihan dan pengujian model KNN. Normalisasi data dengan mengubah skala atau rentang daya menjadi rentang yang lebih terstandarisasi atau normal.

```
# Memisahkan fitur dan label
X = selected_dataset.iloc[:, :-1].values # Fitur
y = selected_dataset.iloc[:, -1].values # Label

# Normalisasi data
scaler = MinMaxScaler()
X_normalized = scaler.fit_transform(X)
```

Gambar 9. Potongan Kode Python untuk Memisahkan Fitur dan Label dan Normalisasi Data

3.2. Memisahkan Data Train, Data Test dan Melatih Metode KNN, Melakukan Prediksi serta Menghitung Akurasi Model

Dengan memisahkan data Train dan data Test, kita dapat menguji kinerja model pada data yang belum pernah dilihat sebelumnya, yang akan memberikan perkiraan lebih realistis tentang seberapa baik model tersebut dapat digunakan untuk memprediksi data baru.

Dalam melatih model KNN kami membuat Objek KNN dengan $k = 5$ menggunakan **'KNeighborsClassifier'** kemudian melatih model KNN menggunakan data training menggunakan metode **'fit'**. Setelah melatih model, kami menggunakan model yang dilatih untuk memprediksi label untuk data test menggunakan metode **'predict'**. Akhirnya akan dihitung akurasi prediksinya menggunakan fungsi **'accuracy_score'** untuk membandingkan label yang diprediksi dengan label sebenarnya pada data set. Lalu hasil dari menghitung akurasi model tersebut adalah 84.08% Akurasi KNN.

```
[ ] # Memisahkan data train dan data test
X_train, X_test, y_train, y_test = train_test_split(X_normalized, y, test_size=0.2, random_state=42)

[ ] # Melatih model KNN
k = 5 # Nilai K
knn = KNeighborsClassifier(n_neighbors=k)
knn.fit(X_train, y_train)

KNeighborsClassifier
KNeighborsClassifier()

[ ] # Melakukan prediksi pada data test
y_pred = knn.predict(X_test)

# Menghitung akurasi model
accuracy = accuracy_score(y_test, y_pred) * 100
print("Akurasi KNN: {:.2f}%".format(accuracy))

Akurasi KNN: 84.08%
```

Gambar 10. Data Train, Test dan Melatih Metode KNN

3.3. Menghitung Precision, Recall dan F-1 Score

Precision mengukur seberapa baik model dalam mengidentifikasi positif yang benar dari semua prediksi positif. Recall mengukur seberapa baik model dalam menemukan semua *instance* positif. F1-Score adalah metrik yang menggabungkan precision dan recall menjadi satu score yang menverminkan keseimbangan antara keduanya. Dapat dilihat hasil dari mengukur ketiga model tersebut pada gambar diatas.

```
[ ] from sklearn.metrics import precision_score, recall_score, f1_score

# Menghitung precision
precision = precision_score(y_test, y_pred, average='macro')
print("Precision:", precision)

# Menghitung recall
recall = recall_score(y_test, y_pred, average='macro')
print("Recall:", recall)

# Menghitung F1-score
f1 = f1_score(y_test, y_pred, average='macro')
print("F1-score:", f1)

Precision: 0.8308191808191808
Recall: 0.854398544520548
F1-score: 0.8352121336339413
```

Gambar 11. Menghitung Precision, Recall, dan F1-Score

3.4. Menampilkan Visualisasi Confusion Matrix Dalam Bentuk Grafik Heatmap

Kami menggunakan `'sns.heatmap'` dari library seaborn untuk membuat heatmap dari confusion matrix. Parameter `'annot=True'` digunakan untuk menampilkan angka-angka pada setiap sel dalam heatmap yang mewakili jumlah instance. Parameter `'cmap=Blues'` digunakan untuk mengukur skema warna heatmap. Parameter `'fmt="d"'` digunakan untuk mengatur format angka pada sel-sel heatmap sebagai bilangan bulat.

Selanjutnya, kami menggunakan fungsi-fungsi dari matplotlib untuk memberikan judul pada heatmap (`'plt.title'`), memberikan label pada sumbu x dan y (`'plt.xlabel'` dan `'plt.ylabel'`) dan menampilkan heatmap (`'plt.show()'`). Pastikan `'y_teset'` dan `'y_pred'` dengan label sebenarnya dan label prediksi yang sesuai dengan data test.



Gambar 12. Menampilkan Visualisasi confusion Matrix Dalam Bentuk Grafik Heatmap

3.5. Memilih dan Menghitung Persentase Data Terhadap Kepuasan Platfrom Shopee

Kode `shopee_df = df.query("E-commerce mana yang sering anda gunakan?" = "Shopee")` ini akan menghasilkan dataframe shopee yang berisi semua baris dari dataframe di mana nilai kolom 'E-commerce mana yang sering anda gunakan?' adalah "Shopee". Kemudian Kode `value_counts = shopee_df['Apakah anda puas menggunakan platfrom tersebut?'].value_counts()` ini akan menghitung jumlah kemunculan setiap nilai kepuasan dalam kolom tersebut, kemudian menghitung total kepuasan berdasarkan aturan tertentu (dalam hal ini, jumlah kemunculan nilai kepuasan dikalikan dengan nilai kepuasan itu sendiri), selanjutnya

menghitung persentase kepuasan dengan membagi total kepuasan dengan jumlah total kemungkinan nilai kepuasan (jumlah pelanggan dikali dengan 5, karena skala kepuasan adalah dari 1 hingga 5). Lalu hasil dari menghitung persentase kepuasan pelanggan terhadap shopee adalah 82.98%

```
[23] #ambil data pelanggan yang menggunakan shopee
shopee_df = df.query("E-commerce mana yang sering anda gunakan? == 'Shopee'")
shopee_df
```

Timestamp	Nama	Umur	Jenis Kelamin	Domisili	Pekerjaan	Apakah anda mengetahui E-commerce?	E-commerce mana yang sering anda gunakan?	Apa alasan anda menggunakan platform tersebut?	Bagaimana kualitas barang/produk yang anda beli?	Bagaimana tampilan dari platform tersebut?	Bagaimana performansi dari platform tersebut?	Bagaimana privasi atau keamanan akun tersebut?	Bagaimana pengemasan barang/produk tersebut?	Bagaimana harga dari produk platform tersebut?	Apakah anda puas menggunakan platform tersebut?
0	risa koesia hufasoft	15-25 tahun	Perempuan	Medan	Mahasiswa	ya	Shopee	Event menarik	5	5	5	5	5	5	4
1	ABDUL WAHID HASIBUAN	15-25 tahun	Laki-laki	Medan	pelajar	ya	Shopee	Event menarik	5	5	4	4	5	5	5
2	ABEDNEGO MARULI HOTASI SIRAGA	15-25 tahun	Laki laki	Binjai	Mahasiswa	ya	Shopee	Event menarik	5	5	5	5	5	5	4
3	ABRAHAM MANALU	15-25 tahun	Laki laki	Medan	Mahasiswa	ya	Shopee	Event menarik	5	5	5	5	5	5	4
4	ANDI SANJAYA SIMANLINTAK	15-25 tahun	Laki laki	Binjai	Mahasiswa	ya	Shopee	Toko yang terjangkau	4	3	4	4	4	4	4
997	Intan Iestari	15-25 tahun	Perempuan	Binjai	Mahasiswa	ya	Shopee	Event menarik	5	5	5	4	5	5	5
998	Guili	15-25 tahun	Laki laki	Medan	Mahasiswa	ya	Shopee	Toko yang terjangkau	4	5	4	4	4	4	5
999	AYUDIA	15-25 tahun	Perempuan	Medan	Mahasiswa	ya	Shopee	Event menarik	4	3	4	4	4	4	4
1000	Muhammad rahul	15-25 tahun	Laki laki	Binjai	Mahasiswa	ya	Shopee	Event menarik	4	5	5	5	5	5	5
1001	Tio Rioma Siburang	15-25 tahun	Perempuan	Medan	Mahasiswa	ya	Shopee	Toko yang terjangkau	4	5	4	4	4	4	4

```
[24] # Menghitung jumlah kemunculan setiap nilai kepuasan
value_counts = shopee_df['Apakah anda puas menggunakan platform tersebut?'].value_counts()

# Menghitung total kepuasan berdasarkan aturan
total = (value_counts * value_counts.index).sum()

# Menghitung persentase kepuasan
percentage = (total / (len(shopee_df) * 5)) * 100

# Menampilkan hasil
print("Kepuasan pelanggan terhadap E-commerce Shopee: %.2f%%" % percentage)

Kepuasan pelanggan terhadap E-commerce Shopee: 82.98%
```

Gambar 13. Menghitung persentase kepuasan shopee

3.6. Menghitung Persentase Data Terhadap Kepuasan Platform TokoPedia

Kode `value_counts = tokopedia_df['Apakah anda puas menggunakan platform tersebut?'].value_counts()` ini akan menghitung jumlah kemunculan setiap nilai kepuasan dalam kolom tersebut, kemudian menghitung total kepuasan berdasarkan aturan tertentu (dalam hal ini, jumlah kemunculan nilai kepuasan dikalikan dengan nilai kepuasan itu sendiri), selanjutnya menghitung persentase kepuasan dengan membagi total kepuasan dengan jumlah total kemungkinan nilai kepuasan (jumlah pelanggan dikali dengan 5, karena skala kepuasan adalah dari 1 hingga 5). Lalu hasil dari menghitung persentase kepuasan pelanggan terhadap TokoPedia adalah 80.30%.

Berdasarkan hasil pelatihan dan perhitungan pengolahan data model KNN didapatkan dua e-commerce terbaik sesuai dengan penilaian responden yaitu Shopee dan TokoPedia, dengan persentase Shopee 82.98% dari 779 jumlah responden dan TokoPedia 80.30% dari 133 jumlah responden. Dengan menggunakan data keseluruhan kuesioner sebanyak 1001 data.

```

# Menghitung jumlah kemunculan setiap nilai kepuasan
value_counts = tokopedia_df['Apakah anda puas menggunakan platfrom tersebut?'].value_counts()

# Menghitung total kepuasan berdasarkan aturan
total = (value_counts * value_counts.index).sum()

# Menghitung persentase kepuasan
percentage = (total / (len(tokopedia_df) * 5)) * 100

# Menampilkan hasil
print("Kepuasan pelanggan terhadap E-commerce Tokopedia: %.2f%%" % percentage)

```

Gambar 14. Menghitung persentase kepuasan tokopedia

Dengan memprediksi kepuasan pelanggan terhadap masing-masing platform yang lebih akurat, dapat disimpulkan bahwa Shopee memiliki nilai kepuasan pelanggan dengan nilai 82.98% dengan jumlah 779 responden, TokoPedia 80.30% dengan jumlah 133 responden, Blibli 80.00% dengan jumlah 15 responden, BukaLapak 80.00% dengan jumlah 23 responden, Lazada 80.78% dengan jumlah 51 responden.

Tabel 1. Hasil Kepuasan Pelanggan

Platafrom	Kepuasan Pelanggan Terhadap E-Commerce (persentase)	Jumlah Responden
Shopee	82.98%	779
TokoPedia	80.30%	133
Blibli	80.00%	15
BukaLapak	80.00%	23
Lazada	80.78%	51

4. SIMPULAN

Dalam menentukan hasil nilai akurasi terhadap kepuasan pelanggan terhadap e-commerce dan pelatihan model KNN, dapat disimpulkan bahwa model KNN baik dalam memprediksi kepuasan pelanggan dengan menggunakan data kuesioner sebanyak 1002 dataset responden. Berdasarkan pelatihan model KNN menunjukkan bahwa metode ini memiliki nilai keakuratan yang cukup baik dengan akurasi KNN adalah 84.08%.

DAFTAR PUSTAKA

- [1] Nugroho Andy 2022 "Pengertian E-Commerce, contoh, Jenis dan perkembangannya": <https://qwords.com/blog/pengertian-e-commerce/> 26 April 2022.
- [2] Warta Ekonomi 2019 "Pertumbuhan E-Commerce pesat di Indonesia" : [https://wartaekonomi.co.id/read216302/pertumbuhan-e-commercepesat?](https://wartaekonomi.co.id/read216302/pertumbuhan-e-commercepesat?di-indonesia) di-indonesia 9/19 Febuari 2019.
- [3] Lindwina Andrea 2021 "Penggunaan E-Commerce Indonesia tertinggi di Dunia": <https://databoks.katadata.co.id/datapublish/2021/06/04/penggunaan-e-commerce-indonesia-tertinggi-di-dunia> /4 Juni 2021.

- [4] Setiawan Rony 2021 "Apa itu Data Mining dan bagaimana Metodenya ?" / [https://www.dicoding.com/blog/apa-itu-data-mining/30 Oktober 2021](https://www.dicoding.com/blog/apa-itu-data-mining/30%20Oktober%202021)'.
- [5] Dika, Onggito et al 2021. " Implementasi Data Mining Menggunakan Metode K-Nearest Neighbor Untuk Menentukan Stok Obat-Obatan Pada Apotek: Studi Kasus Apotek Salaam" Yogyakarta : Universitas PGRI 2021.
- [6] AdminL2PM 2023 "Algoritma K-Nearest Neighbors (KNN) - Pengertian Dan Penerapan"<https://lp2m.uma.ac.id/2023/02/16/algoritma-k-nearest-neighbors-knn-pengertian-dan-penerapan/> 16 Februari 2023.
- [7] Wiyli Yusanti 2012 "Algoritma K-Nearest Neighbour untuk Memprediksi Harga Jual Tanah" Surabaya : Jurnal Matematika, Statistika dan Komputasi., Vol.9 No.1, 57-68, 2012.
- [8] Dika, Onggito et al 2021. " Implementasi Data Mining Menggunakan Metode K-Nearest Neighbor Untuk Menentukan Stok Obat-Obatan Pada Apotek: Studi Kasus Apotek Salaam" Yogyakarta : Universitas PGRI.
- [9] Kawasati Risky 2021 "Teknik pengumpulan data Kualiatif" Sorong : Sekolah Tinggi Agama Islam Negeri (STAIN).
- [10] Rijali Ahmad 2018 "Analisis data Kualiatif" Banjarmasin: UIN Antasari., Vol.7 No.33, 2018.
- [11] Wiyli Yusanti 2012 "Algoritma K-Nearest Neighbour untuk Memprediksi Harga Jual Tanah" Surabaya : Jurnal Matematika, Statistika dan Komputasi., Vol.9 No.1, 57-68, 2012.
- [12] Rombang Mathew Raphael Clinton, Rizal Sengkey 2019 "Purwarupa Sistem Daftar Pelanggaran Lalulintas Berbasis Mini-Komputer Raspberry Pi" Manado : Jurnal Teknik Elektro dan Komputer., Vol.8 No.3, 2019.