

Klasifikasi Kepuasan Pengguna Aplikasi Kopi Kenangan Menggunakan Algoritma Decision Tree

¹Enggar Rahma Agustin, ²Intan Sulistyowati, ³Erna Daniati

¹⁻³Sistem Informasi, Universitas Nusantara PGRI Kediri

E-mail: [1rahmaenggar5@gmail.com](mailto:rahmaenggar5@gmail.com) [2intansulistyowati27@gmail.com](mailto:intansulistyowati27@gmail.com)

[3ernadaniati@unpkediri.ac.id](mailto:ernadaniati@unpkediri.ac.id)

Penulis Korespondens : Erna Daniati

Abstrak— Di tengah persaingan industri minuman yang semakin digital, aplikasi Kopi Kenangan hadir sebagai solusi praktis untuk memesan minuman favorit para pelanggan. Namun, kenyamanan dan kemudahan yang ditawarkan aplikasi belum tentu menjamin kepuasan pengguna. Untuk itu, penting dilakukan analisis lebih lanjut guna memahami sejauh mana aplikasi ini mampu memenuhi harapan penggunanya. Penelitian ini menggunakan algoritma Decision Tree untuk mengklasifikasikan tingkat kepuasan pengguna berdasarkan berbagai faktor, seperti kemudahan penggunaan, kecepatan layanan, hingga fitur yang tersedia. Melalui pendekatan ini, diperoleh gambaran mengenai faktor-faktor utama yang memengaruhi kepuasan, serta masukan berbasis data untuk pengembangan layanan yang lebih baik ke depannya.

Kata Kunci—kepuasan pengguna, aplikasi Kopi Kenangan, klasifikasi, Decision Tree, layanan digital

Abstract—In the midst of competition in the increasingly digital beverage industry, the Kopi Kenangan application is here as a practical solution to order customers' favorite drinks. However, the convenience and convenience offered by the application does not necessarily guarantee user satisfaction. For this reason, it is important to conduct further analysis to understand the extent to which this application is able to meet the expectations of its users. This study uses the Decision Tree algorithm to classify user satisfaction levels based on various factors, such as ease of use, speed of service, and available features. Through this approach, an overview of the main factors that affect satisfaction is obtained, as well as data-driven inputs for better service development in the future.

Keywords—user satisfaction, Kopi Kenangan application, classification, Decision Tree, digital services

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

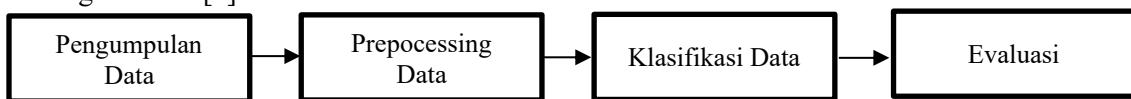
Perkembangan teknologi digital telah secara signifikan mengubah masyarakat dalam mengonsumsi makanan dan minuman, terutama melalui aplikasi mobile. Salah satu aplikasi yang sangat populer di Indonesia adalah "Kopi Kenangan". Dikelola oleh PT Bumi Berkah Boga, kopi kenangan merupakan kedai kopi lokal independen yang tengah mengalami pertumbuhan pesat di tanah air. Mereka menawarkan beragam pilihan minuman kopi berkualitas, mulai dari espresso hingga non-espresso, serta berbagai pastry seperti roti dan kue kering. PT Bumi Berkah Boga

mengintegrasikan teknologi aplikasi dalam operasionalnya dengan mengadopsi konsep "New Retail", di mana perbedaan antara interaksi online dan offline semakin samar. Dengan aplikasi Kopi Kenangan, konsumen dapat memesan kopi dan berbagai minuman lain tanpa perlu mengantri di outlet. Aplikasi ini juga dilengkapi dengan beragam fitur menarik, termasuk promosi, cashback, Kenangan VIP, Daily CheckIn, layanan pesan-antar, chat langsung, serta berbagai metode pembayaran yang lengkap. [1] Penelitian ini akan mencakup pengumpulan dan pemrosesan data dari ulasan di Google Play Store, menggunakan Algoritma Decision Tree dengan bantuan software Google Colab.[2]

Ulasan pengguna memiliki peranan yang sangat penting dalam mengidentifikasi masalah serta mengevaluasi kinerja aplikasi. Melalui penerapan metode decision tree guna mengelompokan respon subjektif dari komentar konsumen, PT Bumi Berkah mencapai pemahaman menyeluruh mengenai tanggapan dan kepuasan konsumen terhadap aplikasi tersebut. Perolehan dari pengkajian ini, akan berkontribusi pada pengembangan dan perbaikan aplikasi Kopi Kenangan, agar bisa menangani keperluan dan keinginan pelanggan di zaman modern. PT Bumi Berkah berkomitmen terus mengembangkan beragam pembaruan kreatif serta meningkatkan infrastruktur layanan yang kreatif serta kolaboratif.

II. METODE

Peneliti menggunakan kumpulan data yang diambil dari aplikasi Google Playstore dan memanfaatkan Google Colab sebagai alat untuk melakukan ekstraksi data. Setelah mengumpulkan data yang diperlukan, langkah berikutnya adalah melakukan proses pembersihan data (Preprocessing). Selanjutnya, klasifikasi akan dilakukan dengan menerapkan metode Decision Tree, yang diakhiri dengan tahap evaluasi, mengikuti alur penelitian yang ditunjukkan dalam gambar. 1.[2]



gambar1. Proses penelitian

2.1 Pengambilan Data

Peneliti memanfaatkan Google Colab sebagai sarana untuk mengumpulkan data pada tahap awal penelitian. Data yang diperoleh berasal dari ulasan pengguna aplikasi Kopi Kenangan yang terdapat di Playstore.

2.2 Preprocessing Data

Preprocessing data merupakan tahap pembersihan data dengan mengecek data yang duplikat serta memastikan konsistensi [3]. Preprocessing data penting dalam analisis data mining untuk membersihkan, mengubah format, dan mempersiapkan data agar lebih mudah dan akurat. [4][5]

- Data Transformation

Tahap transformasi bertujuan untuk memproses data dengan skala, melakukan normalisasi, dan mereduksi. Salah satu tahapan dalam proses transformasi data yaitu mengubah tipe data dari non-numerik menjadi numerik, selain itu, pada tahap ini juga dilakukan penanganan terhadap nilai yang hilang (missing value) dan data penyimpangan (outlier).[6]

2.3 Klasifikasi Data

Klasifikasi adalah teknik penambangan data yang mengelompokkan data ke dalam kategori atau kelas yang telah ditentukan.[7][8] Proses klasifikasi data dapat dilakukan dengan berbagai metode, salah satunya adalah menggunakan pohon keputusan (Decision Tree).[9]

2.4 Decision Tree

Decision Tree adalah Struktur flowchart yang berbentuk seperti pohon, di mana setiap bagian disebut sebagai cabang. Setiap simpul internal, terdapat tes atribut, sedangkan setiap cabang menunjukkan hasil tes, dan simpul daun menunjukkan distribusi kelas.[10][11]

2.5 Data Mining

Data mining adalah suatu proses yang memanfaatkan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengekstrak dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang relavan dari berbagai basis data yang besar.[12]

2.6 Web Scraping

Proses pengambilan dataset dilakukan dengan cara web scraping. Proses web scraping menggunakan bahasa pemrograman Python dengan Google Colab sebagai coding environment-nya. Data yang berhasil dikumpulkan terdiri dari 829 ulasan pengguna mengenai aplikasi kopi kenangan yang terdapat di Google Play store. Contoh ulasan yang diperoleh dari proses web scraping dapat dilihat pada Gambar 2.

2.7 Evaluasi

Tahap Evaluasi merupakan tahap paling akhir yang dilakukan untuk mengukur kinerja model klasifikasi dengan menghitung nilai accuracy, precision, dan recall.

III. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Dataset diperoleh dengan cara *scraping data* menggunakan bahasa *python* kemudian disimpan dalam bentuk csv/excel. Alur visualisasi pengumpulan data ditunjukkan pada gambar 2.



Gambar. 2 Proses Pengumpulan Data

Scraping merupakan metode yang digunakan untuk mengumpulkan informasi dari situs web tertentu, baik secara manual maupun otomatis. Teknik ini berfokus pada pengambilan data dan ekstrasi data.[13] Setelah proses scraping selesai, kami menemukan sebanyak 829 data yang memberikan beragam ulasan dan bintang. Hasil dataset tersebut dapat dilihat pada Gambar 3

Gambar 3. Tampilan Atas Dataset

3.2 Preprocessing Data

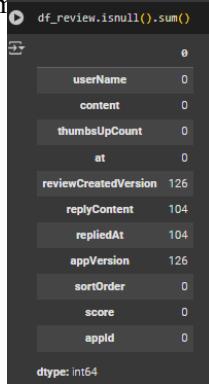
Preprocessing data adalah tahap pembersihan data dengan memeriksa duplikat data dan tingkat konsistensi [14]. Dalam proses preprocessing data memiliki tahapan yang perlu dilakukan seperti melihat missing value atau outliersnya.[15]

```
df_review.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 829 entries, 0 to 828
Data columns (total 13 columns):
 #   Column          Non-Null Count  Dtype  
--- 
 0   reviewId        829 non-null    object  
 1   username        829 non-null    object  
 2   userimage       829 non-null    object  
 3   content         829 non-null    object  
 4   score           829 non-null    int64  
 5   thumbsUpCount  829 non-null    int64  
 6   reviewCreatedAt 783 non-null    object  
 7   at               829 non-null    object  
 8   replyContent    725 non-null    object  
 9   repliedAt       725 non-null    object  
 10  appVersion       783 non-null    object  
 11  sortOrder        829 non-null    object  
 12  appId            829 non-null    object  
dtypes: int64(2), object(11)
memory usage: 84.3+ KB
```

Gambar 4. Isi Dataset

Proses Trasformation Data

Pada proses transformation data, tahapan yang dilakukan adalah mencari missing value di dalam data.



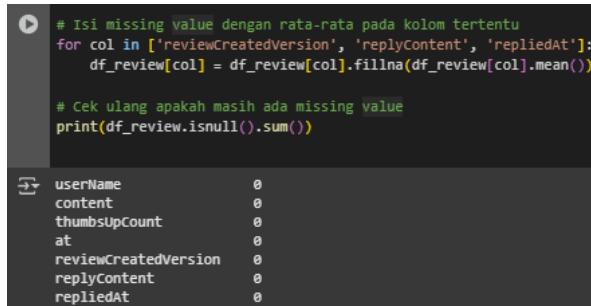
```
df_review.isnull().sum()
```

	0
userName	0
content	0
thumbsUpCount	0
at	0
reviewCreatedVersion	126
replyContent	104
repliedAt	104
appVersion	126
sortOrder	0
score	0
appId	0

dtype: int64

Gambar 5. Kolom Missing Value

Pada Gambar 5, terdapat beberapa kolom yang memiliki data missing value yang cukup signifikan. Oleh karena itu, perlu perbaikan pada data tersebut dengan menambahkan mean(rata-rata) ke dalam kolom tersebut, sehingga diharapkan tidak ada lagi missing value setelah proses ini. Kode untuk penambahan nilai mean (rata-rata) dapat dilihat pada Gambar 6.



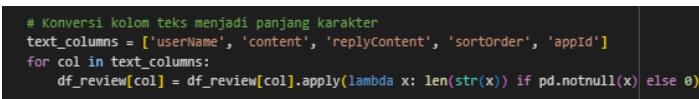
```
# Isi missing value dengan rata-rata pada kolom tertentu
for col in ['reviewCreatedVersion', 'replyContent', 'repliedAt']:
    df_review[col] = df_review[col].fillna(df_review[col].mean())

# Cek ulang apakah masih ada missing value
print(df_review.isnull().sum())
```

	0
userName	0
content	0
thumbsUpCount	0
at	0
reviewCreatedVersion	0
replyContent	0
repliedAt	0

Gambar 6. Kode Penambahan Mean

Setelah melakukan penambahan mean (rata-rata), langkah selanjutnya yaitu mentransformation data dari kolom yang bertipe objek (teks) kedalam numerik (int). Kode untuk perubahan tipe numerik (int) dapat dilihat pada Gambar 7



```
# Konversi kolom teks menjadi panjang karakter
text_columns = ['userName', 'content', 'replyContent', 'sortOrder', 'appId']
for col in text_columns:
    df_review[col] = df_review[col].apply(lambda x: len(str(x)) if pd.notnull(x) else 0)
```

Gambar 7. Kode Perubahan Tipe Data

```
df_review.info()  
  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 829 entries, 0 to 828  
Data columns (total 13 columns):  
 #   Column      Non-Null Count  Dtype     
---    
 0   reviewId    829 non-null   object    
 1   userName    829 non-null   object    
 2   userImage   829 non-null   object    
 3   content     829 non-null   object    
 4   score       829 non-null   int64    
 5   thumbsUpCount 829 non-null   int64    
 6   reviewCreatedVersion 783 non-null   object    
 7   at          829 non-null   object    
 8   replyContent 725 non-null   object    
 9   repliedAt   725 non-null   object    
 10  appVersion   783 non-null   object    
 11  sortOrder   829 non-null   object    
 12  appId       829 non-null   object    
dtypes: int64(2), object(11)  
memory usage: 84.3+ KB
```

Gambar 8. Sebelum Perubahan

Dari Gambar 8 (Sebelum Perubahan) menunjukkan bahwa mayoritas kolom dalam DataFrame masih bertipe data object, termasuk beberapa kolom yang seharusnya bersifat numerik, seperti thumbsUpCount, replyCount, dan reviewCreatedVersion.

```
df_review.info()  
  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 829 entries, 0 to 828  
Data columns (total 11 columns):  
 #   Column      Non-Null Count  Dtype     
---    
 0   userName    829 non-null   int64    
 1   content     829 non-null   int64    
 2   thumbsUpCount 829 non-null   int64    
 3   at          829 non-null   int64    
 4   reviewCreatedVersion 829 non-null   int64    
 5   replyContent 829 non-null   int64    
 6   repliedAt   829 non-null   int64    
 7   appVersion   829 non-null   int64    
 8   sortOrder   829 non-null   int64    
 9   score       829 non-null   int64    
 10  appId       829 non-null   int64    
dtypes: int64(11)  
memory usage: 71.4 KB
```

Gambar 9. Setelah Perubahan

Sementara itu pada Gambar 9 (Setelah Perubahan) memperlihatkan tampilan setelah perubahan tipe data dilakukan. Kolom-kolom yang sebelumnya berformat object kini telah berhasil dikonversi menjadi tipe data yang lebih sesuai, seperti int64 untuk data numerik dan datetime64[ns] untuk informasi tanggal/waktu (at).

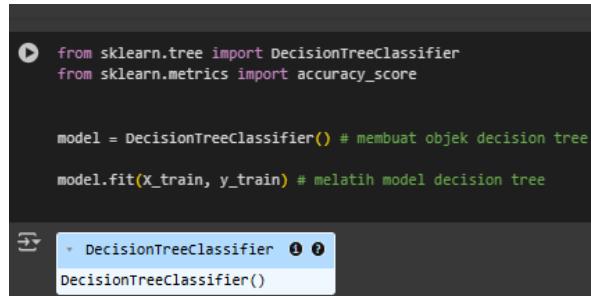
3.3 Klasifikasi Data

Sebelum memasuki tahap klasifikasi data, tahap awal yang perlu dilakukan yaitu pengembangan model klasifikasi melalui pembagian data (data splitting). Dataset dibagi menjadi dua bagian utama, data latih (training set) dan data uji (testing set) dengan proporsi 50:50. Proses ini dilakukan menggunakan fungsi train_test_split dari sklearn.model_selection, untuk memastikan bahwa data yang digunakan untuk pelatihan dan pengujian model terpisah dengan baik. Variabel input (fitur) disimpan di dalam X, sedangkan label target kepuasan pengguna disimpan dalam y. Dengan cara ini, model akan dilatih menggunakan sebagian besar data dan diuji pada data yang sebelumnya tidak pernah diakses, sehingga hasil evaluasi tetap dapat mencerminkan performa model secara lebih adil dan realistik. Kode untuk pengujian data dapat dilihat pada Gambar 10.

```
[ ] # import library train_test_split
      from sklearn.model_selection import train_test_split

      # data splitting
      x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.5)
```

Gambar 10. Tahapan Split Data



```
▶  from sklearn.tree import DecisionTreeClassifier
      from sklearn.metrics import accuracy_score

      model = DecisionTreeClassifier() # membuat objek decision tree
      model.fit(x_train, y_train) # melatih model decision tree
```

Gambar 11. Pelatihan Model Dengan Decision Tree

Pada Gambar 11, langkah krusial berikutnya adalah proses pelatihan model (training) yang dilakukan dengan memanggil fungsi `model.fit(X_train, y_train)`. Fungsi ini bertugas untuk memproses data pelatihan dan mengevaluasi nilai atribut pada `X_train`, serta mempelajari pola-pola yang ada dalam data tersebut berdasarkan label yang sesuai pada `y_train`. Hasil dari proses ini adalah terbentuknya struktur pohon keputusan, di mana setiap node merepresentasikan pertanyaan atau kondisi logis terkait atribut input, hingga akhirnya mengarah pada suatu keputusan klasifikasi. Setelah itu, kita melanjutkan ke proses prediksi untuk data uji, seperti yang ditunjukkan pada Gambar 12.

```
] y_pred = model.predict(x_test)
```

Gambar 12. Prediksi Data Uji

3.4 Evaluasi

Tahap evaluasi menjadi tahap akhir untuk mengetahui hasil dari klasifikasi yang telah dilakukan pada proses diatas.

	precision	recall	f1-score	support
1.0	0.64	0.46	0.54	102
2.0	0.35	0.44	0.39	48
3.0	0.60	0.62	0.61	94
4.0	0.48	0.62	0.54	68
5.0	0.86	0.83	0.84	103
accuracy			0.61	415
macro avg	0.59	0.59	0.58	415
Weighted avg	0.63	0.61	0.61	415

Gambar 13. Hasil dari uji evaluasi

Gambar 13 memperlihatkan hasil evaluasi model Decision Tree dengan tingkat akurasi mencapai 61%, ini berarti sekitar 61% dari total 415 data berhasil diklasifikasikan dengan

tepat oleh model tersebut. Namun, performa model ini bervariasi antara kelas satu dengan yang lainnya. Kelas 5 menonjol dengan precision sejumlah 0.86, tingkat keterpanggilan 0.83, dan index-f1 0.84 yang memperlihatkan dimana metode cukup handal dalam mengenali dan memprediksi kelas ini. Sebaliknya, model mengalami kesulitan dalam mengolah kelas 2, yang hanya mendapatkan f1-score sebesar 0.39, sehingga menjadi sorotan untuk perbaikan lebih lanjut.

Secara keseluruhan, nilai f1-score makro yang mencapai 0.58 menunjukkan bahwa performa model masih berada pada tingkat sedang jika semua kelas dianggap memiliki kepentingan yang sama. Sementara itu, nilai f1-score tertimbang sebesar 0.61 memberikan gambaran tentang keseluruhan model dengan mempertimbangkan jumlah data di setiap kelas.

IV. KESIMPULAN

Dengan memanfaatkan algoritma Decision Tree yang menganalisis data ulasan dari Google Play Store, penelitian ini berhasil mengkategorikan tingkat kepuasan pengguna terhadap aplikasi Kopi Kenangan. Persepsi pengguna dipengaruhi oleh berbagai fitur, seperti kemudahan penggunaan, kecepatan layanan, dan atribut aplikasi itu sendiri. Proses ini melibatkan langkah-langkah preprocessing, transformasi data, dan klasifikasi. Hasil evaluasi model menunjukkan bahwa tingkat akurasi klasifikasi mencapai 61%, dengan kelas ulasan yang mendapatkan skor tertinggi. Temuan ini menunjukkan bahwa pendekatan Decision Tree dapat memberikan wawasan berharga mengenai faktor-faktor yang memengaruhi kepuasan pengguna. Lebih dari itu, penelitian ini dapat dijadikan referensi untuk merancang aplikasi yang lebih responsif terhadap kebutuhan pelanggan di era layanan digital. Penelitian ini juga mendukung penerapan data mining dalam evaluasi kepuasan layanan digital, khususnya dalam sektor makanan dan minuman berbasis aplikasi.

UCAPAN TERIMAKASIH

Dengan penuh rasa syukur, kami mengucapkan terima kasih kepada Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya yang memungkinkan kami menyelesaikan penelitian ini dengan baik. Ucapan terima kasih yang sama kami sampaikan kepada dosen pembimbing kami, yang dengan sabar memberikan arahan dan bimbingan. Kami menghargai setiap pihak yang telah berkontribusi, baik secara langsung maupun tidak langsung dalam kelancaran penelitian ini. Kami juga mengucapkan terima kasih kepada para pengguna aplikasi Kopi Kenangan, karena ulasan mereka menjadi sumber data utama dalam penelitian ini. Semoga

hasil dari penelitian ini dapat memberikan manfaat, khususnya dalam pengembangan layanan digital di masa yang akan datang.

DAFTAR PUSTAKA

- [1] F. I. Komputer and U. D. Nuswantoro, "Menggunakan Metode Decision Tree Pada," vol. 7, no. 6, pp. 3830–3833, 2010.
- [2] D. S. Al Azkiah, E. Erizal, and F. N. Hikmah, "Perbandingan Algoritma SVM dan Decision Tree Dalam Klasifikasi Kepuasan Pengguna Aplikasi Migo E-Bike di Playstore," *Build. Informatics, Technol. Sci.*, vol. 6, no. 1, pp. 158–167, 2024, doi: 10.47065/bits.v6i1.5344.
- [3] E. Herdika *et al.*, "Klasifikasi Risiko Kambuhnya Kanker Tiroid Menggunakan Algoritma Random Forest," vol. 8, pp. 63–74, 2024.
- [4] D. Fitriono, S. A. Wardani, M. N. B. Al, A. Ristyawan, and E. Daniati, "Perbandingan Metode Algoritma Decission Tree dan K-Nearest Neighbors untuk Memprediksi Kualitas Air yang dapat dikonsumsi," vol. 8, pp. 475–484, 2024.
- [5] Ihsan Zulfahmi, "Analisis Sentimen Aplikasi PLN Mobile Menggunakan Metode Decission Tree," *J. Penelit. Rumpun Ilmu Tek.*, vol. 3, no. 1, pp. 11–21, 2023, doi: 10.55606/juprit.v3i1.3096.
- [6] M. R. Muzaki, I. M. Nurfajriana, P. A. Ilahi, A. Ristyawan, and E. Daniati, "Implementasi Data Mining dengan Algoritma K- Nearest Neighbors untuk Memprediksi Risiko Diabetes Menggunakan Chatbot Telegram," vol. 8, pp. 310–318, 2024.
- [7] S. Bahri and A. Lubis, "Metode Klasifikasi Decision Tree Untuk Memprediksi Juara English Premier League," *J. Sintaksis*, vol. 2, no. 1, pp. 63–70, 2020.
- [8] F. K. Wardhani, N. Kamilatutsaniya, A. Alamsyah, E. Daniati, and A. Ristyawan, "Perbandingan Algoritma Naive Bayes Dan Desicion Tree Dalam Pengujian Data Anemia Menggunakan," *Inotek*, vol. 8, 2024, [Online]. Available: <https://proceeding.unpkediri.ac.id/index.php/inotek/>
- [9] A. Tangkelayuk, "The Klasifikasi Kualitas Air Menggunakan Metode KNN, Naïve Bayes, dan Decision Tree," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 2, pp. 1109–1119, 2022, doi: 10.35957/jatisi.v9i2.2048.
- [10] K. Amarya, A. C. A. G, R. Achmad, E. Daniati, and A. Ristyawan, "Analisa Perbandingan Algoritma Classification Berdasarkan Komposisi Label," vol. 8, pp. 32–40, 2024.
- [11] D. Putriani, A. P. A. Prayogi, A. I. Shofyana, A. Ristyawan, and E. Daniati, "Prediksi Customer Churn Menggunakan Algoritma Decision Tree," *Inotek*, vol. 8, pp. 85–94, 2024.
- [12] I. Romli, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Klasifikasi Penyakit Ispa," *Indones. J. Bus. Intell.*, vol. 4, no. 1, p. 10, 2021, doi: 10.21927/ijubi.v4i1.1727.
- [13] F. Djiwadikusumah, G. H. Irawan, and R. Haekal Al-Fadilah, "Web scraping situs e-commerce menggunakan teknik parsing dom," *J. Siliwangi*, vol. 7, no. 2, pp. 52–57, 2021,[Online].Available:<https://jurnal.unsil.ac.id/index.php/jssainstek/article/view/4223/1958>
- [14] K. Sentimen *et al.*, "Klasifikasi Sentimen Pengguna Aplikasi Livin ByMandiri Pada Playstore Menggunakan AlgoritmaNaive Bayes," *Agustus*, vol. 8, pp. 2549–7952, 2024.
- [15] H. T. Santoso, F. A. Felmidi, A. Nur, A. Ristyawan, and E. Daniati, "Analisis Kinerja Algoritma Data Mining pada Klasifikasi Tingkat Obesitas dengan K-Fold Cross Validation dan AUC," vol. 8, pp. 113–122, 2024.