



## Analisis Klasifikasi Keputusan Belanja Konsumen Pada Toko Online XX Menggunakan Algoritma Decision Tree

Putri Maria Theresia Kehi<sup>1\*</sup>, I Wayan Sudiarsa<sup>2</sup>, Maria Oktaviani Suryati<sup>3</sup>, Yosefina Dehadi<sup>4</sup>, Maria Karlinda<sup>5</sup>

<sup>1-5</sup> Fakultas Bisnis dan Desain Kreatif, Institut Bisnis dan Teknologi Indonesia, Indonesia

\*Penulis Korespondensi: [sudiarsa@instiki.ac.id](mailto:sudiarsa@instiki.ac.id)

**Abstract.** *This study aims to analyze consumer purchasing behavior on e-commerce platforms using the Decision Tree algorithm as an easily interpretable classification method. The dataset used consists of 12,330 transaction records with 18 attributes representing visitor characteristics and user activities during interactions with the e-commerce platform. The research stages include data exploration to identify initial patterns, data preprocessing to handle missing values and class imbalance, splitting the data into training and testing sets, training the Decision Tree model, evaluating model performance, and visualizing the tree structure to analyze decision rules. The test results show that the Decision Tree model with a maximum depth of 3 achieves fairly good performance, with an average accuracy of 89.78%, precision of 69.82%, recall of 59.95%, and an F1-score of 64.51% for the buyer class. The visualization of the decision tree provides clear interpretation of the main attributes influencing purchasing decisions, thereby facilitating understanding for non-technical decision makers. Overall, this study demonstrates that the Decision Tree method is effective in modeling consumer purchasing behavior in e-commerce and can be utilized as a basis for data-driven business decision making, particularly in marketing strategies and improving sales conversion rates.*

**Keywords:** Classification; Consumer Behavior; Data Mining; Decision Tree; E-Commerce.

**Abstrak.** Studi ini bertujuan untuk menganalisis perilaku pembelian konsumen pada platform e-commerce menggunakan algoritma Pohon Keputusan (Decision Tree) sebagai metode klasifikasi yang mudah diinterpretasikan. Dataset yang digunakan berisi 12.330 data transaksi dengan 18 atribut yang mempresentasikan karakteristik pengunjung serta aktivitas pengguna selama berinteraksi dengan platform e-commerce. Tahapan penelitian meliputi eksplorasi data untuk memahami pola awal, pra-pemrosesan data guna menangani nilai hilang dan ketidakseimbangan kelas, pembagian data menjadi data dilatih dan data diuji, pelatihan model Pohon Keputusan, evaluasi kinerja model, serta visualisasi struktur pohon untuk analisis aturan keputusan. Hasil pengujian menunjukkan bahwa model Pohon Keputusan dengan kedalaman maksimum 3 menghasilkan kinerja yang cukup baik, dengan akurasi rata-rata 89,78%, presisi 69,82%, recall 59,95%, dan skor f1- sebesar 64,51% untuk kelas pembeli. Visualisasi pohon keputusan memberikan interpretasi yang jelas terhadap atribut-atribut utama yang memengaruhi keputusan pembelian, sehingga memudahkan pemahaman bagi pengambil keputusan non-teknis. Secara keseluruhan, studi ini menunjukkan bahwa metode pohon keputusan (Decision Tree) efektif dalam memodelkan perilaku pembelian konsumen pada e-commerce dan dapat di manfaatkan sebagai dasar pengambilan keputusan bisnis berbasis data, khususnya dalam strategi pemasaran dan peningkatan konversi penjualan.

**Kata Kunci:** E-Commerce; Klasifikasi; Penambangan Data; Perilaku Konsumen; Pohon Keputusan.

### 1. LATAR BELAKANG

Perkembangan teknologi informasi telah mendorong pertumbuhan pesat sektor e-commerce, sehingga menghasilkan volume data transaksi yang sangat besar dan kompleks. Data tersebut menyimpan informasi berharga mengenai perilaku konsumen yang dapat di manfaatkan untuk meningkatkan strategi pemasaran dan pengambilan keputusan bisnis. Menurut Han; Kamber; dan Pei (2012), data mining berperan penting dalam mengekstraksi pola dan pengetahuan dari kumpulan data besar untuk mendukung proses pengambilan keputusan secara efektif. Dalam konteks bisnis digital, pemanfaatan data transaksi menjadi

faktor kunci untuk memahami kebutuhan pelanggan dan meningkatkan daya saing perusahaan (Provost, Foster; Fawcett 2013).

Salah satu pendekatan utama dalam data mining adalah teknik klasifikasi, yaitu proses memetakan data dalam kategori tertentu berdasarkan karakteristik yang dimilikinya (Tan, P.-N.; Steinbach, M.; Kumar 2019). Teknik ini sangat relevan untuk menganalisis perilaku pembelian konsumen, karena perusahaan dapat memprediksi kemungkinan seorang pengunjung akan melakukan pembelian atau tidak. Algoritma Decision Tree merupakan metode klasifikasi yang banyak digunakan karena menghasilkan model yang mudah dipahami, dapat di visualisasikan dalam bentuk struktur pohon, dan mampu menjelaskan hubungan antara atribut secara eksplisit (Quinlan, 1986; Breiman et al., 1984). Keunggulan tersebut menjadikan Decision Tree cocok untuk diterapkan pada permasalahan analisis perilaku konsumen dalam e-commerce.

Namun demikian, permasalahan umum dalam analisis perilaku pembelian adalah ketidakseimbangan data, dimana jumlah konsumen yang tidak melakukan pembelian jauh lebih banyak di bandingkan konsumen yang melakukan pembelian. Kondisi ini dapat menurunkan kemampuan model dalam mengenali kelas minoritas secara akurat . Oleh karena itu, diperlukan penelitian yang tidak hanya membangun model klasifikasi, tetapi juga mengevaluasi kinerjanya secara menyeluruh serta memberikan interpretasi yang jelas terhadap hasil prediksi. Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menerapkan algoritma Decision Tree dalam menganalisis perilaku pembelian konsumen pada data e-commerce dan mengevaluasi kinerjanya sebagai dasar pengambilan keputusan bisnis berbasis data.

## **2. KAJIAN TEORITIS**

### **Data Mining**

Data mining merupakan proses penerapan metode statistik, pembelajaran mesin (*machine learning*), dan kecerdasan buatan untuk mengekstraksi informasi yang bernilai dari data mentah (Larose 2014). Larose juga menjelaskan bahwa proses data mining umumnya meliputi tahapan pemahaman masalah, pengumpulan data, pra-pemrosesan, pemodelan, evaluasi, dan implementasi.

Dalam konteks e-commerce, data mining memainkan peran penting dalam memahami perilaku pelanggan, memprediksi kecenderungan pembelian, mengoptimalkan strategi pemasaran, serta meningkatkan kepuasan pelanggan. Dengan memanfaatkan data transaksi dan aktivitas pengunjung, perusahaan dapat merancang strategi bisnis yang lebih akurat dan

berbasis bukti.

### **Klasifikasi Sebagai Teknik Data Mining**

Klasifikasi bertujuan untuk membangun model prediktif yang mampu memetakan sekumpulan atribut ke dalam kelas tertentu secara akurat (Witten et al. 2017). Model tersebut dipelajari dari data historis dan kemudian digunakan untuk mengklasifikasikan data baru. Pendekatan ini banyak digunakan dalam berbagai bidang, termasuk pemasaran, keuangan, dan sistem rekomendasi

Dalam penelitian ini, klasifikasi digunakan untuk memprediksi apakah seorang pengunjung akan melakukan pembelian atau tidak. Variabel target yang digunakan adalah Revenu, dengan dua kelas yaitu *Pembeli* (x) dan *tidak membeli* (y). Model klasifikasi belajar dari data historis untuk menemukan pola hubungan antara atribut pengunjung dengan keputusan pembelian.

### **Algoritma Decision Tree**

Decision Tree adalah model prediktif yang mempresentasikan hubungan antara atribut input dan kelas target dalam bentuk struktur pohon, dimana setiap simpul internal mempresentasikan pengujian terhadap suatu atribut, setiap cabang menunjukkan hasil dari pengujian tersebut, dan setiap simpul daun menyatakan keputusan atau kelas akhir (Quinlan, 1986).

Keunggulan utama algoritma Decision Tree terletak pada kemampuannya untuk :

- a. Menghasilkan model yang mudah dipahami dan diinterpretasikan.
- b. Menangani data numerik maupun kategorikal.
- c. Memvisualisasikan aturan keputusan secara eksplisit

Proses pembentukan pohon keputusan dilakukan dengan memilih atribut terbaik pada setiap simpul berdasarkan kriteria tertentu, seperti Information Gain atau Gini Index, yang bertujuan untuk memaksimalkan pemisahan antar kelas.

### **Struktur Pohon Keputusan**

Struktur Decision Tree terdiri dari beberapa komponen utama yang saling berkaitan dalam merepresentasikan proses pengambilan keputusan. Menurut Quinlan (1993), pohon keputusan dibangun dalam bentuk struktur hierarkis yang terdiri dari simpul akar (root), simpul internal, cabang, dan simpul daun, yang masing-masing memiliki peran penting dalam proses klasifikasi.

- a. Root Node

Root node merupakan simpul awal dari pohon keputusan yang mewakili atribut paling berpengaruh dalam menentukan hasil klasifikasi. Menurut Han, Kamber,

dan Pei (2012), atribut yang dipilih sebagai root node adalah atribut dengan kemampuan terbaik dalam memisahkan data ke dalam kelas-kelas yang berbeda berdasarkan kriteria tertentu, seperti *Information Gain* atau *Gini Index*.

b. Internal Node

Internal node adalah simpul yang digunakan untuk melakukan pengujian terhadap suatu atribut tertentu. Setiap internal node merepresentasikan kondisi atau aturan yang harus dipenuhi sebelum data dialirkan ke cabang berikutnya. Larose (2014), menjelaskan bahwa internal node berfungsi sebagai titik pengambilan keputusan sementara dalam proses klasifikasi.

c. Branch

Branch atau cabang menunjukkan hasil dari pengujian terhadap atribut pada simpul sebelumnya. Setiap cabang merepresentasikan nilai atau rentang nilai dari atribut yang diuji. Menurut Witten, Frank, dan Hall (2017), cabang berfungsi sebagai penghubung antar simpul yang mengarahkan alur klasifikasi dari root node hingga ke leaf node.

d. Leaf Node

Leaf node merupakan simpul akhir yang menyatakan keputusan atau kelas hasil klasifikasi. Breiman et al. (1984), menyebutkan bahwa leaf node berisi label kelas yang diprediksi atau nilai probabilitas kelas tertentu, yang menjadi output akhir dari model Decision Tree.

Berdasarkan pendapat para ahli tersebut, dapat disimpulkan bahwa struktur pohon keputusan dirancang untuk merepresentasikan proses pengambilan keputusan secara sistematis, transparan, dan mudah dipahami, sehingga sangat efektif digunakan dalam analisis dan klasifikasi data.

### Proses Pembentukan Decision Tree

Proses pembentukan Decision Tree dilakukan secara rekursif melalui langkah-langkah berikut:

- a. Menentukan atribut terbaik sebagai pemisah data pada setiap simpul.
- b. Membagi data ke dalam cabang-cabang berdasarkan nilai atribut tersebut.
- c. Mengulangi proses hingga tercapai kondisi berhenti, seperti:
  - Semua data dalam satu simpul berasal dari kelas yang sama.
  - Kedalam maksimum pohon telah tercapai
  - Jumlah data pada simpul terlalu kecil untuk dibagi

## Decision Tree dalam Analisis Perilaku Konsumen

Dalam analisis perilaku konsumen, Decision Tree sangat efektif karena mampu menampilkan hubungan antara atribut aktivitas pengguna (Jumlah halaman yang dikunjungi, durasi kunjungan, tingkat keluar halaman, dan nilai tambah) dengan keputusan pembelian secara eksplisit. Setiap jalur dari root hingga leaf node membentuk suatu aturan keputusan yang menjelaskan pola perilaku konsumen

### Penelitian Terdahulu

Beberapa penelitian terdahulu menunjukkan bahwa algoritma Decision Tree efektif digunakan untuk menganalisis niat belanja konsumen online. Sakar et al, membuktikan bahwa atribut perilaku kunjungan pengguna berpengaruh signifikan terhadap keputusan pembelian, meskipun penelitian tersebut lebih menekankan perbandingan performa algoritma dan kurang mendalami interpretasi atribut dominan. *Liu dan Shih (2016)* menegaskan keunggulan Decision Tree dalam menghasilkan aturan keputusan yang mudah dipahami, namun penelitian ini terbatas pada jumlah data dan metrik evaluasi. Sementara itu, Nugraheni et al, dalam konteks e-commerce Indonesia menemukan bahwa kualitas produk, kemudahan penggunaan aplikasi, dan kepercayaan konsumen memengaruhi keputusan pembelian, meskipun penggunaan data kuesioner menjadi keterbatasan karena bersifat subjektif.

## METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk membangun dan mengevaluasi model klasifikasi perilaku pembelian konsumen menggunakan algoritma Decision Tree. Tahapan penelitian pada gambar 1, dirancang secara sistematis mulai dari pengumpulan data hingga evaluasi model untuk memastikan keakuratan dan reliabilitas hasil penelitian. Tahapan penelitian dilakukan melalui beberapa langkah utama yaitu, eksplorasi data, pra-pemrosesan data, pembagian dataset, pelatihan model, evaluasi model, interpretasi dan visualisasi model.



**Gambar 1.** Metode penelitian.

### Jenis dan Pendekatan Penelitian

Jenis penelitian yang digunakan adalah penelitian kuantitatif dengan pendekatan eksperimental, di mana data dianalisis menggunakan teknik data mining dan pembelajaran mesin untuk menentukan pola perilaku konsumen. Pendekatan ini memungkinkan peneliti melakukan pengujian objektif terhadap performa algoritma klasifikasi dalam memprediksi keputusan pembelian.

### Sumber Data dan Deskripsi Dataset

Data yang digunakan dalam penelitian ini merupakan dataset transaksi pengunjung e-commerce yang terdiri dari 12.330 data dan 18 atribut, dari repositori [www.kaggle.com](http://www.kaggle.com) dengan user Md. Abdullah Al Mamun. Dataset ini memuat informasi mengenai aktivitas pengguna, karakteristik pengunjung, serta keputusan pembelian yang direpresentasikan oleh variabel *Revenue* sebagai variabel target.

| No | Column                  | Non-Null        | Count Dytpe |
|----|-------------------------|-----------------|-------------|
| 1  | Administrative          | 12330 non-nul   | int64       |
| 2  | Administrative_Duration | 12330 non-null  | int64       |
| 3  | Informational           | 12330 non- null | int64       |
| 4  | Informational_Duration  | 12330 non-null  | float64     |
| 5  | ProductRelated          | 12330 non-null  | int64       |
| 6  | ProductRelated_Duration | 12330 non-null  | float64     |
| 7  | BounceRates             | 12330 non-null  | float64     |
| 8  | ExitRates               | 12330 non-null  | float64     |
| 9  | PageValues              | 12330 non-null  | float64     |
| 10 | SpecialDay              | 12330 non-null  | float64     |
| 11 | Month                   | 12330 non-null  | object      |
| 12 | OperatingSystems        | 12330 non-null  | int64       |
| 13 | Browser                 | 12330 non-null  | int64       |
| 14 | Region                  | 12330 non-null  | int64       |
| 15 | TrafficType             | 12330 non-null  | int64       |
| 16 | VisitorType             | 12330 non-null  | object      |
| 17 | Weekend                 | 12330 non-null  | bool        |
| 18 | Revenue                 | 12330 non-null  | bool        |

dtypes: bool(2), float64(7), int64(7), object(2)

### Perancangan Model Decision Tree

Model klasifikasi yang digunakan adalah Decision Tree Classifier dengan konfigurasi parameter:

- *Max\_Depth* = 3
- *Random\_State* = 43

Pemilihan parameter tersebut bertujuan untuk menjaga keseimbangan antara kompleksitas model dan kemampuan generalisasi, sekaligus mempertahankan kemudahan interpretasi.

### Metode Evaluasi

Evaluasi kinerja model dilakukan menggunakan beberapa metrik klasifikasi, yaitu:

- Akurasi, untuk mengukur tingkat ketepatan prediksi secara keseluruhan.
- Precision, untuk mengukur ketepatan prediksi kelas pembeli
- Recall, untuk mengukur kemampuan model dalam mendeteksi pembeli sebenarnya.
- F1-Score, sebagai harmonisasi antara precision dan recall

Metrik-metrik ini dipilih karena memberikan gambaran komprehensif terhadap performa model dalam menghadapi ketidakseimbangan kelas.

### Confusion Matrix:

|                   | Predicted No Revenue | Predicted Revenue |
|-------------------|----------------------|-------------------|
| Actual No Revenue | 1899                 | 156               |
| Actual Revenue    | 173                  | 238               |

**Gambar 2.** Metode Evaluasi.

## 3. HASIL DAN PEMBAHASAN

### Deskripsi Dataset dan Eksplorasi Data

Dataset yang digunakan dalam penelitian ini terdiri dari 12.330 data transaksi dengan 18 atribut yang merepresentasikan karakteristik pengunjung e-commerce dan aktivitas pengguna. Variabel target yang digunakan adalah *Revenue*, yang menunjukkan apakah seorang pengunjung melakukan pembelian (*True*) atau tidak (*False*).

Berdasarkan eksplorasi awal, tidak ditemukan nilai yang hilang pada seluruh atribut, sehingga dataset dinyatakan lengkap dan layak digunakan untuk pemodelan

### Pra-Pemrosesan Data

Tahapan pra-pemrosesan dilakukan untuk memastikan kualitas data dan kesiapan dataset sebelum digunakan dalam pemodelan.

### Transformasi dan Normalisasi Atribut

Beberapa atribut numerik bertipe object, seperti *Administrative\_Duration*, *Informational\_Duration*, *ProductRelated\_Duration*, dan *PageValues*. Atribut tersebut dikonversi menjadi tipe numerik (float) agar dapat diproses oleh algoritma pembelajaran mesin.

Selain itu, atribut kategorikal seperti *Month* dan *Visitor Type* diubah menjadi bentuk numerik menggunakan teknik one-hot encoding, sedangkan atribut boolean seperti *Weekend* dan *Revenue* dikonversi menjadi nilai biner (0 dan 1).

### Pembagian Data Latih dan Data Uji

Dataset dibagi menjadi 80% data latih (9.864 data) dan 20% data uji (2.466 data)

menggunakan teknik *stratified sampling* untuk menjaga proporsi kelas.

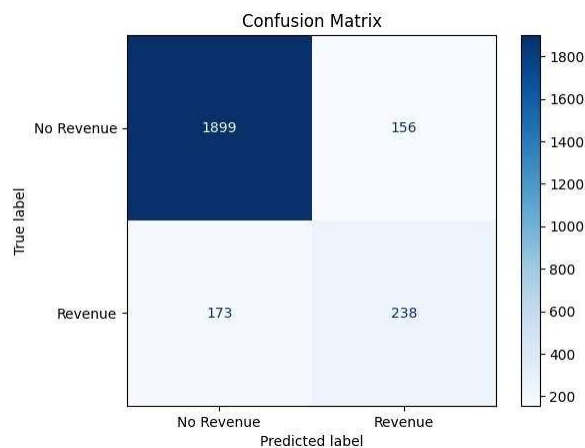
### Pelatihan Model Decision Tree

Model Decision Tree dilatih menggunakan data latih dengan paramter *max\_depth* = 3. Pemilihan kedalaman maksimum ini bertujuan menghindari *overfitting* sekaligus mempertahankan interpretabilitas model.

Dengan konfigurasi tersebut, model mampu membentuk struktur keputusan yang sederhana namun tetap efektif dalam memisahkan kelas pembeli dan non-pembeli.

### Evaluasi Performa Model

Evaluasi dilakukan menggunakan data uji dengan matrik akurasi, precision, recall, dan F1-Score.



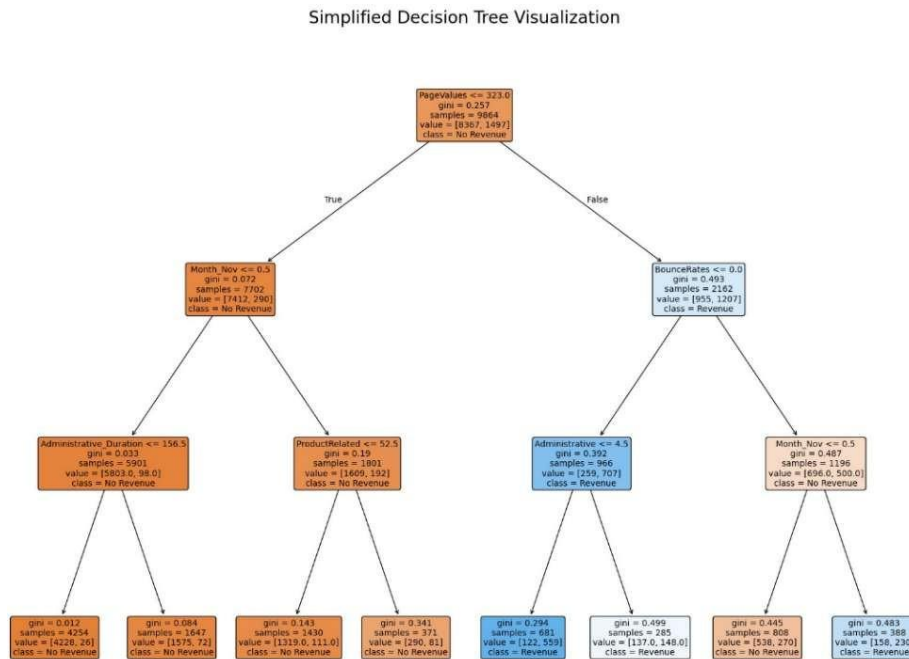
**Gambar 3.** Evaluasi Performa Model.

| Metrik    | Nilai  |
|-----------|--------|
| Akurasi   | 89,78% |
| Precision | 69,82% |
| Recall    | 59,95% |
| F1-Score  | 64,51% |

Hasil tersebut menunjukkan bahwa model memiliki performa yang sangat baik dalam mengklasifikasi nonpembeli (kelas 0), namun performanya masih moderat dalam mendeteksi pembeli (kelas 1). Hal ini disebabkan oleh ketidakseimbangan distribusi kelas pada dataset.

### Interpretasi Model dan Visualisasi Pohon Keputusan

Untuk memahami pola pengambilan keputusan model, dilakukan visualisasi struktur pohon keputusan.



**Gambar 4.** Visualisasi Pohon Keputusan

## Pembahasan

Berdasarkan hasil eksperimen, algoritma Decision Tree terbukti mampu memodelkan perilaku pembelian konsumen secara efektif. Akurasi sebesar 89,79% menunjukkan bahwa model berhasil mempelajari pola umum data, sementara visualisasi pohon keputusan memberikan keunggulan tambahan berupa interpretabilitas yang sangat baik.

Namun demikian, rendahnya nilai recall pada kelas pembeli menunjukkan bahwa masih terdapat sebagian pelanggan potensial yang gagal terdeteksi oleh model. Oleh karena itu, diperlukan pengembangan lanjutan seperti penanganan ketidakseimbangan data dan eksplorasi algoritma klasifikasi lainnya guna meningkatkan kinerja prediksi.

## 4. KESIMPULAN DAN SARAN

### Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma Decision Tree mampu memodelkan perilaku pembelian konsumen dengan tingkat performa yang baik serta menghasilkan model yang mudah dipahami dan diinterpretasikan. Model yang dibangun dengan parameter kedalaman maksimum (*Max\_Depth*) sebesar 3 memperoleh akurasi sebesar 89,79% yang menunjukkan kemampuan klasifikasi yang tinggi dalam membedakan konsumen yang melakukan pembelian dan tidak melakukan pembelian. Selain itu, visualisasi pohon keputusan memberikan gambar yang jelas mengenai faktor-faktor yang berpengaruh terhadap keputusan pembelian, sehingga model ini dapat dimanfaatkan sebagai

alat bantu pengambilan keputusan dalam strategi bisnis berbasis data.

### **Saran**

Meskipun model yang dihasilkan telah menunjukkan kinerja yang baik, penelitian ini masih memiliki beberapa keterbatasan, khususnya terkait dengan ketidakseimbangan distribusi kelas pada dataset yang memengaruhi kemampuan model dalam mendeteksi konsumen yang melakukan pembelian. Oleh karena itu, penelitian selanjutnya disarankan untuk menerapkan teknik penanganan ketidakseimbangan data, seperti oversampling menggunakan metode *Smote*, guna meningkatkan nilai recall pada kelas pembeli. Selain itu, disarankan pula untuk membandingkan kinerja algoritma Decision Tree dengan metode klasifikasi lain, seperti Random Forest, dan Gradient Boosting, agar diperoleh model yang memiliki tingkat akurasi dan generalisasi yang lebih optimal

### **UCAPAN TERIMA KASIH**

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam pelaksanaan penelitian ini. Secara khusus, penulis menyampaikan apresiasi kepada dosen pengampu mata kuliah Artificial Intelligence atas bimbingan dan arahan yang diberikan selama proses penelitian. Ucapan terima kasih juga disampaikan kepada pihak penyedia data serta seluruh rekan yang telah membantu dalam proses pengolahan data, analisis, dan penyusunan artikel ini sehingga dapat diselesaikan dengan baik.

### **DAFTAR PUSTAKA**

- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Belmont: Wadsworth International Group.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321-357. <https://doi.org/10.1613/jair.953>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). San Francisco: Morgan Kaufmann.
- Kelleher, J. D., & Tierney, B. (2018). *Data science: An introduction*. CRC Press. <https://doi.org/10.7551/mitpress/11140.001.0001>
- Larose, D. T. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining*. Hoboken: John Wiley & Sons. <https://doi.org/10.1002/9781118874059>
- Larose, D. T. (2015). *Data Mining and Predictive Analytics*. Hoboken: John Wiley & Sons.
- Liu, D. R., & Shih, Y. Y. (2016). *Integrating AHP and data mining for product r*

ecommodation based on customer lifetime value. *Information & Management*, 53(4), 404-415.

- Nugraheni, R., Santoso, I., & Prabowo, A. (2022). Analisis Faktor yang Mempengaruhi Keputusan Pembelian Konsumen pada E-Commerce di Indonesia. *Jurnal Sistem Informasi*, 18(2), 145-156.
- Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. Sebastopol: O'Reilly Media.
- Quinlan, J. R. (1986). Induction of Decision Trees. *Machine Learning*, 1(1), 81- 106. <https://doi.org/10.1023/A:1022643204877>
- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. San Mateo: Morgan Kaufmann.
- Sakar, C. O., Polat, S. O., Katircioglu, M., & Kastro, Y. (2019). Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks. *Neural Computing and Applications*, 31(10), 6893-6908. <https://doi.org/10.1007/s00521-018-3523-0>
- Tan, P. N., Steinbach, M., & Kumar, V. (2019). *Introduction to Data Mining (2nd ed.)*. Boston: Pearson.
- Witten, I. H., Frank, E., & Hall, M. A. (2016). *Data Mining: Practical Machine Learning Tools and Techniques (4th ed.)*. Cambridge: Morgan Kaufmann.