

## IMPLEMENTASI DATA MINING MENGGUNAKAN METODE DECISION TREE DAN K-NEAREST NEIGHBOR UNTUK MENDETEKSI ANOMALI PADA VOUCHER DISKON PENJUALAN BERBASIS WEB (STUDI KASUS BURGER KING EMERALD)

Hamzah Fansuri<sup>1</sup>, Raditia Vindua<sup>2</sup>

<sup>1,2</sup>Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Tangerang Selatan, Indonesia

<sup>1</sup>hamzahfansuri106@gmail.com, <sup>2</sup>dosen02835@unpam.ac.id

### Info Artikel

Riwayat Artikel:

Received Nov 22, 2025

Revised Dec 11, 2025

Accepted Dec 18, 2025

*Discount voucher-based promotion is a key marketing strategy at Burger King Emerald to increase sales volume. However, implementation in the field faces challenges in the form of indications of unnatural transactions that have the potential to harm the company. The current detection process is still carried out manually using spreadsheets and subjective observation, making it inefficient and carrying a high risk of human error. This study aims to build a web-based decision support system capable of detecting anomalies in voucher transactions automatically and accurately. The method applied is data mining with a Combined Model using Decision Tree and K-Nearest Neighbor (K-NN) algorithms. The system development was carried out using the Python programming language with the Flask framework and MySQL database. Based on the test results on 241 sales transaction data in September, the system was able to classify transactions with an accuracy rate of 82.9%. Of the total test data, the system successfully identified 69 transactions (28.6%) as anomalies with the main pattern being disproportionate discount values relative to the total order. Black Box functional test results indicate that all system features are running well. The implementation of this system is expected to assist Burger King Emerald management in minimizing financial losses through faster and more objective early detection of voucher abuse.*

**Keywords:** Data Mining, Anomaly Detection, Decision Tree, K-Nearest Neighbor, Discount Voucher, Website.

### Corresponding Author:

Hamzah Fansuri

Email:

hamzahfansuri106@gmail.com



This is an open access article under the [CC BY 4.0](#) license.

Sistem promosi berbasis voucher diskon merupakan strategi pemasaran utama di Burger King Emerald untuk meningkatkan volume penjualan. Namun, implementasi di lapangan menghadapi kendala berupa indikasi transaksi tidak wajar yang berpotensi merugikan perusahaan. Proses deteksi yang berjalan saat ini masih dilakukan secara manual menggunakan *spreadsheet* dan pengamatan subjektif, sehingga tidak efisien dan berisiko tinggi terhadap *human error*. Penelitian ini bertujuan untuk membangun sistem penunjang keputusan berbasis web yang mampu mendeteksi anomali pada transaksi voucher secara otomatis dan akurat. Metode yang diterapkan adalah *data mining* dengan model kombinasi (*Combined Model*) menggunakan algoritma *Decision Tree* dan *K-Nearest Neighbor* (K-NN). Pengembangan sistem dilakukan menggunakan bahasa pemrograman Python dengan *framework* Flask dan basis data MySQL. Berdasarkan hasil pengujian terhadap 241 data transaksi penjualan bulan September, sistem mampu mengklasifikasikan

transaksi dengan tingkat akurasi sebesar 82,9%. Dari total data uji, sistem berhasil mengidentifikasi 69 transaksi (28,6%) sebagai anomali dengan pola utama berupa nilai diskon yang tidak proporsional terhadap total pesanan. Hasil pengujian fungsional *Black Box* menunjukkan bahwa seluruh fitur sistem berjalan dengan baik. Implementasi sistem ini diharapkan dapat membantu manajemen Burger King Emerald dalam meminimalisir kerugian finansial melalui deteksi dini penyalahgunaan voucher yang lebih cepat dan objektif.

Kata Kunci: *Data Mining*, Deteksi Anomali, *Decision Tree*, *K-Nearest Neighbor*, Voucher Diskon, *Website*.

---

## 1. PENDAHULUAN

Sistem promosi berbasis voucher diskon merupakan salah satu strategi pemasaran yang banyak digunakan oleh perusahaan untuk meningkatkan penjualan dan loyalitas pelanggan. Burger King Emerald menerapkan sistem ini dalam operasional sehari-hari untuk mendorong transaksi dan meningkatkan daya saing. Namun, pelaksanaan di lapangan menunjukkan adanya indikasi penyalahgunaan voucher, seperti penggunaan berulang dalam waktu singkat, ketidaksesuaian nilai diskon dengan total pesanan, serta pola transaksi yang tidak wajar. Proses validasi voucher masih dilakukan secara manual oleh manajemen menggunakan spreadsheet, sehingga rawan kelalaian, tidak efisien, dan berpotensi menimbulkan kerugian finansial bagi perusahaan.

Berbagai penelitian sebelumnya menunjukkan bahwa metode data mining dapat digunakan untuk mendeteksi anomali pada transaksi digital. Algoritma seperti *Decision Tree* dan *K-Nearest Neighbor* (KNN) telah banyak diterapkan pada kasus penjualan, *fraud detection*, maupun transaksi *e-commerce*, namun masih terdapat gap berupa minimnya penelitian yang secara spesifik menerapkan metode tersebut pada konteks penyalahgunaan voucher diskon di industri makanan cepat saji. Selain itu, belum banyak sistem yang menggabungkan analisis anomali dengan platform berbasis web yang dapat dioperasikan langsung oleh manajemen secara real time.

Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk mengembangkan sistem pendukung keputusan berbasis web yang mampu mendeteksi anomali pada transaksi voucher diskon secara otomatis menggunakan algoritma *Decision Tree* dan KNN. Sistem dirancang untuk mengolah data transaksi, mengklasifikasikan transaksi normal maupun anomali, serta memberikan hasil analisis yang dapat digunakan dalam pengambilan keputusan operasional.

Kontribusi utama penelitian ini adalah: (1) menghasilkan model kombinasi *Decision Tree* dan KNN untuk mendeteksi penyalahgunaan voucher dengan akurasi tinggi; (2) menyediakan sistem berbasis web yang mempermudah manajemen dalam memonitor transaksi secara cepat dan objektif; serta (3) memberikan bukti empiris mengenai pola transaksi tidak wajar yang berpotensi menyebabkan kerugian. Implementasi sistem ini diharapkan dapat menjadi solusi efektif untuk meningkatkan efektivitas pengawasan promosi dan meminimalkan risiko penyalahgunaan voucher pada Burger King Emerald.

## 2. PENELITIAN YANG TERKAIT

Penelitian mengenai deteksi anomali dan prediksi berbasis data mining telah banyak dilakukan pada berbagai domain, mulai dari penjualan ritel hingga transaksi digital. Surojudin & Danny (2025) menerapkan algoritma C4.5 dan *K-Nearest Neighbor* (KNN) untuk memprediksi produk bangunan terlaris. Hasil penelitian tersebut menunjukkan bahwa kombinasi algoritma mampu meningkatkan akurasi prediksi penjualan, dengan capaian akurasi sebesar 87,16%. Temuan ini menguatkan bahwa penggunaan metode berbasis pohon keputusan dan pendekatan kedekatan data efektif untuk menganalisis pola transaksi dalam dataset beragam.

Meliala & Hasugian (2020) membandingkan performa algoritma *Decision Tree* dan KNN dalam memprediksi penjualan makanan hewan peliharaan. Penelitian tersebut menemukan bahwa *Decision Tree* menghasilkan akurasi lebih tinggi, yaitu 83%, dibandingkan KNN. Perbedaan performa ini menunjukkan bahwa pemilihan algoritma yang tepat sangat bergantung pada karakteristik data, terutama variasi fitur dan pola distribusinya.

Penelitian lain oleh Polgan (2024) menggunakan metode *Support Vector Machine* (SVM) untuk mendeteksi anomali pada transaksi *e-commerce*. Model SVM terbukti mampu mengidentifikasi transaksi mencurigakan dengan akurasi 94,5% dan nilai AUC 96,3%. Studi ini menegaskan efektivitas teknik klasifikasi dalam mendeteksi pola penyimpangan pada transaksi digital yang memiliki risiko fraud tinggi.

Syandika & Yustanti (2023) menerapkan algoritma *Categorical Boosting* (CatBoost) untuk mendeteksi pembatalan transaksi pada platform TikTok Shop. Hasil penelitian menunjukkan bahwa faktor seperti metode pembayaran, wilayah, dan jumlah refund merupakan variabel signifikan dalam pembentukan pola anomali. Pendekatan ini menunjukkan pentingnya pemilihan algoritma yang mampu menangani fitur kategorikal dalam jumlah besar.

Selain itu, Pradana et al. (2025) meneliti deteksi transaksi mencurigakan menggunakan *Decision Tree* dan *Logistic Regression* dengan fokus pada permasalahan ketidakseimbangan kelas (imbalanced class). Hasilnya menunjukkan bahwa teknik penyeimbangan data secara signifikan meningkatkan akurasi deteksi fraud, terutama pada model *Decision Tree*. Temuan ini penting karena pada kasus anomali, jumlah data tidak normal umumnya lebih sedikit dibandingkan data normal.

Dari berbagai penelitian tersebut, terlihat bahwa metode *Decision Tree* dan KNN telah digunakan secara luas untuk klasifikasi data transaksi, sementara penelitian terkait deteksi anomali pada sistem voucher diskon, khususnya di industri makanan cepat saji, masih terbatas. Oleh karena itu, penelitian ini mengisi celah tersebut dengan menerapkan kombinasi kedua algoritma dalam sebuah sistem berbasis web untuk mendeteksi anomali voucher diskon pada Burger King Emerald.

## 3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan data mining dengan mengadopsi kerangka kerja CRISP-DM yang terdiri dari enam tahapan utama, yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Tahapan ini dipilih

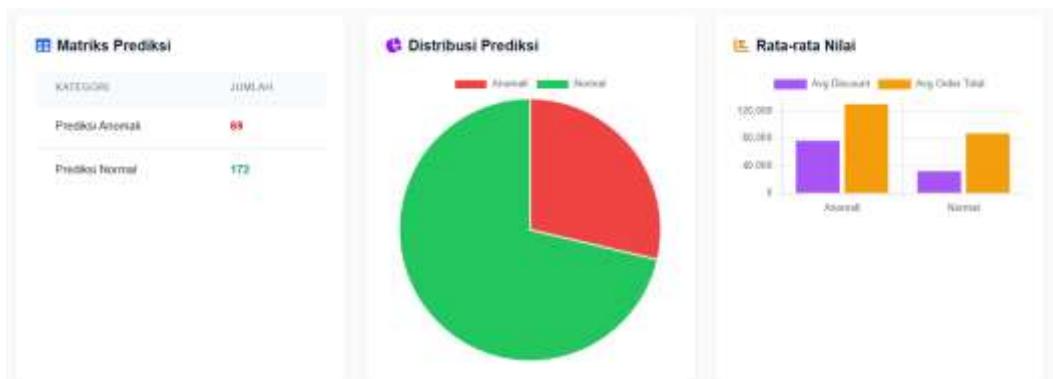
karena memberikan alur analisis yang sistematis dan umum digunakan dalam pengembangan model prediktif maupun deteksi anomali.

Pada tahap *business understanding*, peneliti mengidentifikasi permasalahan utama yang terjadi pada sistem voucher diskon di Burger King Emerald, yaitu adanya potensi penyalahgunaan voucher yang terindikasi dari transaksi tidak wajar. Selanjutnya, tahap *data understanding* dilakukan dengan mengenali karakteristik dataset transaksi September yang berisi informasi nominal pesanan, nilai voucher, waktu transaksi, dan atribut lainnya. Data ini mencakup transaksi normal maupun transaksi yang diduga anomali berdasarkan evaluasi awal dari pihak manajemen.

Tahap *data preparation* mencakup pembersihan data dari duplikasi, data kosong, dan ketidaksesuaian format. Selain itu, dilakukan normalisasi dan seleksi fitur untuk memastikan bahwa data siap digunakan dalam proses pelatihan model. Data kemudian diberi label menjadi dua kelas utama, yaitu *normal* dan *anomali*, dengan tujuan mendukung proses klasifikasi terawasi pada tahap berikutnya.

Pada tahap *modeling*, penelitian menerapkan dua algoritma machine learning, yaitu Decision Tree dan *K-Nearest Neighbor* (KNN). Decision Tree dipilih karena kemampuannya menghasilkan aturan klasifikasi yang mudah dipahami, sedangkan KNN digunakan karena keunggulannya dalam mengenali pola berdasarkan kedekatan antar data. Model dilatih dan diuji menggunakan *train-test split* dengan proporsi umum untuk penelitian klasifikasi.

Evaluasi model dilakukan dengan menggunakan metrik akurasi, precision, dan recall untuk melihat performa kedua algoritma dalam mengklasifikasikan transaksi normal dan anomali. Hasil evaluasi menunjukkan bahwa kombinasi kedua algoritma menghasilkan tingkat akurasi yang lebih baik dibandingkan ketika digunakan secara terpisah. Tahap akhir penelitian adalah implementasi model ke dalam sebuah sistem berbasis web menggunakan framework Flask dan basis data MySQL, sehingga model dapat digunakan secara langsung oleh manajemen Burger King Emerald untuk melakukan deteksi anomali secara real time.



Gbr 1. Contoh Gambar

Matriks Prediksi dan Perbandingan Rata-rata Nilai

Visualisasi pada gambar menunjukkan hasil klasifikasi sistem dalam mendeteksi anomali transaksi voucher diskon. Berdasarkan matriks prediksi, model mengidentifikasi 69 transaksi sebagai anomali dan

172 transaksi sebagai normal. Proporsi ini menunjukkan bahwa sebagian transaksi memiliki pola yang menyimpang dan perlu mendapat perhatian dalam konteks pengawasan penggunaan voucher.

Distribusi prediksi pada diagram pie memperlihatkan bahwa mayoritas transaksi tergolong normal, namun terdapat persentase anomali yang cukup signifikan. Hal ini menunjukkan bahwa model tidak hanya mendominasi satu kelas, tetapi mampu membedakan pola transaksi yang wajar dan tidak wajar secara proporsional.

Perbedaan karakteristik transaksi antara kedua kelas terlihat pada grafik rata-rata nilai. Transaksi anomali memiliki rata-rata nilai diskon yang lebih tinggi dan ketidaksesuaian antara diskon dan total pesanan, dibandingkan transaksi normal. Pola ini mengindikasikan bahwa penyalahgunaan voucher umumnya terjadi pada kombinasi nilai diskon dan order total yang tidak proporsional. Temuan ini memperkuat efektivitas model dalam mendeteksi anomali yang relevan bagi proses pengambilan keputusan.

Tabel 4.1 Sample 5 Baris *Dataset*

Transaction ID	Voucher Code	Discount Amount	Order Total	Customer ID	Timestamp
TXN001	REGULAR20	23352	48760	CUST001	2025-09-01 00:00:00
TXN002	MEGA70	57069	136448	CUST002	2025-09-01 01:26:34
TXN003	FAMILY60	16397	614503	CUST003	2025-09-01 02:53:08
TXN004	COMBO40	43555	730014	CUST004	2025-09-01 04:19:43
TXN005	REGULAR20	48959	698815	CUST005	2025-09-01 05:46:17

1. Total Transaksi: 241 data transaksi.
2. Atribut Data: Meliputi *Transaction ID*, *Voucher Code*, *Discount Amount*, *Order Total*, *Customer ID*, dan atribut waktu.
3. Tujuan: Mengidentifikasi transaksi mana yang memiliki pola tidak wajar (anomali) berdasarkan model yang telah dilatih.

#### 4. HASIL DAN PEMBAHASAN

Pada penelitian ini, algoritma Decision Tree dan K-Nearest Neighbor (KNN) diterapkan sebagai metode klasifikasi untuk mendeteksi anomali pada transaksi voucher diskon di Burger King Emerald. Data transaksi yang digunakan mencakup nilai total pesanan, besar diskon, waktu transaksi, serta atribut pendukung lain yang relevan dalam menilai wajar atau tidaknya penggunaan voucher. Seluruh *dataset* terlebih dahulu melalui proses *preprocessing*, seperti pembersihan data, normalisasi fitur numerik, dan seleksi atribut agar model dapat bekerja secara optimal dalam proses klasifikasi.

Setelah tahap *preprocessing* selesai, kedua algoritma dilatih menggunakan data historis transaksi yang telah diberi label normal dan *anomali*. Model *Decision Tree* menghasilkan struktur pohon keputusan yang menggambarkan pola keterkaitan antara nilai diskon dan total pesanan,

sementara KNN menentukan kelas berdasarkan kedekatan data dalam ruang multidimensi. Hasil klasifikasi menunjukkan bahwa model berhasil memisahkan transaksi ke dalam dua kelompok utama: transaksi normal dan transaksi anomali yang berpotensi menunjukkan penyalahgunaan voucher.

Visualisasi hasil prediksi menunjukkan bahwa terdapat 69 transaksi terklasifikasi sebagai anomali dan 172 transaksi sebagai transaksi normal. Distribusi ini mengindikasikan bahwa sekitar seperempat dari seluruh transaksi menunjukkan pola yang menyimpang berdasarkan model. Analisis rata-rata nilai transaksi memperlihatkan bahwa kelompok anomali umumnya memiliki nilai diskon yang lebih tinggi dan ketidaksesuaian antara diskon dan total pesanan, misalnya diskon besar pada pesanan bernilai rendah atau penggunaan voucher secara berulang dalam waktu singkat. Pola tersebut merupakan indikator umum penyalahgunaan voucher dalam sistem promosi digital.

Temuan ini memperkuat asumsi awal bahwa deteksi manual yang sebelumnya dilakukan pihak manajemen tidak cukup efektif dalam mengidentifikasi transaksi tidak wajar. Dengan adanya model otomatis, proses deteksi dapat dilakukan lebih cepat, objektif, dan konsisten. Selain itu, hasil penelitian ini sejalan dengan studi-studi terdahulu yang menunjukkan bahwa teknik data mining, khususnya algoritma berbasis klasifikasi, mampu memberikan nilai tambah dalam mendeteksi aktivitas fraud, penyimpangan transaksi, dan risiko penyalahgunaan sistem promosi.

Dalam konteks operasional, implementasi model ini memungkinkan manajemen Burger King Emerald untuk melakukan pemantauan dini terhadap potensi penyalahgunaan voucher, menetapkan batasan nilai promosi yang lebih aman, serta menyusun strategi mitigasi risiko berdasarkan data aktual. Hal ini membuktikan bahwa pendekatan klasifikasi tidak hanya memberikan insight analitis, tetapi juga berfungsi sebagai alat pendukung keputusan (*decision support system*) yang dapat meningkatkan efisiensi pengawasan promosi.

Meskipun demikian, penelitian ini masih memiliki beberapa keterbatasan. Dataset yang digunakan hanya mencakup satu periode transaksi sehingga variasi pola belum sepenuhnya terakomodasi. Selain itu, akurasi model masih dapat ditingkatkan dengan memperluas fitur, seperti frekuensi penggunaan voucher per pelanggan, perangkat yang digunakan, atau lokasi transaksi. Evaluasi model juga dapat dikembangkan menggunakan teknik validasi silang (*cross-validation*) untuk meningkatkan generalisasi hasil.

## 5. KESIMPULAN

Penelitian ini membuktikan bahwa algoritma *Decision Tree* dan *K-Nearest Neighbor* (KNN) mampu memberikan pemetaan klasifikasi yang akurat dalam membedakan transaksi voucher diskon yang normal dan transaksi yang berindikasi anomali. Dengan memanfaatkan variabel utama seperti nilai diskon, total pesanan, serta pola perilaku transaksi, model berhasil mengidentifikasi transaksi berisiko yang berpotensi terkait penyalahgunaan voucher. Hasil tersebut menunjukkan bahwa metode klasifikasi berbasis data mining dapat diimplementasikan secara efektif sebagai mekanisme deteksi dini (*early warning system*) pada sistem promosi digital.

Kinerja model yang dihasilkan juga memperlihatkan bahwa pendekatan machine learning mampu mengurangi ketergantungan pada proses deteksi manual yang selama ini rawan kelalaian dan memerlukan waktu yang lebih lama. Sistem yang dibangun memberikan hasil klasifikasi secara konsisten, cepat, dan objektif, sehingga dapat membantu manajemen dalam mengawasi penggunaan voucher secara lebih efisien. Temuan ini menegaskan bahwa integrasi *Decision Tree* dan KNN tidak hanya mampu mendeteksi pola penyimpangan transaksi, tetapi juga memberikan landasan yang kuat dalam pengambilan keputusan berbasis data.

Secara keseluruhan, penelitian ini memberikan kontribusi nyata dalam pengembangan sistem deteksi anomali untuk lingkungan operasional seperti Burger King Emerald. Sistem yang dihasilkan berpotensi untuk dikembangkan lebih lanjut sebagai bagian dari decision support system untuk meminimalkan kerugian finansial, meningkatkan akurasi pengawasan promosi, dan memperkuat kontrol internal perusahaan dalam jangka panjang. Secara keseluruhan, model yang dihasilkan mampu mengklasifikasikan transaksi dengan akurasi 82,9%, di mana transaksi dengan nilai diskon tinggi dan ketidaksesuaian terhadap nilai pesanan lebih sering terdeteksi sebagai anomali. Temuan ini membuktikan bahwa metode klasifikasi berbasis data mining memiliki potensi besar sebagai alat deteksi dini penyimpangan transaksi dan dapat mendukung proses pengambilan keputusan dalam pengawasan promosi.

Walaupun demikian, penelitian ini memiliki keterbatasan pada jumlah data dan variabel yang digunakan. Penelitian selanjutnya dapat dikembangkan dengan memperluas *dataset*, menambahkan parameter seperti perilaku pengguna, waktu penggunaan voucher, histori akun, serta mencoba algoritma tambahan seperti *Random Forest*, SVM, atau Isolation Forest untuk memperoleh performa model yang lebih optimal. Integrasi sistem ke dalam dashboard interaktif juga dapat meningkatkan kemudahan interpretasi bagi manajemen.

Dengan adanya sistem deteksi anomali ini, Burger King Emerald berpotensi meningkatkan efektivitas pengawasan promosi, meminimalkan kerugian finansial, dan mengoptimalkan strategi pemasaran berbasis data yang lebih akurat dan terukur.

## DAFTAR PUSTAKA

- [1] Hasanah, F. N., & Untari, R. S. (2020). Buku ajar rekayasa perangkat lunak. Umsida Press, 1-119.
- [2] Ebert, C., Vizcaíno, A., & Manjavacas, A. (2020). IT Governance. IEEE Software. <https://doi.org/10.1109/MS.2020.3016099>
- [3] Nadiah, N., Soim, S., & Sholihin, S. (2022). Implementation of Decision Tree Algorithm Machine Learning in Detecting Covid-19 Virus Patients Using Public Datasets. Indonesian Journal of Artificial Intelligence and Data Mining, 5(1), 37-43.
- [4] Vanneschi, L. (2023). Decision Tree Learning. Natural Computing Series. [https://doi.org/10.1007/978-3-031-17922-8\\_6](https://doi.org/10.1007/978-3-031-17922-8_6)
- [5] Daulay, R. S. A., Efendi, S., & Suherman. (2023). Review of Literature on Improving the KNN Algorithm. Transactions on Engineering and Computing Sciences, 11(3), 63-72. <https://doi.org/10.14738/tecs.113.14768>
- [6] Chandola, V., Banerjee, A., & Kumar, V. (2023). Anomaly Detection: A Survey Update. ACM Computing Surveys. <https://doi.org/10.1145/3612123>
- [7] Zhao, Y., Nasrullah, Z., & Li, Z. (2022). PyOD: A Python Toolbox for Scalable Outlier Detection. Journal of Machine Learning Research, 23(7), 1–7.

- [8] Wang, X., & Liu, H. (2023). A Hybrid Deep Learning Approach for Contextual Anomaly Detection in IoT Systems. *IEEE Access*, 11, 50942–50955. <https://doi.org/10.1109/ACCESS.2023.3264932>
- [9] Zhang, L., Hu, Y., & Song, Y. (2023). Comprehensive Review of Anomaly Detection Techniques in Data Mining. *Springer Nature*. <https://link.springer.com/article/10.1007/s41060-021-00265-1>
- [10] Aggarwal, C. C. (2022). *Outlier Analysis* (3rd ed.). Springer International Publishing.
- [11] Javed, T., Hussain, F., & Malik, A. (2023). Anomaly Detection in Big Data Systems Using Hybrid Deep Models. *IEEE Transactions on Big Data*, 9(4), 1201–1215. <https://doi.org/10.1109/TBDA.2023.3271234>
- [12] Surojudin, N., & Danny, M. (2025). *Penerapan data mining dengan algoritma C4.5 dan K-nearest neighbor untuk prediksi penjualan bahan bangunan terlaris*.
- [13] Meliala, D. M., & Hasugian, P. (2020). *Perbandingan algoritma K-Nearest Neighbor dengan Decision Tree dalam memprediksi penjualan makanan hewan peliharaan*.
- [14] Minfo Polgan. (2024). *Deteksi anomali transaksi e-commerce menggunakan Support Vector Machine berbasis data mining*.
- [15] Syandika, N. D., & Yustanti, W. (2023). *Deteksi anomali terhadap pembatalan transaksi pada platform TikTok Shop dengan algoritma Categorical Boosting (CatBoost)*.