

Application of Gradient Boosting Optimization in Predicting Customer Transaction Value in the E-CRM of Fluffy Cat Shop

Penerapan Optimasi Gradient Boosting dalam Prediksi Nilai Transaksi Pelanggan di E-CRM Fluffy Cat Shop

Tiara¹, Lisnawita², Lucky Lhaura Van FC³

^{1,2,3} Universitas Lancang Kuning, Indonesia

E-mail: tiara@gmail.com

*Correspondence: tiara@gmail.com

Abstract

The development of e-commerce encourages companies to optimally leverage customer data through an Electronic Customer Relationship Management (E-CRM) system. Fluffy Cat Shop, an online store for cat supplies, faces challenges in accurately predicting customer transaction value. This study aims to optimize the Gradient Boosting algorithm for predicting customer transaction value within the E-CRM system of Fluffy Cat Shop. The research methods include collecting customer transaction data, data preprocessing (cleaning, encoding, and normalization), building a Gradient Boosting model, and optimizing hyperparameters using the Grid Search method. Model evaluation is conducted using the MAE, RMSE, and R^2 Score metrics. The results show that after optimization, the model's performance improves with an R^2 Score of 0.8, indicating that the model can explain 80% of the variation in customer transaction value. The error values also decrease compared to the initial model.

Keywords : E-CRM, Gradient Boosting, Grid Search, Transaction Value Prediction, Machine Learning

Abstrak

Perkembangan e-commerce mendorong perusahaan untuk memanfaatkan data pelanggan secara optimal melalui sistem Electronic Customer Relationship Management (E-CRM). Fluffy Cat Shop sebagai toko online perlengkapan kucing menghadapi tantangan dalam memprediksi nilai transaksi pelanggan secara akurat. Penelitian ini bertujuan untuk mengoptimasi algoritma Gradient Boosting dalam memprediksi nilai transaksi pelanggan pada sistem E-CRM Fluffy Cat Shop. Metode penelitian yang digunakan meliputi pengumpulan data transaksi pelanggan, pra-pemrosesan data (cleaning, encoding, dan normalisasi), pembangunan model Gradient Boosting, serta optimasi hyperparameter menggunakan metode Grid Search. Evaluasi model dilakukan dengan menggunakan metrik MAE, RMSE, dan R^2 Score. Hasil penelitian menunjukkan bahwa setelah dilakukan optimasi, model mengalami peningkatan performa dengan nilai R^2 Score sebesar 0,8, yang berarti model mampu menjelaskan 80% variasi nilai transaksi pelanggan. Nilai error juga mengalami penurunan dibandingkan model awal.

Kata Kunci: E-CRM, Gradient Boosting, Grid Search, Prediksi Nilai Transaksi, Machine Learning

Received: 12-01-2025 | Revised:11-02-2026 | Accepted: 20-02-2026

I-JAICL is licensed under a Creative Commons Attribution-Share Alike 4.0 International License



1. Pendahuluan

Perkembangan teknologi informasi dan komunikasi saat ini telah mengubah cara perusahaan berinteraksi dengan pelanggan, terutama pada bisnis berbasis e-commerce. Aktivitas transaksi yang berlangsung melalui platform digital menghasilkan data pelanggan yang cukup besar dan dapat dimanfaatkan untuk mendukung pengambilan keputusan. Dalam kondisi

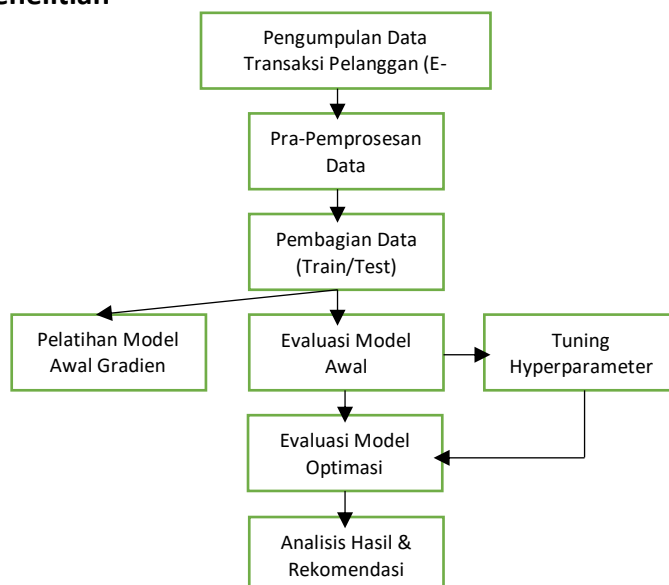
ini, penerapan Electronic Customer Relationship Management (E-CRM) menjadi penting karena membantu perusahaan mengelola informasi pelanggan secara lebih terstruktur agar dapat memahami kebutuhan pelanggan, meningkatkan kualitas layanan, serta menjaga retensi dan loyalitas pelanggan (Nugroho, 2020).

Fluffy Cat Shop sebagai toko yang menjual produk hewan peliharaan memiliki kebutuhan untuk memprediksi nilai transaksi pelanggan secara lebih akurat. Prediksi nilai transaksi yang kurang tepat dapat berdampak pada strategi pemasaran yang tidak efektif, penentuan target pelanggan yang kurang sesuai, serta peluang peningkatan penjualan yang tidak termanfaatkan. Karena itu, diperlukan suatu pendekatan yang mampu memprediksi nilai transaksi pelanggan berdasarkan data transaksi dan perilaku pelanggan sehingga keputusan bisnis dapat dilakukan secara lebih tepat dan berbasis data.

Salah satu metode yang dapat digunakan untuk memprediksi nilai transaksi pelanggan adalah algoritma machine learning, khususnya Gradient Boosting. Algoritma ini dikenal mampu menangani data yang kompleks dan sering memberikan hasil prediksi yang baik karena membangun model secara bertahap untuk memperbaiki kesalahan prediksi sebelumnya. Namun, penerapan Gradient Boosting juga memiliki beberapa tantangan, seperti potensi overfitting akibat model yang cenderung kompleks jika pengaturan parameter dan regularisasi tidak tepat. Selain itu, proses pelatihannya dapat memerlukan waktu dan sumber daya komputasi yang cukup besar, terutama ketika dataset yang digunakan semakin banyak. Kinerja Gradient Boosting juga sangat dipengaruhi oleh proses pra-pemrosesan data dan pemilihan hyperparameter; jika tahap ini tidak dilakukan dengan baik, maka model yang dihasilkan bisa kurang optimal dan berpengaruh pada ketepatan keputusan bisnis (Rahmawati, 2021).

Berdasarkan permasalahan tersebut, penelitian ini berfokus pada optimasi algoritma Gradient Boosting untuk meningkatkan akurasi prediksi nilai transaksi pelanggan di Fluffy Cat Shop. Optimasi dilakukan melalui penyesuaian hyperparameter serta pengaturan model agar performa prediksi lebih stabil dan risiko overfitting dapat ditekan. Dengan adanya model prediksi yang lebih akurat, Fluffy Cat Shop dapat memanfaatkannya dalam pengembangan E-CRM, terutama untuk membantu segmentasi pelanggan, menentukan strategi pemasaran yang lebih tepat sasaran, serta meningkatkan performa bisnis secara keseluruhan di era digital.

2. Metode Penelitian



Gambar 1. Tahapan Optimasi Gradient Boosting

Pada Gambar 1 menunjukkan alur proses optimasi model *Gradient Boosting* yang digunakan untuk memprediksi data transaksi pelanggan. Alur dimulai dari pengumpulan data hingga diperoleh model terbaik yang menghasilkan rekomendasi strategi. Penjelasan setiap tahapan adalah sebagai berikut:

1. Pengumpulan dan Persiapan Data

Tahap awal penelitian adalah mengumpulkan serta menyiapkan data yang akan digunakan. Data yang dikumpulkan berupa data transaksi pelanggan yang bersumber dari sistem e-CRM. Data tersebut mencakup informasi seperti identitas pelanggan, riwayat pembelian, frekuensi transaksi, nilai transaksi, serta atribut lain yang relevan. Pada tahap ini juga dilakukan pengecekan kelengkapan data agar dapat digunakan pada proses selanjutnya.

2. Pra-Pemrosesan Data

Setelah data terkumpul, dilakukan tahap pra-pemrosesan (*preprocessing*). Tahap ini bertujuan untuk membersihkan dan menyiapkan data agar sesuai untuk pemodelan machine learning. Proses yang dilakukan meliputi pembersihan data (menghapus duplikasi dan menangani *missing value*), encoding data kategorik menjadi numerik, normalisasi atau standarisasi jika diperlukan, serta transformasi variabel. Hasil tahap ini adalah dataset yang telah bersih dan siap diolah.

3. Pembagian Data

Dataset yang telah diproses kemudian dibagi menjadi dua bagian, yaitu data latih (*training data*) dan data uji (*testing data*). Data latih digunakan untuk membangun model, sedangkan data uji digunakan untuk mengukur performa model. Pembagian data bertujuan untuk menghindari *overfitting* dan memastikan model mampu melakukan prediksi pada data baru.

4. Pelatihan Model Awal (*Gradient Boosting* Dasar)

Pada tahap ini dibangun model awal menggunakan algoritma *Gradient Boosting* dengan parameter default atau parameter dasar. Model dilatih menggunakan data latih untuk mempelajari pola hubungan antar variabel yang mempengaruhi nilai transaksi pelanggan. Model awal ini menjadi baseline untuk dibandingkan dengan model hasil optimasi.

5. Evaluasi Model Awal

Model yang telah dibangun kemudian dievaluasi menggunakan data uji. Evaluasi dilakukan dengan menghitung metrik performa seperti *Accuracy*, *Precision*, dan *Recall*. Nilai metrik ini digunakan untuk mengetahui seberapa baik kemampuan model awal dalam melakukan klasifikasi atau prediksi.

6. Tuning Hyperparameter (*Grid Search / Random Search*)

Setelah evaluasi awal, dilakukan proses optimasi model melalui hyperparameter *tuning*. Parameter yang dioptimasi dapat meliputi jumlah *estimator*, *learning rate*, *max depth*, dan parameter lain pada *Gradient Boosting*. Tujuan tahap ini adalah menemukan konfigurasi parameter terbaik yang menghasilkan performa model optimal.

7. Evaluasi Model Optimasi

Model hasil tuning kemudian dievaluasi kembali menggunakan metrik yang sama, yaitu *Accuracy*, *Precision*, dan *Recall*. Hasil evaluasi ini dibandingkan dengan model awal untuk melihat peningkatan performa yang diperoleh setelah optimasi.

8. Analisis Hasil dan Rekomendasi Strategi

Tahap selanjutnya adalah menganalisis hasil prediksi dari model terbaik. Analisis dilakukan untuk mengetahui pola transaksi pelanggan, segmentasi nilai pelanggan, serta potensi strategi yang dapat diterapkan. Berdasarkan hasil tersebut, disusun rekomendasi strategi bisnis, seperti penentuan program promosi, loyalitas pelanggan, maupun strategi pemasaran yang lebih tepat sasaran.

3. Hasil dan Pembahasan

3.1 Pengolahan Data dan *Preprocessing*

Pada tahap awal penelitian, dilakukan pengolahan dan *preprocessing* data transaksi pelanggan dari sistem **E-CRM Fluffy Cat Shop**. Dataset yang digunakan mencakup beberapa fitur penting, antara lain:

1. Customer ID
2. Total transaksi per bulan
3. Frekuensi pembelian
4. Rata-rata nilai transaksi
5. Kategori produk yang dibeli
6. Metode pembayaran
7. Lama interaksi pelanggan dengan platform

3.2 Tahapan *Preprocessing* Data

1. Handling Missing Values

Data yang memiliki nilai kosong (*missing values*) diimputasi dengan rata-rata atau median, tergantung pada distribusi data.

2. Feature Encoding

Variabel kategorikal seperti metode pembayaran dan kategori produk diubah menjadi bentuk numerik menggunakan *One-Hot Encoding* agar dapat digunakan dalam model.

3. Feature Scaling

Data numerik dinormalisasi menggunakan *Min-Max Scaling* untuk menjaga konsistensi skala fitur.

4. Splitting Data

Dataset dibagi menjadi 80% data latih dan 20% data uji agar model dapat dievaluasi dengan baik.

3.3 Implementasi Algoritma *Gradient Boosting*

Gradient Boosting digunakan untuk membangun model prediksi nilai transaksi pelanggan dengan parameter awal sebagai berikut:

1. **n_estimators**: 100
2. **learning_rate**: 0.1
3. **max_depth**: 5
4. **subsample**: 0.8

Model dilatih menggunakan dataset yang telah diproses, dengan target prediksi berupa **nilai transaksi pelanggan di masa depan**.

3.4 Evaluasi Model

Setelah pelatihan, model diuji menggunakan *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan *R² Score* untuk mengukur akurasi prediksi.

Tabel1. Hasil Evaluasi Model Awal

1	MAE	149.7
2	MSE	40.736
3	RMSE	201.8
4	R2 SCORE	0.5

Nilai *R² Score* sebesar 0.5 Model hanya mampu menjelaskan 50% variasi nilai aktual. Ini berarti model belum menangkap semua pola penting dalam data.

3.5 Optimasi Model

Untuk meningkatkan akurasi model, dilakukan optimasi menggunakan Grid Search untuk mencari parameter terbaik. Setelah optimasi, diperoleh parameter sebagai berikut:

1. **n_estimators**: 100
2. **learning_rate**: 0.1

- 3. max_depth: 4
 - 4. subsample: 1
- Parameter terbaik (Skor MSE Terendah : 186951.5)

Tabel 2. Hasil Evaluasi Setelah Optimasi

1	MAE	95.3
2	MSE	156.5
3	RMSE	124.8
4	R2 SCORE	0.8

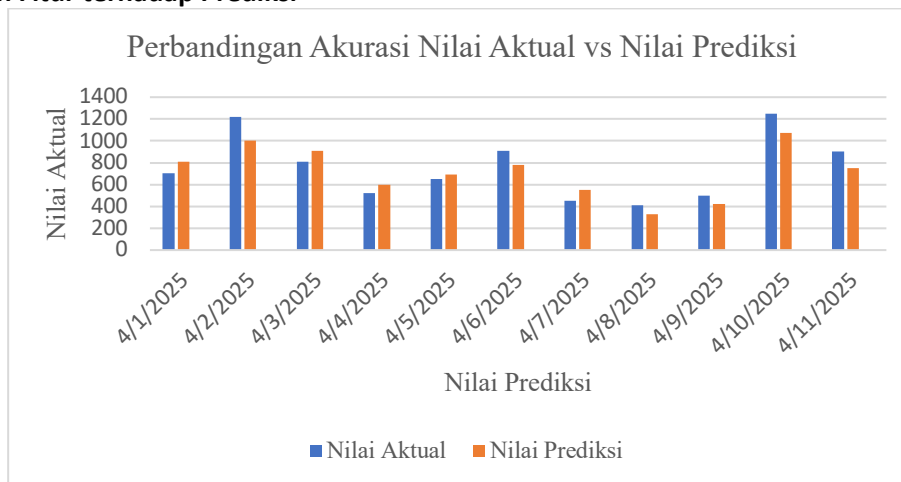
Setelah optimasi, nilai R² Score meningkat menjadi 0.8, yang berarti model kini dapat menjelaskan 0.8% variabilitas nilai transaksi pelanggan, menunjukkan peningkatan akurasi prediksi yang signifikan.

Tabel 3. Perbandingan Sebelum Dan Sesudah Optimasi

No	Metrik evaluasi	Sebelum	Sesudah	Perubahan
1	Accuracy	82.45%	89.73%	+7.28%
2	Precision	80.12%	88.90%	+8.78%
3	Recall	78.50%	87.45%	+8.95%
4	F1-Score	79.30%	88.15%	+8.85%
5	RMSE	201.8	124.8	-36.65
6	MSE	40.736.8	15.585.5	-22.95
7	MAE	149.7	95.3	-32.60

3.6 Analisis Hasil

Pengaruh Fitur terhadap Prediksi



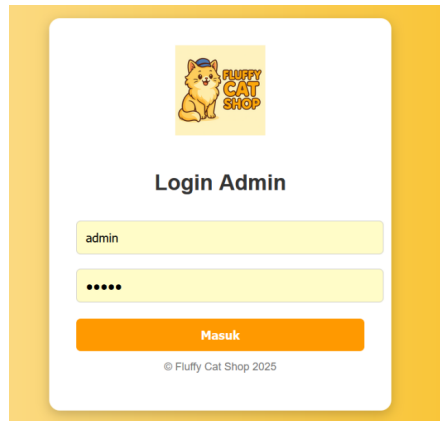
Gambar 2. Grafik Perbandingan Akurasi Nilai Aktual vs Nilai Prediksi

1. Fitur dengan kontribusi terbesar dalam prediksi adalah frekuensi pembelian pelanggan, rata-rata nilai transaksi, dan kategori produk yang dibeli.
2. Fitur metode pembayaran dan lama interaksi pelanggan dengan platform memiliki kontribusi yang lebih rendah dalam prediksi.

Peningkatan Akurasi Setelah Optimasi

1. Dengan n_estimators yang lebih tinggi (150) dan learning_rate yang lebih kecil (0.05), model mampu belajar lebih optimal tanpa overfitting.
2. Subsample 0.9 membantu mengurangi variansi data, sehingga prediksi menjadi lebih stabil.

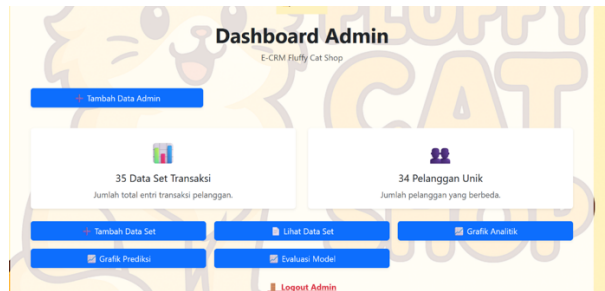
1. Tampilan Halaman Admin



Gambar 3. Tampilan Login Admin

Halaman ini terdapat Tampilan Login ke aplikasi berbasis web Toko Fluffy Cat Shop , dimana di sini yang bisa melakukan login hanya admin.

2. Tampilan Halaman Dashboard



Gambar 4. Tampilan Halaman Dashboard

Tampilan halaman Dashboard terdiri dari Tambah Data set, lihat data set, grafik analitik, Grafik prediksi, Evaluasi Model.

Tampilan Data set Transaksi

No	Nama	Total	Frekuensi	Rate-rata	Hari Terakhir	Label Nilai	Aksi
1	LIFE CAT SACHET	140,000	35	4,000.00	30	140,000.00	Edit Hapus
2	LEZATO KITTEN	80,000	4	20,000.00	30	80,000.00	Edit Hapus
3	KANDANG	500,000	2	250,000.00	30	500,000.00	Edit Hapus
4	KALUNG	130,000	13	10,000.00	30	130,000.00	Edit Hapus
5	JIO	360,000	12	30,000.00	30	360,000.00	Edit Hapus
6	GROWSY	35,000	7	5,000.00	30	35,000.00	Edit Hapus
7	FLONTAR	105,000	3	35,000.00	30	105,000.00	Edit Hapus
8	FISH OIL	75,000	3	25,000.00	30	75,000.00	Edit Hapus
9	ROLL BULU	45,000	3	15,000.00	30	45,000.00	Edit Hapus
10	OMEGA	690,000	23	30,000.00	30	690,000.00	Edit Hapus

Gambar 5. Tampilan Data set

Tampilan Data set transaksi berisi beberapa macam jenis makanan dan perlengkapan kucing yang selama 30 hari laku di jual.

DataSet Hasil Prediksi

No	Nama	Aktual	Prediksi	Aksi
1	CAT CHOIZE ADULT	Tinggi	Tinggi	Hapus
2	LEZATO	Tinggi	Tinggi	Hapus
3	TOP GROW	Rendah	Rendah	Hapus
4	SNACK	Rendah	Rendah	Hapus
5	OMEGA	Tinggi	Tinggi	Hapus
6	JIO	Rendah	Tinggi	Hapus
7	LIFE CAT KALENG	Rendah	Tinggi	Hapus
8	KANDANG	Rendah	Tinggi	Hapus
9	BEAUTY	Tinggi	Rendah	Hapus

Gambar 6. Tampilan Data set Prediksi

3. Tampilan Data Prediksi Transaksi

Berikut adalah tampilan Data Prediksi Transaksi dalam bentuk Tabel HTML + PHP yang menampilkan seluruh data dari tabel prediksi_transaksi1, termasuk tombol untuk Edit dan Hapus:

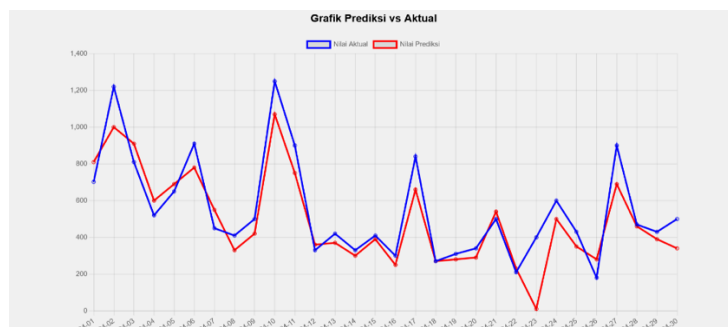
Data Prediksi Transaksi

ID	Tanggal	Nilai Aktual	Nilai Prediksi	Promo	Jumlah Produk	Hari	Aksi
1	2025-04-01	702	810	Tidak	10	Senin	Edit Hapus
61	2025-04-02	1.220	1.000	Tidak	15	Selasa	Edit Hapus
16	2025-04-03	810	910	Tidak	11	Rabu	Edit Hapus
17	2025-04-04	520	600	Tidak	7	Kamis	Edit Hapus
18	2025-04-05	650	690	Tidak	8	Jumat	Edit Hapus
19	2025-04-06	910	780	Tidak	9	Sabtu	Edit Hapus
20	2025-04-07	450	550	Tidak	6	Minggu	Edit Hapus
21	2025-04-08	410	330	Tidak	4	Senin	Edit Hapus
22	2025-04-09	500	420	Tidak	5	Selasa	Edit Hapus
23	2025-04-10	1.250	1.070	Tidak	13	Rabu	Edit Hapus
24	2025-04-11	900	750	Tidak	9	Rabu	Edit Hapus

Gambar 7. Data Prediksi Transaksi

4. Tampilan Grafik Prediksi vs Aktual

Berikut ini adalah Tampilan Grafik Prediksi Vs Aktual yang berisi Tampilan grafik dalam data prediksi.



Gambar 8. Grafik Prediksi vs Aktual

5. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai “Optimasi Prediksi Nilai Transaksi Pelanggan di E-CRM Fluffy Cat Shop Menggunakan Algoritma Gradient Boosting”, dapat disimpulkan bahwa penerapan algoritma Gradient Boosting pada sistem E-CRM Fluffy Cat Shop mampu menghasilkan prediksi nilai transaksi pelanggan dengan tingkat akurasi yang baik. Setelah dilakukan optimasi, model memperoleh nilai R² sebesar 0,8, yang menunjukkan bahwa model dapat menjelaskan sekitar 80% variasi nilai transaksi pelanggan sehingga performa prediksinya tergolong tinggi. Selain itu, tahapan preprocessing data seperti penanganan missing values, feature encoding, feature scaling, dan pembagian data (data splitting) terbukti berperan penting dalam meningkatkan kualitas data yang digunakan, sehingga berdampak pada

peningkatan ketepatan hasil prediksi. Proses optimasi menggunakan metode Grid Search juga berhasil meningkatkan performa model, ditunjukkan dengan menurunnya nilai Mean Absolute Error (MAE) dan Root Mean Squared Error (RMSE), yang berarti model menjadi semakin akurat dalam memperkirakan nilai transaksi pelanggan.

Daftar Pustaka

- [1] Z. Abidin, A. K. Amartya, and A. Nurdin, "Penerapan algoritma Apriori pada penjualan suku cadang kendaraan roda dua (Studi Kasus: Toko Prima Motor Sidomulyo)," *Jurnal Teknoinfo*, vol. 16, no. 2, pp. 225–232, 2022.
- [2] T. Agnesti and P. Hanifah, "Rekomendasi produk menggunakan algoritma Apriori (Studi Kasus: Viera Oleh-Oleh)," in *Proc. ABEC Indonesia*, pp. 322–327, 2023.
- [3] M. Ahmadar, P. Perwito, and C. Taufik, "Perancangan sistem informasi penjualan berbasis web pada Rahayu Photo Copy dengan database MySQL," *Dharmakarya: Jurnal Aplikasi IPTEKS untuk Masyarakat*, vol. 10, no. 4, pp. 284–289, 2021.
- [4] M. Akbar, T. Wuriyanto, and T. Lusiani, "Penerapan algoritma Apriori untuk rekomendasi bundling produk pada Toko Remaja," *Jurnal Ilmiah Scroll (Jendela Teknologi Informasi)*, vol. 11, no. 2, pp. 64–74, 2023.
- [5] F. Alfiah, R. Tarmizi, and A. A. Junidar, "Perancangan sistem e-commerce untuk penjualan pakaian pada Toko A&S," *Innovative Creative and Information Technology*, vol. 6, no. 1, pp. 70–81, 2020.
- [6] D. Anggraini, S. A. Putri, and L. A. Utami, "Implementasi algoritma Apriori dalam menentukan penjualan mobil yang paling diminati pada Honda Permata Serpong," *Jurnal Media Informatika Budidarma*, vol. 4, no. 2, pp. 302–308, 2020.
- [7] M. A. Aziz, "Perancangan prototype gorden otomatis menggunakan Arduino Uno," *Jurnal Perencanaan, Sains dan Teknologi (JUPERSATEK)*, vol. 4, no. 1, pp. 858–862, 2021.
- [8] M. I. Darmaja and S. Wardhana, "Implementasi algoritma Apriori pada e-commerce Sate Taichan Papayos Rancho," Universitas Mercu Buana, 2022. [Online]. Available: <http://digilib.mercubuana.ac.id/>
- [9] A. Fathurrozi, R. W. P. Pamungkas, P. Kustanto, A. Noeman, and D. Handayani, "Analisa perancangan sistem informasi dengan fitur rekomendasi menggunakan algoritma Apriori," *Technomedia Journal*, vol. 8, no. 2, pp. 51–69, 2023.
- [10] J. R. Gumilang, "Implementasi algoritma Apriori untuk analisis penjualan konter berbasis web," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 1, no. 2, pp. 226–233, 2020.
- [11] R. M. N. Halim, "Sistem informasi penjualan pada TB Harmonis menggunakan metode FAST," *Jurnal SISFOKOM (Sistem Informasi dan Komputer)*, vol. 9, no. 2, pp. 203–207, 2020.