

Analisis Faktor Keberhasilan Penjualan Kerajinan Tangan menggunakan Decision Tree dengan Optimasi Grid Search

Nailus Sa'idah Anindia Septiana*, Muhammad Jauhar Vikri, Ita Aristia Sa'ida

Fakultas Sains dan Teknologi, Teknik Informatika, Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro, Indonesia

Email: ^{1,*}nadiaseptiana001@gmail.com, ²vikri@unugiri.ac.id, ³itaaristia@unugiri.ac.id

Email Penulis Korespondensi: nadiaseptiana001@gmail.com

Submitted 22-01-2025; Accepted 09-03-2026; Published 30-04-2026

Abstrak

Penelitian ini dilatarbelakangi oleh keterbatasan kemampuan pelaku Usaha Mikro, Kecil, dan Menengah (UMKM) kerajinan dalam menganalisis faktor-faktor penentu keberhasilan penjualan di platform e-commerce, meskipun data transaksi historis tersedia. Penelitian sebelumnya umumnya menerapkan algoritma klasifikasi tanpa optimasi hyperparameter yang sistematis, sehingga berpotensi menghasilkan model yang kurang optimal dan rentan terhadap overfitting. Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan penerapan algoritma Decision Tree yang dioptimasi menggunakan metode Grid search Cross-validation. Dataset yang digunakan berasal dari platform e-commerce Brasil (Olist Dataset) pada kategori 'artes' sebagai representasi produk kerajinan tangan, dengan pembagian data 80% untuk pelatihan dan 20% untuk pengujian. Proses optimasi mengeksplorasi 576 kombinasi parameter untuk memperoleh konfigurasi terbaik. Hasil pengujian menunjukkan model teroptimasi mencapai akurasi sebesar 97.61% dengan struktur pohon sederhana (max_depth=None), sehingga mudah diinterpretasikan. Analisis feature importance mengidentifikasi tinggi produk sebagai faktor paling dominan (64.23%), diikuti oleh lebar_produk, nilai_pengiriman, berat_produk, dan harga. Temuan ini menunjukkan bahwa kombinasi Decision Tree dan Grid search efektif dalam menghasilkan model prediktif yang akurat dan interpretable sebagai dasar pengambilan keputusan strategis bagi UMKM kerajinan.

Kata Kunci: Decision Tree; Grid search Cross-validation; Penjualan e-commerce; UMKM Kerajinan; Feature Importance

Abstract

This study is motivated by the limited ability of handicraft Micro, Small, and Medium Enterprises (MSMEs) to analyze the key factors influencing sales success on e-commerce platforms, despite the availability of historical transaction data. Previous studies generally applied classification algorithms without systematic hyperparameter optimization, potentially leading to suboptimal models and overfitting issues. To address this gap, this research proposes the implementation of a Decision Tree algorithm optimized using Grid search Cross-validation. The dataset was obtained from the Brazilian e-commerce platform (Olist Dataset), specifically the 'artes' category as a proxy for handicraft products, with an 80:20 split for training and testing data. The optimization process explored 576 parameter combinations to determine the best configuration. The optimized model achieved an accuracy of 97.61% with a simplified tree structure (max_depth=None), enhancing interpretability. Feature importance analysis product_height_cm as the most dominant factor (64.23%), followed by product_height_cm, product_width_cm, Freight_value, product_weight_g, and price. These findings demonstrate that the combination of Decision Tree and Grid search effectively produces an accurate and interpretable predictive model, providing strategic decision-making support for handicraft MSMEs in digital marketplaces.

Keywords: Decision Tree; Grid Search Cross-validation; E-commerce Sales Prediction; Handicraft MSMEs; Feature Importance

1. PENDAHULUAN

Transformasi digital telah mengubah sektor kerajinan tangan Indonesia, di mana adopsi *e-commerce* membuka akses pasar global bagi usaha mikro, kecil, dan menengah (UMKM), namun sekaligus meningkatkan persaingan yang semakin ketat [1]. Karakteristik unik produk kerajinan, seperti nilai estetika, keunikan desain, dan muatan budaya lokal, menyebabkan tantangan spesifik dalam pemasaran online [2]. Penelitian terhadap produk kerajinan kayu menunjukkan bahwa dibutuhkan strategi segmentasi yang tepat berdasarkan cara pelanggan berperilaku agar bisa bersaing [2]. Masalah utama yang dihadapi pelaku UMKM kerajinan adalah keterbatasan kemampuan analisis data untuk mengidentifikasi faktor-faktor determinan keberhasilan transaksi di platform digital, meskipun memiliki akses terhadap data transaksi historis [3]. Pemanfaatan teknik *data mining* dalam analisis penjualan menjadi salah satu pendekatan yang banyak digunakan untuk membantu pengambilan keputusan berbasis data. Selain itu, penelitian sebelumnya umumnya masih menggunakan pendekatan *machine learning* tanpa optimasi *hyperparameter* secara sistematis sehingga berpotensi menghasilkan model yang kurang optimal dan rentan terhadap *overfitting*. Tantangan ini diperparah oleh ketidaktersediaan *dataset* spesifik kerajinan yang terstruktur, sementara algoritma *machine learning* seperti *Decision Tree* yang potensial untuk analisis prediktif sering menghadapi kendala optimasi parameter dan risiko *overfitting* [4], [5].

Penelitian terkait telah mengeksplorasi aspek *e-commerce* dan kerajinan dari berbagai perspektif. Studi kualitatif oleh [1] mengungkap strategi transformasi digital UMKM kerajinan tenun di Bali, meskipun tidak mengintegrasikan pendekatan kuantitatif berbasis data. Dalam bidang *machine learning* untuk UMKM [6], mengimplementasikan *Decision Tree* untuk prediksi penjualan produk fashion dengan akurasi 88,24%, namun tanpa optimasi *hyperparameter*. Mencapai akurasi 81,82% dalam prediksi produk terlaris menggunakan *Decision Tree*, tetapi juga belum mengintegrasikan teknik tuning parameter secara sistematis [7]. Kajian komprehensif oleh [8] menemukan bahwa algoritma *ensemble* seperti *Random Forest* memberikan performa terbaik (akurasi 83,27%) untuk prediksi keberhasilan produk *e-commerce*, namun mengakui kompleksitas interpretasi model bagi pelaku usaha non-teknis. Penelitian lain yang mengaplikasikan *Decision Tree* dalam Konteks pemasaran digital juga dilakukan oleh [9]

Dalam konteks optimasi algoritma, beberapa penelitian telah menunjukkan efektivitas metode *Grid search*[10], menunjukkan peningkatan akurasi *Decision Tree* dalam prediksi diabetes dari 76% menjadi 81% melalui optimasi *hyperparameter*. [11] melaporkan peningkatan performa yang signifikan menggunakan *Grid search* untuk klasifikasi penerima beasiswa sebesar 99%. [4] juga membuktikan efektivitas *Grid search* dalam meningkatkan akurasi *Decision Tree* untuk klasifikasi stunting dari 70,2% menjadi 82,8%. Namun, aplikasi kombinasi *Decision Tree* dan *Grid search* khusus untuk analisis penjualan produk kerajinan di *e-commerce* masih terbatas dan belum banyak dieksplorasi.

Berdasarkan analisis gap penelitian tersebut, empat celah pengetahuan utama teridentifikasi, yaitu: (1) belum ada penelitian yang secara khusus melakukan analisis faktor-faktor penentu keberhasilan penjualan produk kerajinan tangan di *e-commerce* dengan pendekatan *machine learning* yang teroptimasi; (2) implementasi *Decision Tree* dalam kasus ini belum diintegrasikan dengan teknik *hyperparameter tuning* yang sistematis seperti *Grid Search Cross-validation*; (3) kebutuhan akan model prediktif yang tidak hanya akurat tetapi juga mudah diinterpretasi oleh pelaku UMKM; (4) tantangan yang dihadapi dalam penelitian ini membutuhkan solusi kreatif seperti penggunaan *dataset proxy* yang relevan.

Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan kerangka analisis yang integratif, menggabungkan algoritma *Decision Tree* dengan optimasi *Grid Search Cross-validation* khusus pada konteks penjualan kerajinan tangan di *e-commerce*. Pendekatan ini dipilih karena *Decision Tree* menawarkan tingkat interpretabilitas yang tinggi melalui struktur pohon keputusan yang dapat diubah menjadi aturan bisnis yang sederhana, sementara *Grid Search* memastikan pencarian parameter optimal secara eksaustif[3],[12]. *Dataset* yang digunakan didasarkan pada kesamaan karakteristik produk dalam kategori '*artes*' dengan produk kerajinan tangan umum dalam konteks nilai estetika, keunikan, dan target pasar yang niche.

Tujuan penelitian ini adalah: (1) Membangun model klasifikasi untuk menganalisis keberhasilan penjualan produk kerajinan tangan di platform *e-commerce* menggunakan algoritma *Decision Tree*; (2) Mengoptimasi performa model melalui teknik *Grid Search Cross-validation* dengan eksplorasi 576 kombinasi parameter; (3) Mengidentifikasi faktor-faktor dominan yang mempengaruhi keberhasilan penjualan melalui analisis *feature importance*; (4) Merumuskan rekomendasi strategis berbasis data untuk mendukung pengambilan keputusan bisnis. Penelitian ini juga bertujuan untuk memahami bagaimana perubahan dalam aspek ekonomi mempengaruhi operasional UMKM kerajinan. Selain itu, penelitian ini mencakup adaptasi *dataset* global sebagai solusi keterbatasan data lokal, serta penyediaan *insight* bisnis yang dapat diimplementasikan melalui model yang mudah diinterpretasi. Penelitian ini bertujuan untuk membangun model klasifikasi keberhasilan penjualan produk kerajinan tangan pada platform *e-commerce* menggunakan algoritma *Decision Tree* yang dioptimasi dengan metode *Grid Search Cross-validation*. Penelitian ini diharapkan mampu mengidentifikasi faktor-faktor dominan yang mempengaruhi keberhasilan penjualan melalui analisis *feature importance* serta menghasilkan model yang akurat dan mudah diinterpretasikan sebagai dasar pengambilan keputusan strategis bagi pelaku UMKM kerajinan.

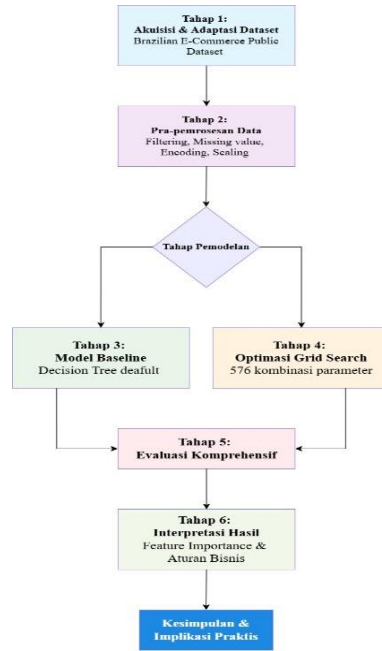
2. METODOLOGI PENELITIAN

2.1 Alur Penelitian

Penelitian ini mengimplementasikan pendekatan analitik sistematis melalui enam tahapan yang saling terhubung. Tahap pertama melibatkan akuisisi dan integrasi *dataset Olist* dari platform *e-commerce* Brasil dengan fokus pada kategori '*artes*' sebagai *proxy* produk kerajinan tangan. Tahap kedua mencakup proses *preprocessing* yang komprehensif, yang meliputi pembersihan data, penanganan nilai hilang, encoding variabel kategorikal, serta normalisasi fitur numerik untuk memastikan kualitas data optimal sebelum pemodelan.

Tahap ketiga adalah pembagian *dataset* dengan rasio 80:20 menjadi data pelatihan dan pengujian, menggunakan teknik *stratified sampling* untuk menjaga keseimbangan distribusi kelas pada data yang tidak seimbang. Tahap keempat merupakan inti penelitian, yaitu implementasi algoritma *Decision Tree* yang dioptimalkan melalui *Grid Search Cross-validation* dengan eksplorasi 576 kombinasi parameter untuk menemukan konfigurasi terbaik. Tahap kelima adalah evaluasi menyeluruh menggunakan metrik *akurasi*, *presisi*, *recall*, *F1-score*, dan *confusion matrix*. Tahap akhir merupakan interpretasi hasil, serta visualisasi *Decision Tree* untuk menghasilkan rekomendasi bisnis yang dapat diimplementasikan.

Alur ini dirancang untuk memastikan validitas metodologis, reproduktibilitas eksperimen, serta relevansi praktis bagi pemangku kepentingan UMKM kerajinan tangan.



Gambar 1. Alur pelaksanaan Penelitian

Berdasarkan Gambar 1, penelitian dimulai dari tahap pengumpulan *dataset*, dilanjutkan dengan *preprocessing* data untuk memilih fitur yang relevan. Selanjutnya dilakukan pembagian data menjadi data *training* dan data *testing*. Model kemudian dibangun menggunakan algoritma *Decision Tree* dan dioptimasi menggunakan *Grid Search Cross-Validation*. Tahap berikutnya adalah evaluasi model menggunakan *confusion matrix* dan *F1-score*, serta analisis *feature importance* untuk mengetahui faktor yang paling berpengaruh terhadap tingkat penjualan..

2.2 Dataset dan preprocessing

Dataset yang digunakan pada penelitian ini berasal dari platform *e-commerce* Brasil (*Olist dataset*) dengan pemfilteran pada kategori '*artes*' sebagai representasi produk kerajinan tangan agar sesuai dengan fokus penelitian pada produk kerajinan tangan. Proses ini juga memperbaiki eksperimen awal sehingga menghasilkan model yang lebih representatif. Pemilihan *dataset* ini didasarkan pada kesamaan karakteristik produk kerajinan dalam konteks bisnis *e-commerce* secara global, sehingga mengatasi keterbatasan data lokal namun tetap relevan secara kontekstual. Empat *dataset* utama yang digunakan meliputi: tabel pesanan (*order status*), detail produk (*product details*), metode pembayaran (*payment methods*), dan atribut produk (*product attributes*).

Proses pra-pengolahan data *scientist* menangani data melalui tiga tahap utama: (1) Penanganan data hilang pada kolom tanggal pengiriman dengan imputasi berbasis estimasi; (2) Variabel kategorikal diubah menggunakan *Label Encoding*; (3) Variabel numerik di Normalisasi dengan *StandardScaler*. Seleksi fitur menghasilkan *dataset* final yang memiliki karakteristik seperti pada Tabel 1. Normalisasi variabel numerik dengan *StandardScaler* dilakukan untuk menyalurkan skala fitur. Pemilihan metode pra-pemrosesan dan transformasi data adalah langkah penting yang secara langsung mengamati hasil model klasifikasi, seperti yang ditunjukkan dalam penelitian[13], yang berhasil meningkatkan akurasi identifikasi kualitas beras hingga 97% dengan menggunakan transformasi data eksponensial. Prinsip yang sama tentang pentingnya pemberian bobot dan transformasi fitur untuk meningkatkan akurasi juga dijelaskan oleh[14] dalam konteks pengembangan algoritma *K-NN*.

Tabel 1. Karakteristik *Dataset* Setelah *Preprocessing*

Parameter	Nilai	Keterangan
Total Sampel	209 transaksi	Setelah filtering kategori ' <i>artes</i> '
Jumlah Fitur	5 variabel	Hasil seleksi fitur akhir yang digunakan dalam model
Variabel Target	<i>is_delivered</i>	Biner (1=sukses 0=gagal)
Distribusi Kelas	94,26% sukses, 5,74% gagal	Tidak seimbang (<i>imbalanced</i>)
Data <i>Training</i>	167 sampel (80%)	Untuk pelatihan model
Data <i>Testing</i>	42 sampel (20%)	Untuk evaluasi akhir

Berdasarkan Tabel 1, *dataset* yang digunakan dalam penelitian ini terdiri dari 209 sampel transaksi setelah proses *preprocessing* pada kategori produk *artes*. Distribusi kelas menunjukkan kondisi tidak seimbang dengan dominasi kelas berhasil sebesar 94,26%. *Dataset* kemudian dibagi menjadi data *training* sebesar 80% dan data *testing* sebesar 20% untuk memastikan proses pelatihan dan evaluasi model berjalan optimal.

2.3 Pembagian Data

Dataset yang telah melalui tahap *preprocessing* selanjutnya dibagi menjadi data *training* dan data *testing* menggunakan metode *hold-out*. Proporsi pembagian data yang digunakan adalah 80% untuk data *training* dan 20% untuk data *testing*. Data *training* digunakan untuk membangun model klasifikasi menggunakan algoritma *Decision Tree*, sedangkan data *testing* digunakan untuk mengevaluasi performa model yang dihasilkan. Dari total 209 data, sebanyak 167 data digunakan sebagai data *training* dan 42 data digunakan sebagai data *testing*.

2.4 Penerapan *Decision Tree*

Algoritma *Decision Tree* merupakan metode klasifikasi berbasis struktur pohon yang bekerja dengan membagi data secara rekursif berdasarkan atribut tertentu untuk meningkatkan homogenitas kelas pada setiap node. Proses pemisahan data dilakukan menggunakan kriteria seperti *Gini Index* atau *Entropy*. Setiap percabangan menghasilkan aturan keputusan yang mudah diinterpretasikan sehingga algoritma ini banyak digunakan dalam analisis prediktif yang membutuhkan transparansi model. Algoritma *Decision Tree* digunakan untuk membangun model klasifikasi tingkat keberhasilan penjualan berdasarkan fitur yang telah dipilih pada tahap *preprocessing*. Proses pembentukan pohon keputusan dilakukan dengan memilih atribut terbaik pada setiap node menggunakan ukuran impurity berupa *entropy* dan *information gain*. Proses *splitting* dilakukan secara rekursif hingga memenuhi parameter penghentian tertentu sehingga dihasilkan struktur pohon yang mampu memisahkan data berdasarkan pola klasifikasi yang terbentuk. Model dibangun menggunakan data *training* sebelum dilakukan proses optimasi hyperparameter menggunakan metode *Grid Search Cross-Validation*. Algoritma *Decision Tree* dipilih sebagai model dasar karena memiliki tingkat interpretabilitas yang tinggi melalui struktur pohon dapat diterjemahkan menjadi aturan *IF-THEN*, sehingga cocok untuk konteks analisis UMKM [12],[6]

2.5 Optimasi *Grid Search Cross-Validation*

Grid Search Cross-validation merupakan teknik optimasi *hyperparameter* yang bekerja dengan menguji seluruh kombinasi parameter secara sistematis menggunakan skema *cross-validation* untuk memperoleh performa model terbaik. Optimasi model dilakukan menggunakan metode *Grid Search Cross-Validation* untuk memperoleh kombinasi parameter terbaik pada algoritma *Decision Tree*. Metode ini bekerja dengan menguji berbagai kombinasi parameter secara sistematis dan mengevaluasi performa model menggunakan teknik *k-fold cross-validation*. Parameter yang diuji dalam penelitian ini meliputi *criterion*, *max_depth*, *min_samples_split*, *min_samples_leaf*, dan *max_features*. Proses optimasi dilakukan menggunakan metrik evaluasi *F1-score macro* untuk mengatasi ketidakseimbangan data pada kelas target. Hasil optimasi digunakan untuk membangun model akhir dengan performa yang lebih optimal. Optimasi *hyperparameter* dilakukan menggunakan *Grid Search Cross-validation (GridSearchCV)* dengan konfigurasi seperti pada **Tabel 2**.

Tabel 2. Konfigurasi *Grid Search Cross-validation*

Parameter	Nilai yang Diuji	Keterangan
Estimator	<i>DecisionTreeClassifier</i>	Algoritma dasar
<i>max_depth</i>	[3,5,7,10,15,None]	Kedalaman pohon
<i>Criterion</i>	['gini','entropy']	Kriteria pemisahan
<i>min_samples_split</i>	[2,5,10,15]	Minimum sampel untuk split
<i>min_samples_leaf</i>	[1,2,4,6]	Minimum sampel di leaf
<i>max_features</i>	['sqrt','log2',None]	Strategi seleksi fitur
<i>Cross-validation</i>	5-fold	Validasi internal
Metrik Evaluasi	Scoring='f1_macro'	Untuk data tidak seimbang
Total Kombinasi	576 kombinasi	Eksplorasi menyeluruh

Tabel 2. menunjukkan konfigurasi parameter yang digunakan pada proses optimasi *Grid Search Cross-validation*. Beberapa parameter utama yang diuji meliputi *max_depth*, *criterion*, *min_samples_split*, *min_samples_leaf*, dan *max_features*. Total kombinasi parameter yang diuji sebanyak 576 kombinasi untuk memperoleh konfigurasi model *Decision Tree* terbaik berdasarkan nilai *F1-score*. Proses optimasi menggunakan *5-fold cross-validation* pada data pelatihan (X_{train} , y_{train}) Untuk setiap kombinasi parameter, model dinilai menggunakan *F1-macro* sebagai metrik utama karena sensitivitasnya terhadap ketidakseimbangan kelas. Kombinasi dengan skor validasi rata-rata tertinggi dipilih sebagai konfigurasi optimal. Pendekatan *Grid Search* dalam mencari parameter terbaik secara sistematis terbukti efektif tidak hanya untuk algoritma machine learning, tetapi juga dalam mengoptimalkan parameter model statistik yang lebih sederhana [15].

2.4 Metrik Evaluasi dan Analisis *Feature importance*

Evaluasi model dilakukan dengan menggunakan empat metrik utama seperti *Akurasi*, *Presisi*, *Recall*, dan *F1-Score*. Memilih berbagai metrik ini penting untuk menghindari kesimpulan yang tidak akurat, terutama pada data yang tidak seimbang, karena akurasi saja bisa memberikan gambaran yang salah. *Confusion matrix* digunakan untuk menganalisis lebih dalam performa model per kelas.

Analisis pentingnya fitur dijelaskan menggunakan atribut *feature importances* dari model *Decision Tree* yang telah dioptimasi. Metrik ini mengukur kontribusi relatif setiap fitur dalam mengambil keputusan, dihitung berdasarkan penurunan ketidakmurnian (*Gini impurity*) yang dihasilkan oleh setiap fitur di semua titik split dalam pohon. Formula perhitungannya adalah:

$$\text{Importance}(X_i) = \frac{\sum (\text{impurity decrease by } X_i \text{ at split } t)}{\sum (\text{impurity decrease by all feature})} \quad (1)$$

Analisis ini membantu mengidentifikasi faktor-faktor kritis yang mempengaruhi keberhasilan penjualan, memberikan wawasan bisnis yang dapat langsung diterapkan oleh UMKM kerajinan.

2.5 Environment dan Tools

Eksperimen dilakukan menggunakan *Python 3.8* dengan *library* utama: *scikit-learn* untuk algoritma *machine learning*, *pandas* dan *numpy* untuk manipulasi data, serta *matplotlib* dan *seaborn* untuk visualisasi. Implementasi dilakukan pada komputer dengan spesifikasi Intel Core i5-9300H, RAM 8GB, dan Windows 11. Seluruh kode dan konfigurasi parameter didokumentasikan secara sistematis untuk memastikan reproduktibilitas penelitian.

3. HASIL DAN PEMBAHASAN

3.1 Implementasi dan Hasil Pengujian Model

3.1.1 Sampel Dataset

Setelah melalui tahap preprocessing, diperoleh dataset yang telah diseleksi berdasarkan kategori produk kerajinan tangan serta fitur yang relevan. Contoh sampel data yang digunakan dalam penelitian ditampilkan pada Gambar 3.

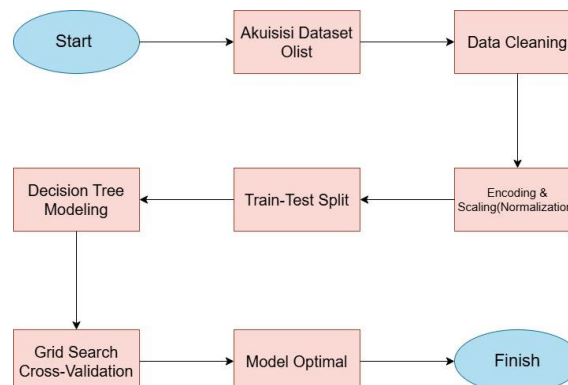
	tinggi_produk_cm	nilai_pengangkutan	harga	waktu_pengiriman_hari	telah_dikirim
357	6.0	33.23	110.99	19.0	1
856	12.0	17.60	48.65	17.0	1
1060	6.0	14.37	139.99	7.0	1
1097	6.0	15.01	99.99	8.0	1
1904	6.0	23.14	79.99	7.0	1
1906	6.0	14.54	99.99	20.0	1
2424	6.0	15.10	36.85	8.0	1
4639	6.0	12.03	99.99	18.0	1
4851	6.0	11.89	79.99	4.0	1
4964	6.0	57.47	99.99	16.0	1

Gambar 3. Contoh Data Sampel *Dataset* Setelah Proses *Preprocessing*

Berdasarkan Gambar 3., menampilkan contoh data sampel *dataset* setelah melalui tahap *preprocessing* yang meliputi pembersihan data, transformasi variabel, dan seleksi fitur. *Dataset* terdiri dari beberapa atribut utama yang merepresentasikan karakteristik produk dan proses transaksi, yaitu *tinggi_produk_cm*, *nilai_pengangkutan*, *harga*, *waktu_pengiriman_hari*, serta variabel target *telah_dikirim*. Setiap baris merepresentasikan satu transaksi produk kerajinan pada platform *e-commerce*. Variabel target *telah_dikirim* digunakan sebagai label klasifikasi dengan nilai biner, yaitu 1 menunjukkan transaksi berhasil dan 0 menunjukkan transaksi tidak berhasil. Visualisasi ini memberikan gambaran awal mengenai struktur data yang digunakan sebagai input dalam proses pelatihan dan pengujian model *Decision Tree*.

3.1.2 Tahapan Implementasi Algoritma *Decision Tree* dalam Penelitian

Dalam penelitian ini, penerapan algoritma *Decision Tree* dilakukan melalui tahapan sistematis untuk memastikan model mampu menyelesaikan permasalahan klasifikasi keberhasilan penjualan produk kerajinan secara optimal. Secara umum, alur teknis tahap ini dapat di lihat pada gambar 2



Gambar 2. Tahapan Implementasi Algoritma *Decision Tree* dalam penelitian

Gambar 2. menunjukkan tahapan implementasi algoritma *Decision Tree* yang digunakan dalam penelitian ini. Proses diawali dengan penggunaan *dataset* Olist yang telah melalui tahap *preprocessing* untuk memastikan kualitas data sebelum pemodelan dilakukan. Selanjutnya *dataset* dibagi menjadi data pelatihan dan data pengujian menggunakan metode *train-test split*. Model *Decision Tree* kemudian dibangun dan dioptimasi menggunakan *Grid Search Cross-validation* untuk memperoleh kombinasi *hyperparameter* terbaik. Tahapan optimasi bertujuan meningkatkan performa klasifikasi sekaligus mengurangi risiko *overfitting*. Model optimal yang dihasilkan selanjutnya digunakan pada tahap evaluasi performa dan analisis faktor penting yang memengaruhi keberhasilan penjualan produk kerajinan. Alur ini memastikan proses pemodelan berlangsung secara sistematis, terstruktur, dan dapat direproduksi.

3.1.3 Performa Model *Baseline* dan Masalah *Overfitting*

Sebagai tahap awal, model *Decision Tree* dibangun menggunakan parameter default (*criterion='gini'* dan *max_depth=None*) untuk memperoleh gambaran performa dasar sebelum dilakukan proses optimasi. Model dengan parameter default memungkinkan pohon tumbuh tanpa pembatasan kedalaman, sehingga berpotensi menghasilkan struktur yang kompleks. Fenomena serupa dilaporkan dalam penelitian[3], Sebagai tahap awal, model *Decision Tree* dibangun menggunakan parameter default (*criterion='gini'* dan *max_depth=None*) untuk memperoleh gambaran performa dasar sebelum dilakukan proses optimasi. Model dengan parameter default memungkinkan pohon tumbuh tanpa pembatasan kedalaman, sehingga berpotensi menghasilkan struktur yang kompleks.

3.1.4 Hasil Optimasi *Grid Search Cross-validation*

Proses optimasi menggunakan *Grid Search Cross-validation* dengan 5-fold berhasil mengeksplorasi 576 kombinasi parameter dan menemukan konfigurasi optimal: *max_depth=None*, *criterion='entropy'*, *min_samples_split=2*, *min_samples_leaf=4*, *max_features=None*. Meskipun parameter *max_depth* tidak dibatasi (*None*), kompleksitas model tetap terkontrol melalui parameter *min_samples_leaf=4* sehingga pembentukan pohon tidak berlebihan dan risiko *overfitting* dapat diminimalkan, terutama pada *dataset* berukuran kecil. Konfigurasi ini menghasilkan skor validasi rata-rata (*F1-macro*) sebesar 0,8403 diperoleh dari proses *cross-validation* pada tahap *Grid Search*. Sedangkan nilai *F1-score* sebesar 0.83 merupakan hasil evaluasi model pada data testing. seperti ditunjukkan pada Tabel 3.

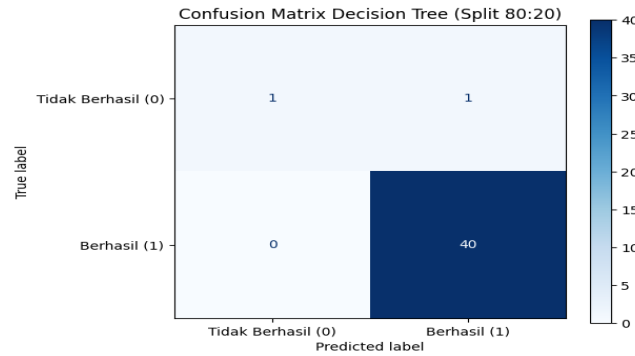
Tabel 3. Perbandingan Peforma Model Sebelum dan Sesudah Optimasi

Aspek	Model Baseline (Default)	Model Teroptimasi (<i>Grid Search</i>)	Interpretasi
Parameter (<i>max_depth</i>)	None (tidak terbatas)	None	Model lebih sederhana
Kompleksitas (Jumlah Node)	15	6	Struktur lebih ringkas
Skor Validasi (<i>F1_macro</i>)	-	0,8403	Estimasi kinerja baik
Akurasi pada Data Uji	Tidak diukur	97,61%	Peforma nyata
<i>F1-Score</i> pada Data Uji	Tidak diukur	0,83	Baik untuk data tidak seimbang

Pada Tabel 3. Dijelaskan bahwa Peningkatan signifikan setelah proses optimasi menunjukkan efektivitas metode *Grid Search* dalam mengatasi masalah *overfitting*, sejalan dengan temuan [4] dan [10] yang berhasil meningkatkan akurasi *Decision Tree* dalam klasifikasi stunting dari 70,2% menjadi 82,8% melalui optimasi serupa. Hasil penurunan jumlah node dari 15 menjadi 6 menunjukkan bahwa model yang lebih sederhana sering kali memiliki kemampuan generalisasi yang lebih baik. Penelitian ini berfokus pada model hasil optimasi sehingga evaluasi baseline tidak ditampilkan secara terpisah.

Meskipun akurasi model awal sudah cukup tinggi, proses optimasi *Grid Search* tetap dilakukan untuk memastikan bahwa parameter yang digunakan merupakan konfigurasi optimal serta untuk mengurangi risiko *overfitting*. Selain itu, optimasi juga bertujuan meningkatkan stabilitas model melalui proses *cross-validation* sehingga performa model tidak hanya bergantung pada satu pembagian data saja. Dengan demikian, penggunaan *Grid Search* tidak hanya berfokus pada peningkatan akurasi, tetapi juga pada validitas dan generalisasi model. Optimasi tidak selalu bertujuan meningkatkan akurasi secara signifikan, tetapi memastikan model berada pada konfigurasi parameter terbaik dan menghasilkan performa yang lebih stabil. Optimasi *Grid Search* dalam penelitian ini tidak hanya berfokus pada peningkatan akurasi, tetapi juga pada pencarian kombinasi parameter terbaik untuk meningkatkan stabilitas model pada *dataset* yang relatif kecil dan tidak seimbang.

3.1.5 Analisis *Confusion Matrix*



Gambar 4. Visualisasi *Confusion Matrix Decision Tree*

Pada Gambar 4. hasil *confusion matrix* menunjukkan bahwa dari total 42 data pengujian, model berhasil mengklasifikasikan 40 transaksi berhasil (*True Positive*) dengan benar. Pada kelas tidak berhasil, terdapat 1 data yang terklasifikasi dengan benar sebagai tidak berhasil (*True Negative*) dan 1 data yang salah diprediksi sebagai berhasil (*False Positive*). Tidak terdapat kesalahan prediksi pada kelas berhasil (*False Negative* = 0), yang menunjukkan bahwa model mampu mengenali seluruh transaksi berhasil secara akurat. Secara keseluruhan, hasil ini menghasilkan tingkat akurasi sebesar 97,61% dengan nilai *F1-macro* sebesar 0,8403 dan ini sejalan dengan penelitian [16] Meskipun performa model sangat baik dalam mendeteksi kelas berhasil, masih terdapat keterbatasan dalam mendeteksi kelas tidak berhasil yang jumlahnya jauh lebih sedikit, sehingga menunjukkan karakteristik *dataset* yang tidak seimbang (*imbalanced dataset*).

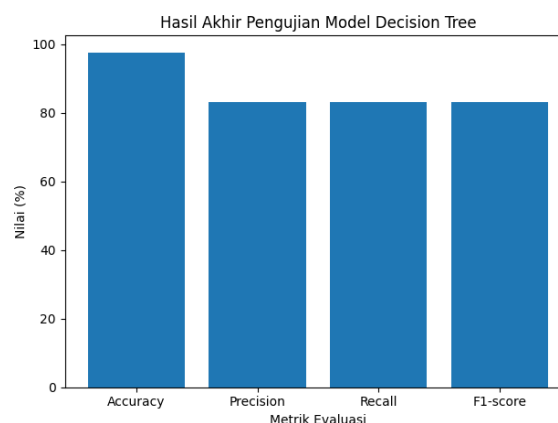
3.1.6 Hasil Pengujian Model *Decision Tree*

Setelah dilakukan proses optimasi menggunakan *Grid Search Cross-validation*, model *Decision Tree* terbaik selanjutnya diuji menggunakan data testing untuk mengetahui performa akhir model. Hasil pengujian dirangkum dalam bentuk tabel dan grafik untuk memberikan gambaran komprehensif terhadap kinerja algoritma.

Tabel 4. Hasil Akhir Pengujian Model *Decision Tree*

Metrik Evaluasi	Nilai
Accuracy	97.61%
Precision	0.83
Recall	0.83
F1-Score (Macro)	0.83

Tabel 4. menunjukkan hasil akhir pengujian model *Decision Tree* optimal pada data testing. Model menghasilkan akurasi sebesar 97,61% dengan nilai *precision*, *recall*, dan *F1-score* yang seimbang. Hasil ini menunjukkan bahwa model mampu melakukan klasifikasi dengan baik meskipun *dataset* memiliki distribusi kelas yang tidak seimbang.



Gambar 5. Grafik Hasil Akhir Pengujian Model *Decision Tree*

Gambar 5. menampilkan visualisasi hasil akhir pengujian model *Decision Tree* pada data testing. Terlihat bahwa seluruh metrik evaluasi menunjukkan performa yang tinggi dan relatif seimbang. Hal ini mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik serta mampu mempertahankan performa klasifikasi pada kondisi *dataset* yang tidak seimbang.

3.2 Analisis Faktor Penentu Keberhasilan penjualan

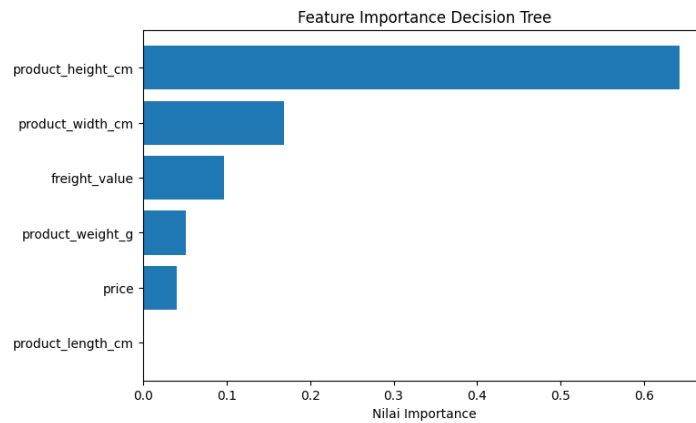
3.2.1 *Feature importance* dan Identifikasi Faktor Kunci

Analisis *Feature importance* menunjukkan bahwa hanya 6 dari 14 fitur yang berkontribusi secara signifikan terhadap prediksi, seperti terlihat pada Tabel 4 dan Gambar 6.

Tabel 4. *Feature importance* Model Optimal

Peringkat	Fitur	Importance Score	Kontribusi	Interpretasi Bisnis
1	<i>product_height_cm</i>	0,6423	64,23%	Dimensi tinggi produk sangat memengaruhi keberhasilan transaksi
2	<i>Produk_width_cm</i>	0,1690	16,90%	Lebar produk berpengaruh pada proses pengiriman
3	<i>freight_value</i>	0,0966	9,66%	Biaya pengiriman memengaruhi keputusan transaksi
4	<i>Produk_weight_g</i>	0,0514	5,14%	Berat produk berpengaruh terhadap logistik
5	<i>price</i>	0,0407	4,07%	Harga produk memiliki pengaruh relatif kecil
6	<i>Product length cm</i>	0,0000	0,000%	Tidak memberikan kontribusi dalam model

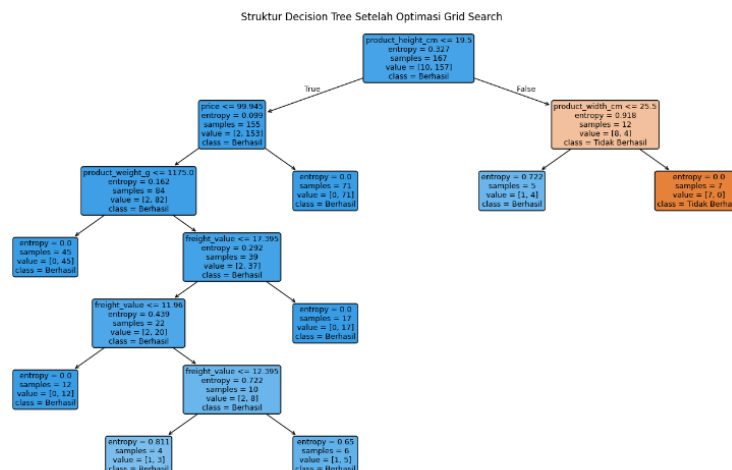
Pada Tabel 4. Hasil analisis menunjukkan bahwa atribut dimensi produk, khususnya *product_height_cm*, memiliki kontribusi paling dominan sebesar 64,23% dalam proses klasifikasi. Hal ini mengindikasikan bahwa karakteristik fisik produk memiliki pengaruh yang lebih besar dibandingkan faktor harga dalam menentukan keberhasilan transaksi pada kategori *artes*.



Gambar 6. Visualisasi *Feature importance*

Berdasarkan Gambar 6., terlihat bahwa atribut *product_height_cm* memiliki nilai kepentingan paling tinggi dibandingkan fitur lainnya dengan kontribusi sebesar 64,23%. Hal ini menunjukkan bahwa tinggi produk menjadi faktor paling dominan dalam proses klasifikasi keberhasilan transaksi. Atribut berikutnya yang berpengaruh adalah *product_width_cm* dan *freight_value*, meskipun kontribusinya jauh lebih kecil dibandingkan fitur utama. Sementara itu, *product_length_cm* tidak memberikan kontribusi terhadap pembentukan model, yang ditunjukkan dengan nilai importance sebesar 0,00%. Temuan ini mengindikasikan bahwa karakteristik dimensi fisik produk lebih berpengaruh dibandingkan faktor harga dalam menentukan hasil klasifikasi [8], [2].

3.2.2 Interpretasi Struktur Pohon Keputusan



Gambar 7. Visualisasi Pohon Keputusan

Berdasarkan Gambar 7., struktur *Decision Tree* hasil optimasi *Grid Search* menunjukkan bahwa atribut *product_height_cm* menjadi akar pohon (*root node*), yang berarti fitur ini merupakan faktor paling dominan dalam menentukan klasifikasi keberhasilan transaksi. Jika tinggi produk $\leq 19,5$ cm, mayoritas transaksi diklasifikasikan sebagai *Berhasil*, dengan pemisahan lanjutan berdasarkan variabel *price*, *product_weight_g*, dan *freight_value*. Sebaliknya, jika *product_height_cm* $> 19,5$ cm, model melakukan pemisahan berdasarkan *product_width_cm*. Pada cabang ini terlihat kecenderungan kelas *Tidak Berhasil*, terutama ketika lebar produk melebihi batas tertentu. Struktur pohon yang relatif ringkas (*node* lebih sedikit dibanding model awal) menunjukkan bahwa proses optimasi berhasil menyederhanakan model tanpa mengurangi kemampuan klasifikasi. Aturan-aturan yang terbentuk dapat diterjemahkan ke dalam logika bisnis berbasis *IF-THEN*, sehingga model tidak hanya akurat tetapi juga mudah diinterpretasikan untuk mendukung pengambilan keputusan pada UMKM, dan juga untuk mengatasi keterbatasan model *black-box* seperti *Random Forest* atau *XGBoost* [11], [17].

3.1 Pembahasan Integratif

3.3.1 Keunggulan *Decision Tree* + *Grid Search* vs Algoritma Lain

Perbandingan dengan penelitian sebelumnya menunjukkan kelebihan dari kombinasi *Decision Tree* dan *Grid Search* dalam konteks analisis usaha kecil menengah (UMKM) kerajinan:

- Interpretabilitas dan Akurasi: Meskipun algoritma *ensemble* seperti *Random Forest* mencapai tingkat akurasi yang lebih tinggi (89,7%) menurut penelitian[18], model tersebut sulit diinterpretasi oleh pelaku usaha non-teknis. Temuan serupa dari perbandingan algoritma klasifikasi menunjukkan bahwa meskipun *Random Forest* atau *XGBoost*[19] sering kali lebih baik dalam akurasi, *Decision Tree* tetap dipilih utama ketika kemudahan dalam memahami hasil dan kemudahan dalam menerjemahkan ke dalam aturan bisnis menjadi prioritas utama[20]. Sementara itu, *Decision Tree* yang dioptimasi menawarkan keseimbangan yang ideal antara akurasi 97.61% dan interpretabilitas tinggi. Dalam konteks ini, meskipun meningkatnya akurasi dapat menghasilkan prediksi yang lebih baik, interpretabilitas tetap menjadi faktor penting bagi pelaku usaha. Oleh karena itu, pendekatan yang memadukan keduanya dianggap lebih efektif untuk mencapai hasil yang seimbang dalam analisis UMKM kerajinan.
- Efisiensi komputasi: proses optimasi dengan 576 kombinasi parameter memakan waktu lebih singkat dibandingkan teknik kompleks seperti *Bayesian Optimization* atau *HyperOpt*, namun tetap menghasilkan konfigurasi optimal. Hal ini sejalan dengan temuan [21] yang membandingkan berbagai teknik *tuning hyperparameter*. Efisiensi ini penting untuk konteks UMKM dengan sumber daya komputasi terbatas.
- Relevansi untuk Data Kecil: *Dataset* dengan 209 sampel cukup untuk membangun model yang kuat setelah proses optimasi, membuktikan bahwa ukuran data bukan hambatan utama jika *preprocessing* dan optimasi dilakukan dengan tepat. Ini sesuai dengan penelitian sebelumnya yang berhasil membuat model klasifikasi yang tepat menggunakan *Decision Tree* pada *dataset* dengan jumlah sampel yang tidak terlalu besar[22],[23]

3.3.2 Implikasi Praktis untuk UMKM Kerajinan

Temuan penelitian memberikan implikasi praktis yang langsung dapat diterapkan:

- Optimasi Dimensi Produk: Fokus pada tinggi produk optimal ($\leq 19,5$ cm) karena *product_height_cm* merupakan faktor paling dominan (64,23%). Pengelolaan dimensi yang efisien dapat meningkatkan peluang keberhasilan transaksi sekaligus menekan risiko kendala logistik.
- Pengendalian Lebar dan Berat Produk: Selain tinggi produk, *product_width_cm* dan *product_weight_g* juga berkontribusi dalam proses klasifikasi. Oleh karena itu, UMKM perlu memastikan ukuran dan berat produk tetap proporsional agar tidak meningkatkan biaya distribusi.
- Strategi Pengelolaan Biaya Pengiriman: Variabel *freight_value* berperan dalam percabangan keputusan model. Pengendalian ongkos kirim misalnya melalui efisiensi kemasan atau kerja sama dengan jasa logistik dapat membantu menjaga tingkat keberhasilan transaksi.
- Pendekatan *Value-Based Pricing*: Harga (*price*) bukan faktor utama dalam model (kontribusi relatif kecil dibanding dimensi produk). Hal ini menunjukkan bahwa keberhasilan transaksi lebih dipengaruhi oleh karakteristik fisik produk daripada sekadar harga, sehingga UMKM dapat menekankan nilai estetika, kualitas, dan keunikan produk dalam strategi pemasaran.

Temuan ini memperkuat dan memperluas penelitian sebelumnya:

- Konsisten dengan mengenai efektivitas *Grid Search* untuk optimasi *Decision Tree*.
- Melengkapi penelitian dengan menawarkan pendekatan kuantitatif untuk strategi digital UMKM kerajinan.
- Berbeda dengan penelitian [8] yang menemukan reputasi penjual sebagai faktor dominan dalam kesuksesan produk *e-commerce* pada umumnya, hasil penelitian ini menunjukkan analisis yang lebih spesifik terhadap produk kerajinan tangan, yang sangat bergantung pada atribut fisiknya..
- Sejalan dengan penelitian tentang keunggulan *Decision Tree* dibandingkan Logistic Regression dalam klasifikasi berbasis gejala.serta penelitian [22] yang juga menerapkan *Decision Tree* untuk analisis perilaku konsumen.

3.3.4 Validasi dan Transformasi Data

Tingkat akurasi sebesar 97.61% yang dicapai dalam penelitian ini memperkuat temuan sebelumnya tentang efektivitas optimasi teknis dalam machine learning[13] yang menggunakan transformasi data eksponensial untuk meningkatkan

akurasi identifikasi mutu beras hingga 97%. Selain itu, pendekatan sistematis dalam meningkatkan performalgoritma melalui modifikasi fungsi juga selaras dengan penelitian[14] yang berhasil meningkatkan akurasi K-NN dengan pembobotan eksponensial. Meskipun domain dan metode teknis yang digunakan berbeda, ketiga penelitian (termasuk penelitian ini) sama-sama menunjukkan bahwa pendekatan optimasi yang terstruktur baik melalui tuning *hyperparameter* (*Grid Search*) maupun transformasi data dapat menghasilkan peningkatan performa model yang signifikan.

4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan algoritma *Decision Tree* yang dioptimasi menggunakan metode *Grid Search* *Cross-validation* untuk menganalisis faktor penentu keberhasilan penjualan produk kerajinan pada platform *e-commerce*. Hasil pengujian menunjukkan bahwa model optimal dengan parameter *max_depth* = None menghasilkan akurasi sebesar 97,61% dan nilai *F1-score* sebesar 0,83, yang menunjukkan peningkatan performa dibandingkan model baseline serta mampu mengurangi risiko *overfitting*. Analisis *feature importance* mengidentifikasi dimensi tinggi produk sebagai faktor paling dominan dengan kontribusi sebesar 64,23%, diikuti oleh lebar produk, biaya pengiriman, berat produk, dan harga produk. Struktur pohon keputusan yang dihasilkan juga lebih sederhana dan mudah diinterpretasikan sehingga dapat digunakan sebagai dasar rekomendasi strategi penjualan berbasis data bagi pelaku UMKM kerajinan dalam meningkatkan daya saing pada marketplace digital[16].

REFERENCES

- [1] N. L. W. S. Telagawathi, N. M. Suci, and K. K. Heryanda, "Strategi Transformasi Digital Umkm Kerajinan Tangan Pada Masa Pandemi Covid-19 Di Kabupaten Gianyar, Bali," *J. Ilmu Sos. dan Hum.*, vol. 11, no. 2, pp. 204–212, 2022, doi: 10.23887/jish.v11i2.39734.
- [2] M. G. F. Christine, M. Dachyar, and R. Nurcahyo, "Product Segmentation of Wooden Handicraft Micro , Small and Medium Enterprises (MSMEs) in Indonesia," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 598, pp. 1–8, 2019, doi: 10.1088/1757-899X/598/1/012063.
- [3] F. I. Sanjaya and A. D. Kalifa, "Precision Marketing Model using *Decision Tree* on SME *e-commerce* Case Study Orebae.com," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, no. 5, pp. 1033–1039, 2023.
- [4] R. S. T. Rahmayani and F. Budiman, "Analisa Optimasi *Grid Search* pada Algoritma Random Forest dan *Decision Tree* untuk Klasifikasi Stunting," *BUILD. Informatics, Technol. Sci.*, vol. 6, no. 3, pp. 1537–1546, 2024, doi: 10.47065/bits.v6i3.6128.
- [5] A. Basir and F. A. Tyas, "Akuisisi Pengetahuan untuk Sistem Pakar Pencegahan Stunting (SIPENTING) Menggunakan *Decision Tree* dan *Grid Search*," *Sist. Sist. Inf.*, vol. 14, no. 3, pp. 1165–1175, 2025.
- [6] D. I. Suranda and A. Nugroho, "KLASIFIKASI DATA PENJUALAN UNTUK MEMPREDIKSI TINGKAT PENJUALAN PRODUK MENGGUNAKAN METODE *DECISION TREE*," *J. TEKINKOM*, vol. 7, no. 1, pp. 370–376, 2024, doi: 10.37600/tekinkom.v7i1.1269.
- [7] S. O. Usnat, Y. P. K. Kelen, and K. J. T. Seran, "IMPLEMENTASI METODE *DECISION TREE* DALAM MEMPREDIKSI PENJUALAN PRODUK TERLARIS (STUDI KASUS : KREATIF MART ATAMBUA)," *Zo. J. Sist. Inf.*, vol. 7, no. 1, pp. 317–330, 2025.
- [8] F. U. Aulya and Kusnawi, "Evaluating Classification Models for Predicting Product Success in Indonesian *E-Commerce*," *J. Tek. Inform.*, vol. 6, no. 4, pp. 2723–2739, 2025, doi: <https://doi.org/10.52436/1.jutif.2025.6.4.5071>.
- [9] A. D. Prastika and Y. S. Nugroho, "Klasifikasi Efektivitas Iklan di *E-Commerce* dan Media Sosial terhadap Minat Beli Mahasiswa UMS dengan *Decision Tree*," *J. Pendidik. dan Teknol. Indones.*, vol. 5, no. 4, pp. 919–933, 2025, doi: <https://doi.org/10.52436/1.jpti.727>.
- [10] D. Anggreani, Hamdani, Nurmisda, and Lukman, "*Grid Search Hyperparameter* Analysis in Optimizing The *Decision Tree* Method for Diabetes Prediction," *Indones. J. Data Sci.*, vol. 5, no. 3, pp. 190–197, 2024, doi: <https://doi.org/10.56705/ijodas.v5i3.190> Research.
- [11] T. Suprapti, B. Nurhakim, B. W. A. Hermina, and V. A. Syahputra, "A *Decision Tree* Model with *Grid Search* Optimization for Scholarship Recipient Classification," *J. Tek. Inform.*, vol. 6, no. 5, pp. 3800–3813, 2025, doi: <https://doi.org/10.52436/1.jutif.2025.6.5.5235>.
- [12] A. Arista, "Comparison *Decision Tree* and Logistic Regression Machine Learning Classification Algorithms to determine Covid-19," *Sink. J. dan Penelit. Tek. Inform.*, vol. 6, no. 1, pp. 56–62, 2022, doi: <https://doi.org/10.33395/sinkron.v7i1.11243>.
- [13] M. J. Vikri, I. W. Prastya, Dwi, U. P. Sanjaya, and M. A. Barata, "RICE QUALITY IDENTIFICATION FOR INDONESIAN FOOD STANDARDS BASED ON ELECTRONIC NOSE," *J. INOVTEK POLBENG - SERI Inform.*, vol. 10, no. 1, pp. 49–60, 2025, doi: <https://doi.org/10.35314/0y0xct32>.
- [14] M. J. Vikri and R. Rohmah, "Penerapan Fungsi Exponential Pada Pembobotan Fungsi Jarak Euclidean Algoritma K-Nearest Neighbor," *Gener. J.*, vol. 6, no. 2, pp. 98–105, 2022, doi: <https://doi.org/10.29407/gj.v6i2.18070>.
- [15] W. A. Medyanti, M. Faisal, and H. Nurhayati, "Optimasi Metode Single Exponential Smoothing Dengan *Grid Search* Pada Prediksi Nilai Ekspor Migas," *SINTECH J.*, vol. 7, no. 1, pp. 59–69, 2024, doi: <https://doi.org/10.31598>.
- [16] Y. Aprillia, Z. Alawi, I. T. A. Aristia, and S. A. Ida, "Implementasi Algoritma Naive Bayes Dengan Feature Selection Backward Elimination Dalam Pengklasifikasian Status Penderita Stunting Pada Balita," *J. Multidiscip. Appl. Quantum Inf. Sci.*, vol. 4, no. 02, pp. 1–6, 2024.
- [17] A. H. Anshor and A. T. Zy, "Implementasi Metode *Decision Tree* pada Sistem Prediksi Status Kualitas Produk Minuman A," *J. Ilm. Inform. Glob.*, vol. 15, no. 1, pp. 17–22, 2024.
- [18] S. Ramadhan, Z. Alamin, M. Jannah, M. Akbar, and R. Fikriyansah, "Data-driven MSME Success Prediction Using *Decision Tree*-Based Machine Learning Techniques," *Journix J. Informatics Comput.*, vol. 1, no. 1, pp. 1–9, 2025, doi: 10.63866/journix.v1i1.3.
- [19] W. Nugraha and A. Sasongko, "*Hyperparameter* Tuning pada Algoritma Klasifikasi dengan *Grid Search*," *Sist. Sist. Inf.*, vol.

- 11, no. 2, pp. 391–401, 2022.
- [20] T. Sugihartono, Marini, K. Irfan, and H. A. Pradana, “Optimalisasi Kinerja Extreme Gradient Boosting melalui *Grid Search* untuk Prediksi Risiko Penyakit Jantung,” *J. Pendidik. dan Teknol. Indones.*, vol. 5, no. 10, pp. 3074–3084, 2025, doi: <https://doi.org/10.52436/1.jpti.1346>.
- [21] M. Arifin and S. Adiyono, “*Hyperparameter* Tuning in Machine Learning to Predicting Student Academic Achievement,” *Int. J. Artif. Intelligence Res.*, vol. 8, no. 1.1, pp. 1–8, 2024, [Online]. Available: <http://ijair.id>
- [22] Mukhlisin and H. W. Nugroho, “Customer Loyalty Classification Using KNN and *Decision Tree* for Sales Strategy Development,” *Sink. J. dan Penelit. Tek. Inform.*, vol. 9, no. 3, pp. 1159–1166, 2025, doi: <https://doi.org/10.33395/sinkron.v9i3.15110> e-ISSN.
- [23] H. S. Putra, H. Okyprana, and R. K. Sormin, “IMPLEMENTASI ALGORITMA *DECISION TREE* DALAM MENENTUKAN LOKASI PENJUALAN DAGANGAN PADA DINAS PASAR SERBELAWAN SIMALUNGUN,” *SmartEDU J.*, vol. 1, no. 2, pp. 61–69, 2022.