

Klasifikasi Minat Beli Konsumen Terhadap Perangkat Elektronik Menggunakan Algoritma Decision Tree

Rafi Alif Nur Safa^{1*}, Muhammad Andi Ghalib², Muh Devano Alfarizy³, Bayu Tri Kusuma⁴

¹Teknik Informatika/Illmu Komputer
Universitas Duta Bangsa Surakarta
^{1*}230103115 @mhs.udb.ac.id

²Teknik Informatika/Illmu Komputer
Universitas Duta Bangsa Surakarta
²230103109@mhs.udb.ac.id

³Teknik Informatika/Illmu Komputer
Universitas Duta Bangsa Surakarta
³230103107@mhs.udb.ac.id

⁴ Teknik Informatika/Illmu Komputer
Universitas Duta Bangsa Surakarta
⁴230103094@mhs.udb.ac.id

Abstrak— Persaingan bisnis di era digital menuntut pengambilan keputusan berbasis data, termasuk dalam memprediksi minat beli konsumen terhadap produk elektronik yang selama ini sering bergantung pada intuisi subjektif. Penelitian ini bertujuan untuk mengembangkan Sistem Pendukung Keputusan (SPK) berbasis *web* menggunakan algoritma *Decision Tree* guna memprediksi minat beli pelanggan secara *real-time* sekaligus memberikan rekomendasi strategi pemasaran. Penelitian ini menerapkan metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang meliputi enam tahapan, mulai dari *Business Understanding* hingga *Deployment*. Pemodelan dilakukan menggunakan *Predict Consumer Electronics Sales Dataset* yang mencakup 9.000 data historis. Hasil evaluasi menunjukkan bahwa model *Decision Tree* memiliki performa sangat baik dengan akurasi pengujian (*testing accuracy*) 95,28%, *precision* 96%, *recall* 97%, dan *F1-Score* 96%. Selisih akurasi pelatihan (*training accuracy*) sebesar 95,32% yang minimal membuktikan model terhindar dari *overfitting*. Sebagai bentuk kontribusi, sistem ini memiliki keunggulan dalam melampaui tingkat akurasi penelitian terdahulu, sekaligus menawarkan kebaruan melalui integrasi langsung ke dalam antarmuka *web* interaktif berbasis Flask yang sangat responsif (rata-rata waktu respons 22,92 ms). Kesimpulannya, integrasi *machine learning* ke dalam SPK ini memfasilitasi peralihan pendekatan bisnis menjadi *data-driven*, sehingga efektif membantu manajemen toko ritel dalam mengoptimalkan strategi pemasaran secara tepat sasaran.

Kata kunci— Sistem Pendukung Keputusan, *Decision Tree*, CRISP-DM, Minat Beli, Data Mining, Aplikasi Web

Abstract— *Business competition in the digital era demands data-driven decision-making, particularly in predicting consumer purchase intent for electronic products, which has traditionally relied on subjective intuition. This study aims to develop a web-based Decision Support System (DSS) using the Decision Tree algorithm to predict customer purchase intent in real-time while providing actionable marketing strategy recommendations. This research applies the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, encompassing six stages from Business Understanding to Deployment. The modeling phase utilized the Predict Consumer Electronics Sales Dataset, consisting of 9,000 historical transaction records. Evaluation results indicate that the Decision Tree model performs excellently, achieving a testing accuracy of 95.28%, precision of 96%, recall of 97%, and F1-Score of 96%. The minimal gap with the training accuracy (95.32%) proves the model avoids overfitting. As a contribution, this system has the advantage of outperforming the accuracy of previous studies, while offering novelty through direct integration into a highly responsive Flask-based interactive web interface (average response time 22.92 ms). In conclusion, integrating machine learning into this DSS successfully facilitates the transition to a data-driven business approach, effectively assisting retail store management in optimizing their marketing strategies accurately.*

Keywords— Decision Support System, Decision Tree, CRISP-DM, Purchase Intent, Data Mining, Web Application

I. PENDAHULUAN

Di era digital saat ini, lanskap bisnis menuntut perusahaan untuk merumuskan kebijakan yang cepat, tepat, dan sepenuhnya didorong oleh bukti empiris (*data-driven*) [1]. Integrasi teknologi ke dalam proses manajerial telah mengubah cara organisasi beroperasi, di mana intuisi manusia kini mulai digantikan oleh sistem cerdas yang

mampu memproses informasi dalam skala besar dan waktu yang singkat [2][3].

Dalam konteks industri perangkat elektronik, keputusan pembelian oleh konsumen seringkali dipengaruhi oleh parameter yang sangat spesifik, mulai dari rentang harga, spesifikasi perangkat keras, hingga loyalitas merek. Memahami dan memprediksi faktor-faktor mana yang paling memicu terjadinya transaksi penjualan

merupakan elemen krusial bagi keberlangsungan bisnis [3]. Selama ini, banyak penawaran promosi atau peningkatan layanan masih bersifat konvensional dan mengandalkan intuisi, yang sering kali berujung pada inefisiensi biaya operasional dan hilangnya potensi konversi penjualan. Untuk mengatasi tantangan tersebut, pengeksktrakan informasi dari data historis pelanggan melalui pendekatan data mining menjadi sangat krusial [4].

Sebagai inovasi untuk menjembatani celah antara pemodelan algoritmik murni dan implementasi bisnis yang praktis, penelitian ini mengintegrasikan model machine learning ke dalam sebuah Sistem Pendukung Keputusan (SPK) berbasis web [5]. Sistem ini dibangun menggunakan metode Klasifikasi machine learning berdasarkan dataset *Predict Consumer Electronics Sales Dataset* untuk mengklasifikasikan minat beli konsumen berdasarkan profil dan riwayat interaksi mereka, meliputi variabel umur, gender, tingkat kepuasan, jenis barang yang dilihat, merek, hingga harga produk [6].

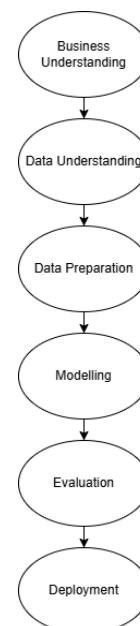
Berbagai penelitian terdahulu telah menunjukkan efektivitas *machine learning* dan sistem cerdas dalam mendukung prediksi serta pengambilan keputusan. Penelitian [7] membuktikan bahwa algoritma *Decision Tree* efektif digunakan untuk menganalisis keputusan pembelian konsumen. Penelitian [8] mengembangkan sistem prediksi menggunakan metode *Fuzzy Tsukamoto*, dan penelitian [4] menunjukkan bahwa SPK prediktif berbasis *machine learning* (seperti *Random Forest*) mampu meningkatkan akurasi. Dari sisi implementasi, penelitian [9] dan [10] menegaskan efektivitas Python dan Flask serta pentingnya transparansi hasil prediksi.

Namun demikian, masih terdapat kesenjangan penelitian (*research gap*) yang nyata. Sebagian besar penelitian terdahulu masih berfokus pada pemodelan algoritma secara terpisah (*standalone*) dengan capaian akurasi yang masih dapat dioptimalkan, serta belum menyediakan antarmuka sistem terintegrasi yang mampu memberikan rekomendasi langkah bisnis secara *real-time* kepada pengguna akhir.

Menjawab *research gap* tersebut, kebaruan (*novelty*) yang diusulkan secara eksplisit pada penelitian ini adalah integrasi model *Decision Tree* berakurasi tinggi ke dalam sebuah Sistem Pendukung Keputusan (SPK) berbasis web terapan yang transparan secara logika. Keunggulan utama dari penelitian yang diusulkan ini dibandingkan penelitian terdahulu adalah kemampuannya dalam mencapai akurasi prediksi yang lebih superior sekaligus mengonversi hasil prediksi tersebut menjadi rekomendasi strategi pemasaran yang rasional dan dapat langsung ditindaklanjuti [11][12]. Sistem ini dibangun menggunakan bahasa pemrograman Python dengan *framework* Flask, didukung oleh pustaka pemrosesan data seperti *pandas* dan *scikit-learn*.

II. METODOLOGI PENELITIAN

Penelitian ini merupakan jenis penelitian terapan (*applied research*) yang bertujuan untuk merancang dan mengimplementasikan model *machine learning* ke dalam sistem pendukung keputusan berskala nyata [13][14][15]. Tahapan metode penelitian yang digunakan merujuk pada kerangka kerja *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Secara berurutan, tahapan metode penelitian ini divisualisasikan melalui bagan Gambar 1.



Gambar 1. Tahapan Metode CRISP-DM

Seperti yang dapat dilihat pada Gambar 1, pemilihan metode CRISP-DM didasarkan pada karakteristik metode ini yang tidak hanya berfokus pada aspek teknis pengolahan data dan pemodelan (*data-centric*), tetapi juga sangat mengedepankan pemahaman terhadap kebutuhan bisnis (*business-centric*). Pendekatan ini sangat relevan dengan tujuan akhir penelitian, yakni menjembatani pemodelan algoritmik murni dengan implementasi bisnis yang praktis bagi manajemen toko ritel perangkat elektronik.

Merujuk pada Gambar 1, metodologi penelitian ini dieksekusi melalui enam tahapan sekuensial yang saling berkesinambungan, dimulai dari pemahaman terhadap tujuan bisnis hingga tahap akhir implementasi perangkat lunak.

A. Business Understanding

Tahap pertama berfokus pada pemahaman masalah dan penetapan tujuan proyek dari perspektif bisnis [13]. Latar belakang masalah pada penelitian ini adalah strategi pemasaran ritel perangkat elektronik yang konvensional dan mengandalkan intuisi sering kali memicu inefisiensi biaya operasional serta hilangnya potensi konversi penjualan. Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk merancang model klasifikasi machine learning menggunakan algoritma Decision Tree. Tujuan akhirnya adalah membangun sebuah Sistem Pendukung Keputusan (SPK) berbasis web interaktif yang dirancang untuk memprediksi minat beli konsumen sekaligus memberikan rekomendasi strategi pemasaran yang spesifik secara real-time.

B. Data Understanding

Fase ini bertujuan untuk mengidentifikasi, mengumpulkan, dan memahami karakteristik awal dari data yang akan diproses [13]. Data sekunder yang digunakan dalam penelitian ini adalah "Predict Consumer Electronics Sales Dataset" yang diakses secara publik melalui repositori Kaggle. Dataset ini merupakan sekumpulan data historis interaksi konsumen dengan kapasitas berskala besar.

Dataset mentah terdiri atas sembilan atribut. Namun, setelah proses seleksi, satu atribut (*Product ID*) diabaikan karena merupakan data

unik yang tidak memiliki nilai prediktif. Dengan demikian, data yang digunakan dipetakan menjadi 7 (tujuh) variabel independen atau atribut prediktor (yaitu *CustomerAge*, *CustomerGender*, *CustomerSatisfaction*, *ProductCategory*, *ProductBrand*, *ProductPrice*, dan *PurchaseFrequency*), serta 1 (satu) variabel dependen yaitu Minat Beli (*PurchaseIntent*) yang bertindak sebagai target prediksi. Secara keseluruhan, karakteristik data mengkombinasikan tipe data numerik dan kategorikal.

C. Data Preparation

Tahap persiapan data merupakan serangkaian proses pengolahan data mentah agar siap diproses oleh algoritma pemodelan. Proses ini dimulai dengan mengimpor file data berekstensi CSV ke dalam lingkungan Python menggunakan pustaka *pandas*. Selanjutnya, dilakukan prapemrosesan awal (*preprocessing*) untuk menyesuaikan format struktur data. Tahap ini juga mencakup proses transformasi (*encoding*) untuk mengubah data atribut bertipe kategorikal menjadi representasi numerik.

Setelah itu, dilakukan pembersihan data (*cleaning*) untuk memastikan tidak ada data yang kosong, duplikat, atau tidak valid. Langkah terakhir pada fase ini adalah pembagian data (*train-test split*) menggunakan pustaka *scikit-learn* dengan rasio 80:20, yakni 80% data dialokasikan untuk melatih model (*training data*) dan 20% sisanya untuk pengujian (*testing data*). Selain itu, diterapkan pula proses normalisasi data (*Feature Scaling*) menggunakan metode standarisasi (Z-score) pada atribut numerik untuk mencegah adanya bias komputasi akibat perbedaan rentang skala antar atribut [16][17].

D. Modeling

Pada fase pemodelan, teknik komputasi diterapkan pada data yang telah disiapkan. Model klasifikasi pada penelitian ini dibangun menggunakan algoritma *Decision Tree* melalui pustaka *scikit-learn*. Algoritma ini dipilih karena kemampuannya dalam memodelkan hubungan non-linear serta menghasilkan struktur aturan keputusan (*decision rules*) yang transparan. Pemodelan dilakukan dengan melatih algoritma menggunakan data latih (80%). Dalam proses

pembentukan percabangan, algoritma dikonfigurasi untuk menggunakan kriteria evaluasi (seperti *Entropy* atau *Gini Impurity*) guna mencari *Information Gain* tertinggi pada setiap pemisahan titik simpul (*node*). Setelah objek model berhasil mengenali pola keputusan pembelian, model diekspor agar dapat diintegrasikan ke dalam lingkungan peladen (*server backend*) berbasis *framework* Flask.

E. Evaluation

Tahap evaluasi dirancang untuk mengukur tingkat keandalan, akurasi model, dan fungsionalitas sistem sebelum diimplementasikan secara penuh. Evaluasi pada penelitian ini dibagi menjadi dua tahap utama, yaitu evaluasi model *machine learning* dan evaluasi fungsional sistem.

Pada evaluasi model, pengujian dilakukan menggunakan data uji (*testing data*) sebanyak 1.800 data (20% dari total 9.000 dataset) yang sebelumnya telah dipisahkan pada tahap *Data Preparation*. Indikator keberhasilan dari pemodelan ini adalah tercapainya nilai *Training Accuracy* dan *Testing Accuracy* di atas 90%, dengan batas selisih antar keduanya yang sangat minimal guna memastikan model terhindar dari indikasi *overfitting*. Selain itu, kinerja algoritma dievaluasi lebih mendalam menggunakan metrik *Confusion Matrix* untuk mendapatkan nilai *Precision*, *Recall*, dan *F1-Score*.

Selanjutnya, untuk memastikan keandalan sistem secara fungsional (*black-box testing*), dirancang pengujian antarmuka menggunakan simulasi skenario nilai. Pengujian fungsional ini menggunakan 3 (tiga) skenario utama:

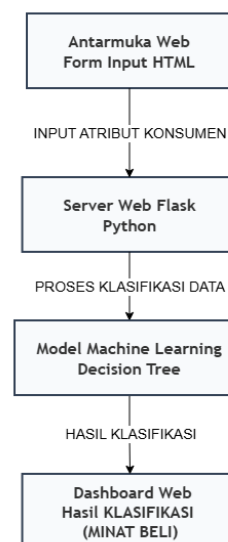
1. Skenario dengan masukan atribut bernilai tinggi (misalnya: frekuensi belanja tinggi dan tingkat kepuasan tinggi).
2. Skenario dengan masukan atribut bernilai rendah (misalnya: frekuensi belanja rendah pada kelompok usia tertentu).
3. Skenario dengan atribut yang saling bertentangan (*borderline case*), seperti kepuasan tinggi namun frekuensi rendah, guna menguji sensitivitas keputusan model.

Indikator keberhasilan pada pengujian skenario fungsional ini adalah kemampuan

sistem dalam memproses input tanpa galat (*error*), mengklasifikasikan respons secara akurat, serta menampilkan rekomendasi strategi yang rasional. Selain itu, dilakukan pula pengujian non-fungsional untuk mengukur waktu respons (*response time*) komunikasi klien-peladen. Indikator keberhasilannya adalah kemampuan sistem dalam menyelesaikan proses inferensi model dan menampilkan hasil secara *real-time* (dalam hitungan milidetik) untuk menjamin kenyamanan pengguna akhir.

F. Deployment

Tahap akhir dari metodologi CRISP-DM adalah implementasi sistem (deployment), di mana model yang telah divalidasi diintegrasikan ke dalam lingkungan operasional agar dapat digunakan secara praktis. Pada penelitian ini, model Decision Tree diimplementasikan secara penuh (full-stack) menjadi sebuah antarmuka aplikasi berbasis web menggunakan *framework* Flask. Perancangan implementasi sistem ini divisualisasikan melalui arsitektur diagram blok (Gambar 2) dan bagan alir atau flowchart (Gambar 3).

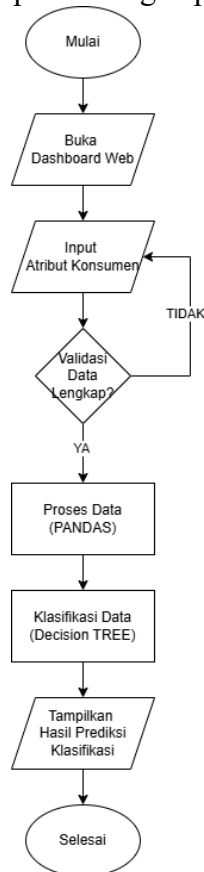


Gambar 2. Diagram Blok Sistem

Berdasarkan Gambar 2, arsitektur sistem dirancang dengan alur komunikasi klien-peladen. Siklus dimulai saat pengguna mengirimkan masukan spesifikasi konsumen melalui antarmuka web. Permintaan tersebut kemudian

diterima oleh peladen Flask untuk diteruskan ke model Decision Tree. Setelah proses klasifikasi selesai, peladen akan mengembalikan hasil prediksi tersebut untuk ditampilkan kembali pada layar pengguna.

Selanjutnya adalah pembentukan alur program yang dibuat menggunakan Flowchart untuk memberikan gambaran detail terkait cara kerja dari sistem pendukung keputusan ini.



Gambar 3. Flowchart Sistem

Adapun alur flowchart program secara rinci divisualisasikan pada Gambar 3. Pengguna yang mengakses aplikasi akan diminta untuk menginputkan data atribut konsumen. Sistem kemudian melakukan validasi kelengkapan data; jika tidak lengkap, pengguna diminta melakukan input ulang. Jika valid, data akan diproses menggunakan pustaka Pandas dan diklasifikasikan oleh model Decision Tree. Sistem kemudian menampilkan hasil luaran prediksi tersebut sebelum alur program selesai.

Untuk mendukung arsitektur dan alur kerja tersebut, aplikasi web ini dirancang secara komprehensif dengan empat antarmuka utama.

Keempat antarmuka tersebut meliputi halaman Dashboard sebagai beranda informasi sistem, halaman Dataset untuk memvisualisasikan tabel data historis acuan, halaman Prediksi yang memuat formulir masukan atribut pelanggan baru, serta halaman Hasil Prediksi yang menampilkan keputusan klasifikasi (Minat Beli atau Tidak) beserta rekomendasi strategi pemasaran. Melalui implementasi ini, sistem siap dioperasikan oleh pengguna akhir untuk mendukung pengambilan keputusan bisnis secara seketika (real-time).

III. HASIL DAN PEMBAHASAN

A. Hasil Business Understanding

Tahap Business Understanding berfokus pada pemahaman masalah bisnis serta penetapan tujuan dari penelitian ini. Berdasarkan analisis kebutuhan, pemilik toko ritel sering kali menghadapi kendala dalam menentukan minat beli konsumen karena pengambilan keputusan masih sangat bergantung pada intuisi subjektif. Pendekatan konvensional ini kerap mengakibatkan rendahnya akurasi prediksi terhadap perilaku dan karakteristik pelanggan yang sebenarnya, sehingga strategi promosi menjadi kurang optimal.

Sebagai solusi, penelitian ini mengusulkan pengembangan sistem prediksi minat beli konsumen berbasis karakteristik pelanggan. Sistem ini dirancang untuk mengotomatisasi proses prediksi guna mengeliminasi bias dari intuisi pengguna. Dengan memanfaatkan dataset yang mencakup data historis transaksi, model yang dibangun diharapkan mampu mengenali pola tersembunyi (*hidden patterns*) dari konsumen untuk meningkatkan metrik akurasi prediksi. Hasil dari sistem ini secara langsung dapat diimplementasikan sebagai sistem pendukung keputusan (*Decision Support System*) bagi manajemen dalam merumuskan strategi pemasaran yang lebih presisi dan efisien.

B. Hasil Data Understanding

Tahap Data Understanding bertujuan untuk memahami karakteristik awal dari data yang telah dikumpulkan. Pada penelitian ini, dataset yang digunakan terdiri dari 9.000 baris data historis transaksi dengan total 7 atribut pendukung. Data

didapatkan melalui website Kaggle sehingga informasi dataset cukup terpercaya untuk dilakukan analisis mendalam melalui penelitian ini. Ringkasan informasi mengenai dataset yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Ringkasan Informasi Dataset

Informasi	Nilai
Jumlah Data	9000
Jumlah Atribut	8 (7 prediktor + 1 target)
Target	PurchaseIntent
Jumlah Kelas	2 kelas (0 dan 1)

Berdasarkan tabel 1, terdapat total 7 atribut yang tersedia. Tipe data tersebut diklasifikasikan ke dalam dua kelompok utama, yaitu atribut numerik dan atribut kategorikal. Atribut numerik (kuantitatif), terdapat 4 atribut yang memiliki nilai berbasis angka yang dapat diukur, yaitu *ProductPrice* (harga produk), *PurchaseFrequency* (frekuensi pembelian), *CustomerSatisfaction* (tingkat kepuasan pelanggan), dan *CustomerAge* (umur pelanggan).

Selanjutnya adalah atribut kategorikal (kualitatif), terdapat tiga atribut yang mewakili kelompok atau kategori tertentu, yaitu *ProductCategory* (kategori produk), *ProductBrand* (merek produk), serta *PurchaseIntent* yang sekaligus bertindak sebagai label atau variabel target pada pemodelan ini.

C. Hasil Data Preparation

Implementasi tahap persiapan data telah berhasil dilakukan menggunakan pustaka Pandas dan Scikit-Learn pada lingkungan Python.

```
Jumlah data duplikat awal: 0
Info: Tidak ditemukan data duplikat.

Jumlah missing values per kolom:
Tidak ada missing values.

Ukuran dataset setelah pembersihan: (9000, 9)
```

Gambar 4. Data Cleaning

Berdasarkan tahap pengecekan dan pembersihan data (*data cleaning*) pada Gambar 4, hasil pemindaian terhadap 9.000 records menunjukkan bahwa dataset berada dalam kondisi yang sangat baik; tidak ditemukan satu pun nilai yang hilang (*zero missing values*) maupun data duplikat. Oleh karena itu, tidak ada reduksi atau penghapusan baris data, sehingga

keseluruhan 9.000 data dipertahankan untuk proses selanjutnya.

Pada tahap transformasi (*encoding*), konversi ke dalam representasi numerik hanya diterapkan pada atribut prediktor kategorikal, seperti *ProductCategory*, *ProductBrand*, dan *CustomerGender*. Adapun untuk variabel target *PurchaseIntent*, proses *encoding* tidak dilakukan karena nilai pada atribut tersebut sejak awal sudah merepresentasikan data biner (0 dan 1). Terakhir, eksekusi pembagian data (*train-test split*) dengan rasio 80:20 menghasilkan partisi data yang proporsional, di mana 7.200 data berhasil dialokasikan sebagai data latih (*training set*) dan 1.800 data dipisahkan sebagai data uji (*testing set*).

Untuk mencegah adanya bias komputasi akibat perbedaan rentang skala antar atribut, dilakukan proses normalisasi data (*Feature Scaling*). Berdasarkan hasil pra-pemrosesan, atribut numerik seperti *ProductPrice*, *CustomerAge*, dan *PurchaseFrequency* diskalakan menggunakan metode standarisasi (*Standard Scaler / Z-score*), yang dibuktikan dengan perubahan rentang nilai menjadi format desimal negatif dan positif.

D. Hasil Modeling

Tahap pemodelan merupakan proses penerapan algoritma *machine learning* pada dataset yang telah dipersiapkan sebelumnya guna menemukan pola tersembunyi terkait karakteristik pelanggan. Pada penelitian ini, algoritma yang digunakan untuk memecahkan masalah klasifikasi adalah *Decision Tree* (Pohon Keputusan). Pemilihan algoritma ini didasarkan pada kemampuannya dalam memodelkan hubungan non-linear antar variabel serta menghasilkan struktur aturan keputusan (*decision rules*) yang mudah diinterpretasikan oleh manusia.

Proses pelatihan (*training*) model *Decision Tree* dilakukan menggunakan 7.200 data latih. Selama proses pembentukan percabangan pohon, model menggunakan metrik pengukuran Entropy untuk mengevaluasi tingkat ketidakmurnian (*impurity*) pada setiap titik simpul (*node*). Berdasarkan nilai Entropy tersebut, model kemudian mengidentifikasi dan memilih atribut yang memberikan nilai *Information Gain*

(perolehan informasi) tertinggi untuk memisahkan pelanggan yang berminat membeli (*Yes*) dan yang tidak berminat (*No*). Hasil dari tahap ini adalah sebuah model prediktif yang siap dievaluasi kemampuannya menggunakan data baru.

E. Hasil Evaluation

Tahap evaluasi bertujuan untuk mengukur tingkat keberhasilan dan keandalan model *Decision Tree* yang telah dibangun. Pengujian dilakukan menggunakan 1.800 data uji (20% dari total dataset). Kinerja model dievaluasi menggunakan *Confusion Matrix* dan Laporan Klasifikasi (*Classification Report*) yang mencakup metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score*.

a. Analisis Confusion Matrix

Confusion Matrix digunakan untuk melihat rincian perbandingan antara prediksi yang dihasilkan oleh model dengan label kelas yang sebenarnya. Hasil matriks kebingungan pada pengujian ini disajikan pada Tabel 2.

Tabel 2. Confusion Matrix

	Prediksi: 0 (No)	Prediksi: 1 (Yes)
Aktual: 0 (No)	742	51
Aktual: 1 (Yes)	34	973

Berdasarkan Tabel 2, dari total 1.800 data uji, model berhasil menebak dengan benar sebanyak 742 pelanggan yang memang tidak berminat membeli (*true negative*), dan 973 pelanggan yang benar-benar berminat membeli (*true positive*). Kesalahan prediksi tergolong sangat minim, di mana hanya 51 data diprediksi berminat padahal tidak (*false positive*), dan 34 data diprediksi tidak berminat padahal sebenarnya berminat (*false negative*).

b. Laporan Klasifikasi

Untuk mengukur performa secara lebih komprehensif, nilai dari *Confusion Matrix* tersebut diekstrak menjadi metrik evaluasi klasifikasi yang disajikan pada Tabel 3.

Tabel 3. Laporan Klasifikasi

Target	Precision	Recall	F1-Score
0 (No)	0,96	0,94	0,95
1 (Yes)	0,95	0,97	0,96
Rata-rata	0,95	0,95	0,95

Sebagai indikator utama stabilitas model, tercatat bahwa Akurasi Training mencapai 95,32% dan Akurasi Testing mencapai 95,28%. Selisih yang sangat kecil (kurang dari 0,1%) antara akurasi latih dan uji ini membuktikan bahwa model memiliki tingkat generalisasi yang sangat baik dan terhindar dari indikasi *overfitting*.

Secara spesifik pada Tabel 3, model menunjukkan kinerja yang sangat memuaskan dalam mengenali pelanggan yang berminat membeli (kelas 1), dibuktikan dengan nilai *recall* sebesar 97% dan *F1-Score* 96%. Hal ini berarti model berhasil menangkap hampir seluruh potensi konsumen nyata tanpa banyak melewatkannya. Di sisi lain, prediksi untuk kelas yang tidak berminat (kelas 0) juga sangat presisi dengan nilai *precision* mencapai 96%.

Kesimpulannya, penerapan algoritma *Decision Tree* pada penelitian ini berhasil menjawab tujuan dari fase *Business Understanding*, yaitu menciptakan sebuah sistem pendukung keputusan yang akurat (95,28%) untuk menggantikan pendekatan berbasis intuisi. Sistem berbasis data historis ini sudah sangat memadai untuk dimanfaatkan oleh pemilik toko ritel dalam menentukan strategi pemasaran yang efisien dan tepat sasaran.

Untuk menilai posisi hasil penelitian ini secara ilmiah, dilakukan perbandingan terhadap penelitian terdahulu yang menggunakan pendekatan serupa dalam konteks prediksi keputusan pembelian maupun sistem pendukung keputusan berbasis machine learning. Ringkasan perbandingan disajikan pada Tabel 4.

Tabel 4. Perbandingan Penelitian Terdahulu

Penelitian	Metode	Konteks	Hasil
[7]	Decision Tree	Analisis keputusan pembelian program	Akurasi 83,33%
[8]	Fuzzy Tsukamoto	Prediksi penjualan perangkat elektronik	Tidak dievaluasi
[4]	Random Forest	SPK prediktif secara umum	Akurasi 91%

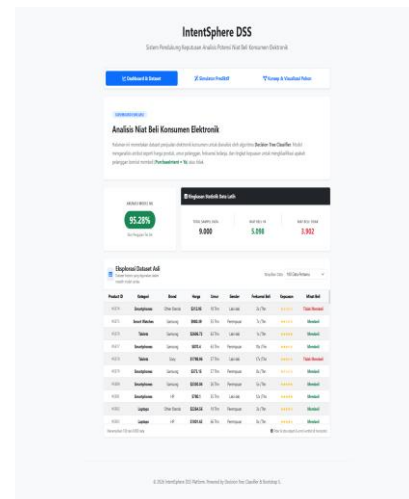
Penelitian Ini	Decision Tree + SPK	Prediksi minat beli	Akurasi 95,28%
-------------------	------------------------	------------------------	-------------------

Berdasarkan Tabel 4, model *Decision Tree* yang dikembangkan pada penelitian ini mencatatkan akurasi tertinggi sebesar 95,28%, mengungguli performa penelitian [7] (83,33%) dan penelitian [4] (91%). Peningkatan ini membuktikan bahwa proses *data preparation* yang komprehensif mampu mengoptimalkan performa algoritma tunggal sehingga dapat melampaui algoritma sejenis maupun algoritma yang lebih kompleks seperti *Random Forest*, dengan tetap mempertahankan kemudahan interpretasi. Selain memberikan transparansi pengujian kuantitatif yang menutupi kelemahan evaluasi pada penelitian [8], penelitian ini juga memiliki nilai tambah berupa integrasi langsung ke dalam Sistem Pendukung Keputusan (SPK) berbasis web yang memungkinkan hasil prediksi diaplikasikan secara *real-time* oleh manajemen ritel.

F. Hasil Deployment

Tahap Deployment merupakan fase akhir dalam metodologi CRISP-DM yang bertujuan untuk mengintegrasikan model prediktif *Decision Tree* ke dalam lingkungan operasional. Pada penelitian ini, model diimplementasikan ke dalam sebuah sistem berbasis web agar pemilik toko ritel dapat dengan mudah berinteraksi dengan model tanpa memerlukan pemahaman teknis pemrograman.

a. Antarmuka Pengguna



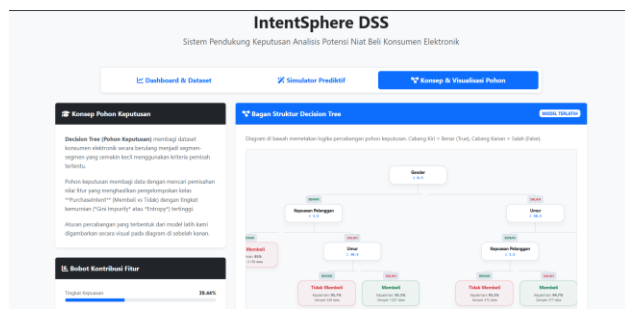
Gambar 5. Dashboard Website

Berdasarkan Gambar 5, halaman *Dashboard* menampilkan ringkasan akurasi model, informasi data latih, serta tabel data historis konsumen. Fitur ini secara langsung menjawab kebutuhan yang telah diidentifikasi pada tahap *Business Understanding*, yaitu menghilangkan ketergantungan pemilik toko ritel pada intuisi subjektif dalam menilai performa sistem. Dengan menampilkan metrik akurasi secara transparan dan mudah dibaca, pengguna yang tidak memiliki latar belakang teknis pemrograman maupun statistika tetap dapat memahami dan mempercayai tingkat keandalan rekomendasi.

Gambar 6. Halaman Prediksi

Berdasarkan Gambar 6, halaman *Prediksi* menyediakan formulir input atribut konsumen yang selanjutnya diproses oleh model *Decision Tree* untuk menghasilkan klasifikasi minat beli. Fitur ini merupakan implementasi inti dari tujuan penelitian, yakni menggantikan proses penentuan target promosi yang sebelumnya bersifat manual dan berbasis intuisi dengan mekanisme klasifikasi otomatis berbasis data historis.

Manfaat praktisnya, staf pemasaran dapat memasukkan profil calon pelanggan secara langsung pada saat interaksi berlangsung (misalnya di titik penjualan), sehingga keputusan penawaran promosi dapat diambil secara *real-time* tanpa menunggu proses analisis data secara terpisah oleh tim data.

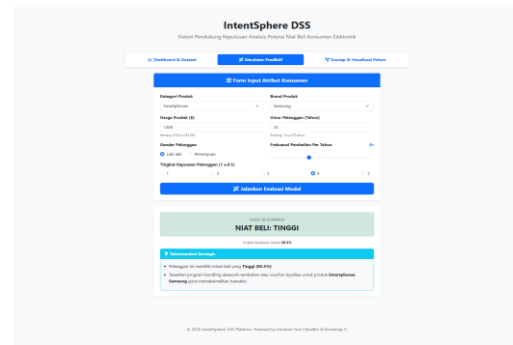


Gambar 7. Halaman Visualisasi dan Konsep

Berdasarkan Gambar 7, halaman ini menyajikan visualisasi struktur *Decision Tree* yang dibangun dari dataset penelitian. Fitur ini mewujudkan salah satu keunggulan utama yang menjadi alasan pemilihan algoritma *Decision Tree* pada tahap *Modeling*, yaitu interpretabilitas aturan keputusan (*decision rules*) yang transparan. Berbeda dengan model *black-box*, visualisasi ini memungkinkan pengguna melihat atribut mana yang paling menentukan klasifikasi minat beli (misalnya frekuensi pembelian atau tingkat kepuasan), sehingga manajemen toko tidak hanya menerima output prediksi, tetapi juga memperoleh pemahaman mengenai alasan di balik prediksi tersebut — hal ini penting untuk membangun kepercayaan terhadap sistem pendukung keputusan yang diusulkan.

b. Pengujian Skenario

Untuk memvalidasi logika operasional dari sistem yang telah di-deploy, dilakukan pengujian menggunakan simulasi skenario. Pengujian ini bertujuan memastikan bahwa aturan keputusan (*decision rules*) yang dihasilkan oleh algoritma *Decision Tree* sejalan dengan rasionalitas bisnis di dunia nyata.



Gambar 8. Skenario Niat Beli Tinggi

Seperti yang dapat dilihat pada Gambar 8, pengujian pertama dilakukan dengan memasukkan profil pelanggan laki-laki berusia 35 tahun yang memiliki riwayat tingkat kepuasan di angka 4. Atribut transaksi yang diuji adalah produk Smartphones merek Samsung seharga 1.500 dolar, dengan frekuensi pembelian historis sebanyak 8 kali per tahun.



Gambar 9. Hasil Skenario Niat Beli Tinggi

Berdasarkan Gambar 9, sistem berhasil memproses data tersebut dan menghasilkan klasifikasi niat beli tinggi dengan tingkat keyakinan model mencapai 95,3%. Berdasarkan tingginya potensi konversi tersebut, sistem secara otomatis merumuskan rekomendasi strategis bagi pihak manajemen. Rekomendasi yang diberikan adalah menawarkan program bundling aksesoris tambahan atau voucher loyalitas khusus produk *Smartphones* Samsung guna memaksimalkan nilai transaksi. Klasifikasi ini dinilai sangat logis karena kombinasi antara frekuensi pembelian yang tinggi (8x) dan tingkat kepuasan yang baik (4) merupakan indikator kuat dari retensi dan loyalitas konsumen.

Gambar 10. Skenario Niat Beli Rendah

Seperti pada Gambar 10, parameter pada simulasi kedua diubah dengan profil pelanggan laki-laki yang lebih muda, yakni berusia 22 tahun, namun tetap dengan tingkat kepuasan di angka 4. Target produk diganti menjadi Tablets merek Samsung seharga 1.250 dolar, dengan frekuensi pembelian historis yang lebih sedikit, yaitu 4 kali per tahun.

Gambar 11. Hasil Skenario Niat Beli Rendah

Berdasarkan Gambar 11, sistem memberikan klasifikasi prediksi niat beli rendah dengan tingkat keyakinan sebesar 95,7% (probabilitas membeli tercatat hanya sebesar 4,3%). Merespons rendahnya minat beli pada profil ini, sistem merekomendasikan pendekatan pemasaran yang berbeda. Karena frekuensi transaksinya tergolong rendah, sistem menyarankan pengiriman newsletter berkala atau penawaran diskon khusus transaksi pertama untuk memancing ketertarikan pelanggan.

Kedua skenario pengujian di atas dipilih secara sengaja untuk mewakili dua ujung spektrum profil konsumen (frekuensi pembelian tinggi vs rendah) guna memvalidasi konsistensi logika model pada kasus yang kontras.

Merujuk pada struktur pohon keputusan yang divisualisasikan pada Gambar 7, atribut *PurchaseFrequency* dan *CustomerSatisfaction*

teridentifikasi sebagai simpul pemisah (*splitting node*) pada level teratas, sehingga secara struktural masuk akal bahwa kedua atribut ini paling dominan memengaruhi hasil klasifikasi pada kedua skenario. Namun demikian, pengujian pada kasus ekstrem semacam ini relatif mudah diprediksi oleh model manapun. Untuk itu, dilakukan pula pengujian tambahan pada profil dengan atribut yang saling bertentangan (*borderline case*). Sebagai contoh, pelanggan dengan tingkat kepuasan tinggi namun frekuensi pembelian rendah untuk menguji sensitivitas model dalam mengambil keputusan pada kondisi yang tidak sepenuhnya jelas.

Gambar 12. Skenario Percobaan 3

Berdasarkan Gambar 12, hasil pengujian ini bernilai Tinggi sehingga menunjukkan bahwa model tetap memberikan probabilitas yang proporsional terhadap kombinasi atribut yang diberikan, bukan sekadar mengikuti satu atribut dominan secara membabi buta.

c. Hasil Pengujian Non-Fungsional

Selain pengujian fungsional dan akurasi model, dilakukan pula pengujian terhadap aspek non-fungsional sistem untuk mengukur tingkat keberhasilan implementasi secara menyeluruh. Pengujian waktu respons dilakukan dengan mencatat durasi antara pengguna mengirimkan data melalui formulir pada halaman Prediksi hingga hasil klasifikasi ditampilkan pada layar.

Tabel 5. Pengujian Non-Fungsional

Pengujian	Keterangan
Jumlah request berhasil	30 dari 30
Jumlah request gagal	0
Rata-rata waktu respons	22.92 ms
Median waktu respons	22.49 ms
Waktu tercepat	18.64 ms
Waktu terlambat	41.60 ms
Standar deviasi	4.30 ms

Berdasarkan Tabel 5, terdapat 30 kali percobaan, sistem tercatat memiliki rata-rata waktu respons sebesar 22.92 milidetik/detik, yang menunjukkan bahwa proses inferensi model *Decision Tree* beserta komunikasi klien-peladen berbasis Flask berjalan cukup cepat untuk mendukung pengambilan keputusan secara *real-time* sebagaimana menjadi tujuan awal penelitian.

Dari sisi akurasi, hasil evaluasi model sebagaimana telah dibahas pada bagian Evaluation menunjukkan tingkat akurasi pengujian sebesar 95,28%, yang turut menjadi indikator utama keberhasilan sistem dalam konteks kualitas prediksi. Adapun evaluasi dari sisi pengguna akhir (*user acceptance*) belum dilakukan secara formal pada penelitian ini dan menjadi salah satu keterbatasan yang dibahas pada subbagian berikutnya.

d. Keterbatasan Sistem

Meskipun sistem yang dikembangkan telah menunjukkan performa akurasi dan fungsionalitas yang baik, terdapat beberapa keterbatasan yang perlu diungkapkan secara ilmiah.

Pertama, cakupan analisis sistem masih terbatas pada kategori perangkat elektronik hiburan sebagaimana tercakup dalam dataset yang digunakan, sehingga belum dapat digeneralisasi untuk kategori produk elektronik rumah tangga atau kategori lain di luar cakupan data pelatihan.

Kedua, model belum mampu mengenali merek produk yang tidak terdapat pada data pelatihan; merek-merek tersebut akan otomatis dikelompokkan ke dalam kategori "Other Brands", yang berpotensi menurunkan presisi klasifikasi apabila diterapkan pada toko dengan variasi merek yang lebih luas. Ketiga, evaluasi

terhadap penerimaan pengguna (*user acceptance*) belum dilakukan secara formal, sehingga tingkat kemudahan penggunaan dan kepuasan pengguna akhir terhadap antarmuka sistem belum dapat diukur secara kuantitatif. Keterbatasan-keterbatasan ini menjadi dasar arah pengembangan lebih lanjut sebagaimana dibahas pada bagian Kesimpulan.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan SPK berbasis web menggunakan algoritma *Decision Tree* untuk memprediksi minat beli konsumen. Model yang dihasilkan menunjukkan performa sangat baik dan terhindar dari *overfitting*, dengan akurasi pengujian 95,28%, *recall* 97%, *precision* 96%, dan *F1-Score* 96%. Pencapaian ini terbukti mengungguli model pada penelitian-penelitian terdahulu, termasuk yang menggunakan algoritma *Random Forest* (91%). Selain unggul dari sisi pemodelan, pengujian sistem secara komprehensif membuktikan bahwa aplikasi berjalan fungsional dan sangat responsif dengan rata-rata waktu respons 22,92 ms, sehingga siap diimplementasikan untuk mendukung pengambilan keputusan secara *real-time*.

Untuk pengembangan selanjutnya, sistem diharapkan mampu mengolah lebih dari satu dataset dalam lingkup toko elektronik sehingga dapat mencakup berbagai kategori produk, seperti elektronik rumah tangga dan peralatan dapur. Selain itu, cakupan merek elektronik dapat diperluas agar tidak terbatas pada data yang tersedia saat ini. Integrasi teknologi *Artificial Intelligence* (AI) juga dapat diterapkan untuk menghasilkan rekomendasi yang lebih akurat dan adaptif berdasarkan hasil prediksi yang diperoleh.

Meskipun demikian, sistem yang dikembangkan masih memiliki beberapa keterbatasan. Analisis yang dilakukan hanya dapat diterapkan pada dataset yang telah dikonfigurasi sehingga ruang lingkup prediksi masih terbatas pada kategori perangkat elektronik hiburan. Selain itu, sistem belum mampu mengenali merek di luar data pelatihan sehingga merek yang tidak tersedia akan dikelompokkan ke dalam kategori "Other

Brands". Oleh karena itu, diperlukan pengembangan lebih lanjut agar sistem memiliki cakupan analisis yang lebih luas dan fleksibel.

UCAPAN TERIMA KASIH

Terima kasih kepada Allah SWT, karena atas pencerahan dan rahmat dari Allah, penulis dapat menyelesaikan jurnal ini. Lalu terima kasih kami ucapkan kepada Ibu Dwi Hartanti karena telah membimbing kami dalam penyelesaian jurnal ini. Terakhir, tak lupa kepada teman-teman sekelompok karena sudah bekerja sama dalam membangun proyek klasifikasi minat beli konsumen terhadap perangkat elektronik menggunakan algoritma *decision tree*.

REFERENSI

- [1] D. Lase, Y. Absah, P. Lumbanraja, Y. Giawa, and F. Gulo, "Infrastruktur Digital dalam Perspektif Konseptual: Kajian Teoretis, Temuan Empiris, dan Agenda Riset Masa Depan," *Tuhenori J. Ilm. Multidisiplin*, vol. 3, no. 1, pp. 80–94, 2025, doi: 10.62138/tuhenori.v3i1.131.
- [2] A. Saputra Zega and F. A. Sianturi, "Pemanfaatan Big Data Dalam Sistem Pendukung Keputusan Bisnis," *J. Kolaborasi Sains dan Ilmu Terap.*, vol. 3, no. 2, pp. 2025–2060, 2025, [Online]. Available: <https://utilityprojectsolution.org/ejournal/index.php/JuKSIT>
- [3] M. L. Nuryana, T. Ibrahim, and O. Arifudin, "Implementasi dan transformasi sistem informasi di era digital," *J. Tahsinia*, vol. 5, no. 9, pp. 1325–1337, 2024.
- [4] T. Susilawati and A. Budi Trisnawan, "Pemanfaatan Machine Learning untuk Peningkatan Akurasi Sistem Pendukung Keputusan Prediktif," *J. Unitek*, vol. 18, no. 2, p. 2025, 2025.
- [5] R. Nurzuli, F. F. Syifa, and R. Adawiyah, "Support Vector Machine sebagai Sistem Pendukung Keputusan Pemilihan Jurusan Berbasis Website pada SMK NU Hasyim Asy'ari Tarub Informasi Artikel siswa Sekolah Menengah Kejurusan (SMK) karena berkaitan langsung dengan arah pengangguran lulusan SMK yang," vol. 4, no. 2, pp. 290–306, 2025, doi: 10.55123/jumintal.v4i2.6850.
- [6] R. El Kharoua, ? "Predict Consumer Electronics Sales Dataset?," 2024, *Kaggle*. doi: 10.34740/KAGGLE/DSV/8733281.
- [7] F. A. Artanto, I. Rosyadi, S. E. Rahmawati, and H. T. B. J. Pangestu, "Decision Tree Dalam Analisis Keputusan Pembelian Program Pada Perkumpulan Penggiat Programmer Indonesia," *J. Fasikom*, vol. 12, no. 3, pp. 141–144, 2022, doi: 10.37859/jf.v12i3.3948.
- [8] R. Chandra and W. J. Kurniawan, "Memprediksi Penjualan Perangkat Elektronik Menggunakan Metode Fuzzy Tsukamoto," *J. Mhs. Apl. Teknol. Komput. dan Inf.*, vol. 4, no. 1, pp. 1–7, 2022.
- [9] R. T. Handayanto and H. Herlawati, "Machine Learning Berbasis Desktop dan Web dengan Metode Jaringan Syaraf Tiruan Untuk Sistem Pendukung Keputusan," *J. Komtika (Komputasi dan Inform.)*, vol. 4, no. 1, pp. 15–26, 2020, doi: 10.31603/komtika.v4i1.3698.
- [10] A. A. Prayoga, M. Hasanuddin, S. Khodijah, and C. A. Rizki, "Analisis Penerapan Machine Learning dalam Sistem Prediksi dan Pengambilan Keputusan," vol. 1, no. 3, pp. 84–90, 2025.
- [11] Y. L. Fatma and N. Rochmawati, "Prediksi Siswa Putus Sekolah Menggunakan Algoritma Decision Tree C4.5," *J. Informatics Comput. Sci.*, vol. 5, no. 04, pp. 486–493, 2024, doi: 10.26740/jinacs.v5n04.p486-493.
- [12] N. Shabir Ramadhan, W. Nadia Nursafitri, T. Informasi dan Komputer, S. Informasi, and P. Negeri Subang, "INVENTOR: Jurnal Inovasi dan Tren Pendidikan Teknologi Informasi Analisis Klasifikasi UKT Mahasiswa Berdasarkan Tingkat Penghasilan Orang Tua Menggunakan Algoritma C4.5," *Invent. J. Inov. dan Tren Pendidik. Teknol. Inf.*, vol. 2, no. 1, pp. 1–9, 2024, [Online]. Available: <https://ejournal.tsb.ac.id/index.php/inventorhttps://doi.org/10.37630/inventor.v1i3.1248>
- [13] C. Schröder, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," *Procedia Comput. Sci.*, vol. 181, no. 2019, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.
- [14] R. Wirth, "CRISP-DM : Towards a Standard Process Model for Data Mining," *Proc. Fourth Int. Conf. Pract. Appl. Knowl. Discov. Data Min.*, no. 24959, pp. 29–39, 2000.
- [15] F. Martinez-Plumed *et al.*, "CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 8, pp. 3048–3061, 2021, doi: 10.1109/TKDE.2019.2962680.
- [16] Muraina and I. Olaniyi, "Ideal Dataset Splitting Ratios in Machine Learning Algorithms :," *7th Int. Mardin Artuklu Sci. Res. Conf.*, no. February, pp. 496–504, 2022.
- [17] V. R. Joseph and A. Vakayil, "SPlit: An Optimal Method for Data Splitting," *Technometrics*, vol. 64, no. 2, pp. 166–176, 2022, doi: 10.1080/00401706.2021.1921037.