

Analisa Kepuasan Pelanggan Terhadap Layanan Aplikasi E-Commerce Menggunakan Algoritma C4.5

Eko Setia Budi¹, Abdul Rahman Kadafi¹, Yasdi Kharismawan², Randi Fadillah², Desy Sasqia Putri^{2*}

¹Fakultas Teknologi Informasi, Program Studi Sistem Informasi, Universitas Nusa Mandiri, Jakarta, Indonesia

²Fakultas Teknik dan Informatika, Program Studi Sistem Informasi, Universitas Bina Sarana Informatika, Jakarta, Indonesia

Email: ¹eko.etb@nusamandiri.ac.id, ²abdul.alk@nusamandiri.ac.id, ³19215037@bsi.ac.id, ⁴19215105@bsi.ac.id,

^{5,*}19215155@bsi.ac.id,

Email Penulis Korespondensi: 19215155@bsi.ac.id

Abstrak—Kepuasan pelanggan merupakan salah satu faktor kunci yang sangat mempengaruhi loyalitas dan keberlanjutan bisnis dari suatu aplikasi e-commerce. Penelitian ini berfokus pada analisis tingkat kepuasan pelanggan terhadap pelayanan e-commerce dengan menggunakan algoritma C4.5. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor kunci yang mempengaruhi kepuasan pelanggan serta memberikan rekomendasi yang dapat membantu perusahaan e-commerce dalam meningkatkan kualitas layanan mereka. Data kepuasan pelanggan dikumpulkan melalui survei yang mencakup berbagai atribut seperti kualitas produk, kecepatan pengiriman, dan responsivitas layanan pelanggan. Hasil analisis menunjukkan bahwa atribut pengiriman produk merupakan faktor yang paling mempengaruhi kepuasan pelanggan, dengan nilai gain sebesar 0,337981562. Model yang dihasilkan memiliki akurasi sebesar 93%, menunjukkan kemampuan yang baik dalam memprediksi kepuasan pelanggan. Temuan ini diharapkan dapat memberikan wawasan praktis bagi perusahaan e-commerce dalam upaya mereka meningkatkan kepuasan dan loyalitas pelanggan.

Kata Kunci: Kepuasan Pelanggan; E-Commerce; Algoritma C4.5; Analisis Data

Abstract—Customer satisfaction is one of the key factors that greatly influences loyalty and business sustainability of an e-commerce application. This research focuses on analyzing the level of customer satisfaction with e-commerce services using the C4.5 algorithm. This research aims to identify key factors that influence customer satisfaction and provide recommendations that can help e-commerce companies improve the quality of their services. Customer satisfaction data is collected through surveys that cover various attributes such as product quality, delivery speed and customer service responsiveness. The analysis results show that the product delivery attribute is the factor that most influences customer satisfaction, with a gain value of 0.337981562. The resulting model has an accuracy of 93%, showing good ability in predicting customer satisfaction. These findings are expected to provide practical insights for e-commerce companies in their efforts to increase customer satisfaction and loyalty.

Keywords: Customer Satisfaction; E-Commerce; C4.5 Algorithm; Data Analysis

1. PENDAHULUAN

Dalam era digital yang berkembang pesat, e-commerce telah menjadi salah satu sektor bisnis yang mengalami pertumbuhan paling signifikan, e-commerce berhasil merevolusi pengalaman berbelanja, menawarkan kenyamanan, pilihan, dan fleksibilitas yang belum pernah ada sebelumnya kepada konsumen[1]. Peningkatan akses internet dan penggunaan perangkat mobile telah mendorong semakin banyak konsumen untuk berbelanja online. Namun, keberhasilan bisnis e-commerce tidak hanya ditentukan oleh keberagaman produk atau harga yang kompetitif, tetapi juga oleh tingkat kepuasan pelanggan. Banyaknya aplikasi e-commerce yang tersedia, tentu semua aplikasi berusaha untuk memberikan kualitas pelayanan yang terbaik mulai dari harga, produk, user interface aplikasi, kemudahan mencari produk, kecepatan dalam merespon masalah, kecepatan konfirmasi produk, kecepatan pengiriman barang dan aspek-aspek lainnya. Kualitas pelayanan merupakan hal penting yang harus diperhatikan dan dimaksimalkan agar mampu bertahan dan dijadikan pilihan oleh pelanggan[2]. Untuk memastikan bahwa aplikasi tersebut bagus atau tidak tentu dibutuhkan umpan balik dari para pelanggan untuk mengukur seberapa puas pelanggan terhadap aplikasi e-commerce tersebut.

Kepuasan pelanggan sendiri merupakan salah satu faktor kunci yang sangat mempengaruhi loyalitas dan keberlanjutan bisnis dari suatu aplikasi e-commerce. Pelanggan yang puas cenderung kembali berbelanja dan merekomendasikan platform e-commerce tersebut kepada orang lain, sementara pelanggan yang tidak puas dapat beralih ke pesaing dan menyebarkan ulasan negatif. Oleh karena itu, memahami dan meningkatkan kepuasan pelanggan menjadi prioritas utama bagi perusahaan e-commerce untuk terus berinovasi agar tidak kehilangan pelanggan. Pengamatan dan pengukuran terhadap kepuasan pelanggan telah menjadi hal yang sangat esensial bagi setiap perusahaan karena cara tersebut dapat memberikan umpan balik dan masukan bagi keperluan pengembang dan penerapan strategi peningkatan kepuasan pelanggan[3].

Peneliti sebelumnya mengungkapkan pada hasil penelitiannya bahwa Harga (keterjangkauan harga, kesesuaian dengan kualitas, daya saing harga, kesesuaian dengan manfaat), Promosi (Promosi konsumen, Promosi dagang, Promosi wiraniaga) dan Kualitas pelayanan (tangible, reliability, responsiveness, assurance, dan empathy) mempunyai pengaruh signifikan yang cukup kuat terhadap kepuasan pelanggan[4]. Peneliti lainnya menggunakan beberapa atribut lain pada penelitiannya seperti kualitas produk, harga, promosi, ketersediaan produk, ketepatan pengiriman dan penanganan keluhan [5]. Adapun atribut yang digunakan pada penelitiannya seperti interface, feedback, harga, kualitas barang dan waktu pengiriman dengan gain tertinggi ada pada atribut harga sebesar 0,3793926 [6]. Berdasarkan hasil dari penelitiannya maka dilakukan penelitian ini dengan tujuan untuk mendapatkan informasi yang berbeda dengan cara menambah jumlah atribut pada penelitian seperti atribut mudah mencari produk, kelancaran proses pembelian, pengiriman produk, kejelasan informasi produk, respon tim layanan, solusi yang diberikan, tampilan dan navigasi, sering

menerima promosi, harga produk dan kecepatan konfirmasi. Penelitian ini menggunakan atribut yang lebih banyak dari penelitian sebelumnya dengan tujuan apakah akan mendapatkan hasil yang berbeda jika menambahkan beberapa atribut baru.

Untuk mengukur tingkat kepuasan pelanggan, perusahaan sering kali mengumpulkan data melalui survei atau ulasan pelanggan. Data ini kemudian dianalisis untuk mengidentifikasi faktor-faktor yang mempengaruhi kepuasan pelanggan dan untuk mengembangkan strategi peningkatan layanan. Dalam konteks ini, teknologi dan metode analisis data memainkan peran penting. Salah satu metode yang efektif untuk analisis data kepuasan pelanggan adalah algoritma C4.5. Algoritma C4.5 merupakan salah satu algoritma pembelajaran mesin yang populer untuk membuat pohon keputusan. Algoritma ini mampu menangani data dengan atribut numerik dan kategorikal, serta mampu menangani data yang hilang. Dengan menggunakan algoritma C4.5, perusahaan dapat mengidentifikasi pola-pola dalam data pelanggan yang menunjukkan faktor-faktor yang paling berpengaruh terhadap kepuasan. Pohon keputusan yang dihasilkan dapat digunakan untuk memprediksi tingkat kepuasan pelanggan berdasarkan berbagai atribut layanan.

Penelitian ini bertujuan untuk menganalisis tingkat kepuasan pelanggan terhadap pelayanan e-commerce dengan menggunakan algoritma C4.5. Melalui analisis ini, diharapkan dapat diidentifikasi faktor-faktor kunci yang mempengaruhi kepuasan pelanggan serta memberikan rekomendasi yang dapat membantu perusahaan e-commerce dalam meningkatkan kualitas layanan mereka. Hasil dari penelitian ini diharapkan tidak hanya memberikan kontribusi akademis dalam bidang analisis data dan pembelajaran mesin, tetapi juga memberikan manfaat praktis bagi perusahaan e-commerce dalam usaha mereka meningkatkan kepuasan dan loyalitas pelanggan.

Untuk mencapai tujuan penelitian ini, beberapa langkah akan dilakukan. Pertama, pengumpulan data kepuasan pelanggan melalui survei. Data ini akan mencakup berbagai atribut yang berhubungan dengan pengalaman pelanggan, seperti kualitas produk, kecepatan pengiriman, responsivitas layanan pelanggan, dan harga. Data yang dikumpulkan kemudian akan diproses dan dianalisis menggunakan algoritma C4.5. Dalam penelitian ini hasil survey kepuasan pengguna akan dianalisis dengan menggunakan metode algoritma C4.5 yang merupakan metode klasifikasi pohon keputusan yang mudah dipahami dan diinterpretasikan, selain itu algoritma C4.5 memiliki tingkat akurasi yang tinggi dibandingkan algoritma klasifikasi lainnya, dan juga merupakan algoritma yang efisien untuk menangani atribut bertipe diskret dan numerik.

Proses analisis akan melibatkan pembuatan pohon keputusan yang dapat mengidentifikasi faktor-faktor utama yang mempengaruhi kepuasan pelanggan. Pohon keputusan ini akan membantu dalam memvisualisasikan hubungan antara berbagai atribut layanan dan tingkat kepuasan pelanggan. Dengan menggunakan model ini, perusahaan dapat memprediksi bagaimana perubahan dalam atribut tertentu dapat mempengaruhi kepuasan pelanggan secara keseluruhan. Selain itu, penelitian ini juga akan mengevaluasi efektivitas model yang dihasilkan dalam memprediksi tingkat kepuasan pelanggan. Evaluasi akan dilakukan dengan menggunakan data uji untuk memastikan bahwa model tersebut memiliki akurasi yang tinggi dan dapat diandalkan dalam berbagai kondisi. Hasil evaluasi ini akan memberikan gambaran tentang seberapa baik model tersebut dapat digunakan dalam situasi nyata untuk membantu perusahaan e-commerce meningkatkan kualitas layanan mereka.

Di akhir penelitian, akan diberikan rekomendasi yang didasarkan pada temuan analisis. Rekomendasi ini akan mencakup strategi-strategi spesifik yang dapat diterapkan oleh perusahaan e-commerce untuk meningkatkan kepuasan pelanggan. Misalnya, jika analisis menunjukkan bahwa kecepatan pengiriman adalah faktor yang paling berpengaruh, perusahaan dapat fokus pada peningkatan logistik dan distribusi. Jika kualitas produk menjadi faktor utama, perusahaan dapat bekerja sama dengan pemasok untuk memastikan standar kualitas yang lebih tinggi. Penelitian ini juga akan mengidentifikasi area-area di mana perusahaan e-commerce mungkin memerlukan peningkatan lebih lanjut dan memberikan saran praktis tentang bagaimana meningkatkan atribut-atribut layanan yang kurang memuaskan. Dengan menerapkan rekomendasi ini, diharapkan perusahaan e-commerce dapat mencapai tingkat kepuasan pelanggan yang lebih tinggi, yang pada gilirannya akan meningkatkan loyalitas pelanggan dan mendukung pertumbuhan bisnis yang berkelanjutan.

Secara keseluruhan, penelitian ini bertujuan untuk memberikan wawasan yang mendalam tentang faktor-faktor yang mempengaruhi kepuasan pelanggan dalam layanan e-commerce dan menawarkan solusi praktis untuk meningkatkan pengalaman pelanggan. Dengan memahami dan mengatasi faktor-faktor tersebut, perusahaan e-commerce dapat menciptakan lingkungan belanja yang lebih memuaskan bagi pelanggan mereka, yang pada akhirnya akan meningkatkan loyalitas dan daya saing di pasar yang semakin kompetitif.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining merupakan bidang multidisiplin yang berada di antara teknologi basis data, statistik, ML, dan pengenalan pola yang mengambil keuntungan dari semua disiplin ilmu ini[7]. Data mining digunakan untuk menganalisis data besar untuk menemukan pola, relasi, dan wawasan yang berguna. Ini melibatkan penggunaan teknik statistik, matematika, dan algoritma pembelajaran mesin untuk mengidentifikasi informasi yang belum diketahui dalam data yang dapat membantu dalam pengambilan keputusan dan prediksi di masa depan. Data mining disebut juga dengan Knowledge Discovery in Database (KDD) ataupun pattern recognition digunakan untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna[8].

Data mining dapat digunakan untuk mengklasifikasikan. Klasifikasi data mining melibatkan sistem penyusunan dalam kelompok dan golongan menurut standar yang ditetapkan. Klasifikasi data adalah proses yang menemukan fitur yang sama pada himpunan objek dalam basis data dan mengklasifikasikannya ke dalam kelas-kelas yang berbeda berdasarkan model klasifikasi yang ditetapkan. Tujuan dari klasifikasi adalah agar model yang dihasilkan dapat digunakan untuk memprediksi kelas data yang tidak konsisten[9].

2.2 Klasifikasi

klasifikasi adalah metode untuk mengidentifikasi suatu bentuk atau pola yang menggambarkan serta membedakan jenis atau tipe data[10]. Klasifikasi merupakan salah satu teknik dalam data mining yang sering digunakan karena cukup sederhana. Ada dua tahap dalam teknik klasifikasi, yaitu mengembangkan model untuk pelatihan dan mengevaluasi model dengan menggunakan data pelatihan[11]. Model yang berbeda akan memiliki kekuatan yang berbeda dalam klasifikasi data yang akan mempengaruhi kinerja klasifikasi[12].

2.3 Decision Tree

Decision Tree adalah teknik berbasis pohon di mana setiap jalur mulai dari akar dijelaskan oleh data yang memisahkan urutan sampai hasil Boolean pada simpul daun tercapai. Ini adalah contoh hirarkis dari hubungan pengetahuan yang berisi node dan koneksi. Ketika relasi digunakan untuk mengklasifikasikan, node mewakili tujuan[13]. Decision Tree memakai kinerja dalam bentuk pohon (tree). Dalam pohon keputusan, masing-masing node yang dimiliki menjelaskan atribut yang ada, cabangnya menafsirkan nilai pada atribut yang ada dan leaf mempresentasikan kelas[14]. Node ini dinamakan pohon karena kemiripan bentuknya. Simpul akar adalah titik awal yang kemudian meluas ke berbagai cabang sehingga membentuk struktur seperti pohon. Pohon keputusan hanya memecah pohon menjadi sub-pohon berdasarkan jawaban atas pertanyaan, yaitu apakah ya atau tidak[15].

2.4 Algoritma C4.5

Algoritma C4.5 adalah salah satu metode klasifikasi dari data mining yang memberi kontribusi satu set data berlabel dan menghasilkan pohon keputusan sebagai keluaran yang mempunyai input berupa training samples dan sample training samples[16]. Banyak peneliti menggabungkan dan membandingkan algoritma c45 dengan algoritma lain seperti analisis komparatif metode Naïve Bayes, K Nearest Neighbor dan C4.5[17]. Algoritma C4.5 merupakan penerapan dari teknik pohon keputusan yang digunakan dalam klasifikasi data mining. Secara umum, algoritma C4.5 membangun pohon keputusan dengan langkah-langkah berikut:

- Memilih atribut sebagai simpul akar.
- Membuat cabang untuk setiap nilai atribut.
- Memisahkan kasus ke dalam cabang berdasarkan nilai atribut.
- Mengulangi proses untuk setiap cabang hingga semua cabang memiliki kelas yang sama.

Nilai keuntungan dari setiap atribut menentukan simpul dalam pohon keputusan menggunakan suatu rumus[18]:

$$\text{Entropy}(X) = \sum_{i=1}^n -P_i * \log_2 P_i \quad (1)$$

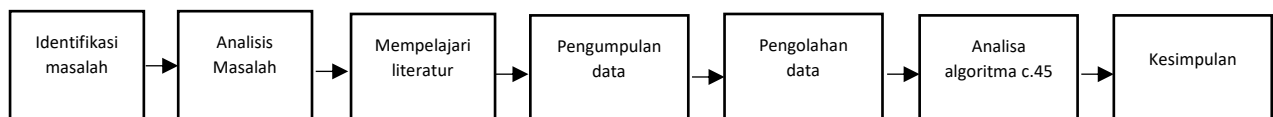
Dengan keterangan X merupakan jumlah himpunan kasus, N merupakan jumlah partisi X, dan P_i merupakan proporsi dari X_i dan X.

$$\text{Gain}(X, A) = \text{Entropy}(X) - \sum_{i=1}^n \frac{|X_i|}{|X|} * \text{Entropy}(X_i) \quad (2)$$

Dengan keterangan X merupakan jumlah himpunan kasus, A merupakan atribut, N merupakan jumlah partisi atribut A, $|X_i|$ merupakan jumlah kasus partisi ke -i, dan $|X|$ merupakan jumlah kasus dalam.

2.5 Metode Penelitian

Pada penelitian ini akan dilakukan dengan melalui beberapa tahapan metode penelitian yang dapat dilihat pada gambar 1 dibawah ini.



Gambar 1. Metode penelitian

Berikut adalah penjelasan dari gambar 1 diatas yaitu:

A. Identifikasi Masalah

Identifikasi permasalahan penelitian salah satu langkah yang paling penting dalam penulisan skripsi, tesis atau disertasi[19]. Dalam identifikasi masalah, penulis menemukan permasalahan dan merumuskan masalah. Bagaimana mengetahui kepuasan pelanggan terhadap aplikasi e-commerce.

B. Analisis Masalah

Analisis masalah terstruktur sangat relevan dalam menghadapi masalah yang kompleks dalam praktiknya[20]. Setelah mengetahui masalah, maka harus ditentukan ruang lingkup dari penelitian agar analisis masalah dapat lebih terarah.

C. Mempelajari Literatur

Literatur diambil dari jurnal dan buku yang berkaitan dengan kepuasan pelanggan terhadap aplikasi e-commerce dan tahun yang digunakan adalah maksimal 5 tahun terakhir.

D. Mengumpulkan Data

Pengumpulan data bertujuan untuk mendapatkan informasi yang diperlukan dalam rangka mencapai tujuan penelitian[21]. Data yang dikumpulkan adalah data kuisisioner menggunakan google form sebanyak 150 kuisisioner yang diberikan kepada responden secara acak. Kuisisioner berisi pertanyaan-pertanyaan yang mempengaruhi kepuasan pelanggan. Adapun pertanyaan yang diambil dalam penelitian ini terdiri dari 10 pertanyaan dengan jawaban skala 1-5 dengan keterangan sangat tidak puas sampai sangat puas, yakni:

- Seberapa mudah Anda menemukan produk yang Anda cari di platform e-commerce ini?
- Seberapa lancar proses pembelian produk di platform e-commerce ini?
- Bagaimana pendapat Anda tentang proses pengiriman produk yang Anda beli?
- Bagaimana pendapat Anda tentang kejelasan informasi produk yang tersedia di platform ini?
- Seberapa responsif tim layanan pelanggan dalam menanggapi pertanyaan atau keluhan Anda?
- Sejauh mana Anda puas dengan solusi yang diberikan oleh layanan pelanggan untuk masalah yang Anda hadapi?
- Sejauh mana Anda puas dengan tampilan dan navigasi situs web/aplikasi e-commerce ini?
- Seberapa sering Anda menerima promosi atau penawaran khusus dari platform ini?
- Apakah Anda merasa bahwa harga produk di platform ini bersaing?
- Seberapa cepat Anda menerima konfirmasi setelah melakukan pembelian?

E. Pengolahan Data

Pengolahan data menjelaskan prosedur pengolahan data analisis sesuai dengan pendekatan yang dilakukan, kemudian sebagai bahan penelitian disesuaikan dengan pembahasan pada penelitian ini menggunakan metode C4.5. Hasil pengolahan terhadap data ini nantinya dapat menjadi informasi[22].

F. Analisis Algoritma C4.5

Analisis yang dilakukan dalam penelitian ini yaitu pengolahan data dengan metode C4.5, melakukan analisis dan pengkajian secara mendalam terhadap hasil penelitian untuk diuraikan dengan melakukan pemecahan masalah menggunakan metode algoritma C4.5, kemudian melalui tools rapid miner memberikan gambaran yang jelas dan dapat menarik kesimpulan yang berkenaan dengan permasalahan yang diteliti.

3. HASIL DAN PEMBAHASAN

3.1 Analisa Data

Pengumpulan data set kepuasan pelanggan atau pengguna aplikasi e-commerce yaitu dengan menyebarkan kuisisioner untuk melakukan penelitian. Jumlah responden yang di gunakan pada penelitian ini adalah 151 responden. Berikut adalah perhitungan manual untuk mencari nilai entropy total.

A. Entropy

Diketahui:

- Jumlah kasus = 151
- Jumlah ya = 34
- Jumlah tidak = 117

$$H(P_{ya}, P_{tidak}) = \left(-\left(\frac{34}{151}\right) * \log_2 \left(\frac{34}{151}\right) \right) + \left(-\left(\frac{117}{151}\right) * \log_2 \left(\frac{117}{151}\right) \right)$$

$$\log_2(0.225165562) = \frac{\log_{10}(0.225165562)}{\log_{10}(2)} \approx -\frac{0.647207013}{0.301029995} \approx -2.150970197$$

$$\log_2(0.774834438) = \frac{\log_{10}(0.774834438)}{\log_{10}(2)} \approx -\frac{0.111815406}{0.301029995} \approx -0.371043837$$

$$-0.225165562 \log_2(0.225165562) = -0.225165562 \times -2.150970197 \approx 0.48455773$$

$$-0.774834438 \log_2(0.774834438) = -0.774834438 \times -0.371043837 \approx 0.284930393$$

$$H(p_{ya}, p_{tidak}) = 0.484557732 + 0.284930393 \approx 0.769488125$$

Setelah menemukan entropy total maka bisa dilakukan untuk mencari entropy dari masing-masing subset pada masing-masing atribut serta mencari semua nilai gain pada atribut-atribut yang ada. Perhitungan entropy dan gain selanjutnya dihitung menggunakan software microsoft excel agar mempermudah perhitungan dan meminimalisir kesalahan terhadap perhitungan nilai entropy dan gain.

Tabel 1. Data Record

Mudah mencari produk	Kelancaran proses pembelian	Pengiriman produk	Kejelasan informasi produk	Respon tim layanan	Solusi yang diberikan	Tampilan dan navigasi	Sering menerima promosi	Harga produk	Kecepatan konfirmasi	Kategori Kepuasan
Puas	Sangat Puas	Sangat Puas	Sangat Puas	Cukup	Cukup	Puas	Cukup	Puas	Puas	YA

Mudah mencari produk	Kelancaran proses pembelian	Pengiriman produk	Kejelasan informasi produk	Respon tim layanan	Solusi yang diberikan	Tampilan dan navigasi	Sering menerima promosi	Harga produk	Kecepatan konfirmasi	Kategori Kepuasan
Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Puas	Sangat Puas	Puas	Sangat Puas	YA
Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Buruk	Puas	Sangat Puas	YA
Cukup	Cukup	Cukup	Cukup	Cukup	Cukup	Buruk	Cukup	Puas	Cukup	TIDAK
Sangat Puas	Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	YA
Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Cukup	Cukup	Sangat Puas	Sangat Puas	Puas	Sangat Puas	YA
Sangat Puas	Puas	Cukup	Cukup	Buruk	Cukup	Puas	Puas	Sangat Puas	Cukup	YA
Sangat Puas	Puas	Puas	Puas	Puas	Puas	Puas	Sangat Buruk	Sangat Puas	Puas	YA
Puas	Puas	Puas	Puas	Puas	Puas	Cukup	Cukup	Puas	Puas	YA
Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	YA
Sangat Puas	Puas	Puas	Puas	Puas	Puas	Sangat Puas	Puas	Puas	Puas	YA
Sangat Puas	Puas	Cukup	Cukup	Cukup	Cukup	Puas	Cukup	Cukup	Cukup	TIDAK
Sangat Puas	Puas	Puas	Puas	Buruk	Sangat Buruk	Puas	Sangat Puas	Puas	Cukup	YA
Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	YA
Sangat Puas	Puas	Puas	Puas	Cukup	Cukup	Puas	Cukup	Sangat Puas	Puas	YA
Puas	Puas	Puas	Puas	Puas	Cukup	Cukup	Buruk	Cukup	Puas	YA
Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	YA
Sangat Puas	Sangat Puas	Puas	Puas	Cukup	Puas	Sangat Puas	Sangat Puas	Sangat Puas	Puas	YA
Buruk	Puas	Puas	Puas	Puas	Buruk	Buruk	Sangat Puas	Buruk	Buruk	TIDAK
Cukup	Cukup	Cukup	Cukup	Cukup	Cukup	Cukup	Buruk	Puas	Puas	TIDAK
Sangat Puas	Sangat Puas	Puas	Puas	Puas	Puas	Sangat Puas	Puas	Puas	Cukup	YA
Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	YA
Sangat Puas	Sangat Puas	Puas	Puas	Puas	Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	YA
Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Cukup	Puas	Sangat Puas	Cukup	Sangat Puas	Sangat Puas	YA
Puas	Puas	Cukup	Cukup	Puas	Puas	Cukup	Puas	Cukup	Puas	YA
Puas	Puas	Puas	Puas	Cukup	Puas	Puas	Cukup	Cukup	Puas	YA
Puas	Puas	Puas	Puas	Puas	Puas	Puas	Cukup	Cukup	Puas	YA
Puas	Puas	Puas	Puas	Cukup	Puas	Puas	Sangat Puas	Puas	Puas	YA
Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Cukup	Sangat Puas	Sangat Puas	YA

3.2 Pembahasan

Berdasarkan dari kuisioner yang telah dikumpulkan diambil 151 sampel untuk diolah. Hasil dari perhitungan menggunakan rumus 1 dan 2 dengan Gain tertinggi terdapat pada atribut pengiriman produk dengan nilai 0,337981562. Rinciannya sebagai berikut:

Tabel 2. Node awal

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain
1	Total	151	34	117	0,76948812	5
	Mudah mencari produk					0,23898196
	Sangat Buruk	5	3	2	0,97095059	4
	Buruk	8	8	0	0,99863596	0
	Cukup	23	12	11	0,47357292	4
	Puas	69	7	62	0,42622865	5
	Sangat Puas	46	4	42	0,42622865	7
	Kelancaran proses pembelian					0,27973183
	Buruk	5	5	0	0	5

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain		
Pengiriman produk	Cukup	21	13	8	0,95871188 3	0,33798156 2		
	Puas	74	15	59	0,72730067 9			
	Sangat Puas	51	0	51	0			
	Sangat Buruk	2	2	0	0			
	Buruk	7	6	1	0,59167277 9			
	Cukup	39	20	19	0,99952568 9			
	Puas	69	4	65	0,31933738 7			
	Sangat Puas	34	0	34	0			
	Kejelasan informasi produk	Sangat Buruk	2	2	0		0	0,31350383 9
		Buruk	7	6	1		0,59167277 9	
Cukup		39	21	18	0,99572745 2			
Puas		69	5	64	0,37505140 6			
Sangat Puas		34	0	34	0			
Respon tim layanan		Sangat Buruk	1	1	0	0	0,24262139 9	
		Buruk	11	9	2	0,68403843 6		
		Cukup	49	19	30	0,96333554 6		
		Puas	60	5	55	0,41381685 9		
		Sangat Puas	30	30	0	0		
	Solusi yang diberikan	Sangat Buruk	1	0	1	0		0,32358143 9
		Buruk	10	10	0	0		
		Cukup	44	19	25	0,98654463 0,22536363		
		Puas	55	2	53	9		
		Sangat Puas	41	2	39	0,28119379 6		
Tampilan dan navigasi		Sangat Buruk	3	3	0	0	0,32666236 7	
		Buruk	8	7	1	0,54356444 3		
		Cukup	33	17	16	0,99933750 4		
		Puas	70	6	64	0,42200051 7		
		Sangat Puas	37	37	0	0		
	Sering menerima promosi	Sangat Buruk	2	1	1	1		0,14975762 9
		Buruk	15	9	6	0,97095059 4		
		Cukup	41	14	27	0,92621221 3		

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain
Harga produk	Puas	48	5	43	0,482066148	0,266289818
	Sangat Puas	45	3	42	0,353359335	
	Sangat Buruk	1	1	0	0	
	Buruk	9	6	3	0,918295834	
	Cukup	29	16	13	0,992266639	
Kecepatan konfirmasi	Puas	71	9	62	0,548486609	0,265204674
	Sangat Puas	41	41	0	0	
	Sangat Buruk	2	2	0	0	
	Buruk	14	8	6	0,985228136	
	Cukup	34	16	18	0,997502546	
	Puas	62	6	56	0,458685816	
	Sangat Puas	39	39	0	0	

Berdasarkan tabel 2 dapat disimpulkan bahwa “jika pengiriman produk = sangat buruk / sangat bagus” maka nilai sudah final. Sedangkan jika nilai dari pengiriman produk = buruk / cukup / puas maka harus dicari akar selanjut nya

Tabel 3. Node 1.2

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain
1.2 Total Kriteria = Buruk		7	6	1	0,591672779	0,305958493
	Mudah mencari produk					
	Sangat Buruk	1	1	0	0	
	Buruk	1	1	0	0	
	Cukup	3	3	0	0	
Kelancaran proses pembelian	Puas	2	1	1	1	0,198117421
	Buruk	2	2	0	0	
	Cukup	2	2	0	0	
Kejelasan informasi produk	Puas	3	2	1	0,918295834	0
	Buruk	7	6	1	0,591672779	
Respon tim layanan		4	4	0	0	0,448815636
	Buruk	1	1	0	0	
	Puas	2	2	1	0,5	
Solusi yang diberikan						0,305958493
	Cukup	2	2	0	0	
	Buruk	2	2	0	0	
	Sangat Puas	2	1	1	1	
	Puas	1	1	0	0	

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain
Tampilan dan navigasi	Sangat Buruk	1	1	0	0	0,59167277
	Buruk	2	2	0	0	9
	Puas	2	2	0	0	
	Cukup	1	1	0	0	
	Sangat Puas	1	0	1	0	
Sering menerima promosi	Buruk	2	2	0	0	0,30595849
	Sangat Puas	1	1	0	0	3
	Puas	2	2	0	0	
	Cukup	2	1	1	1	
Harga produk	Buruk	3	3	0	0	0,19811742
	Sangat Buruk	1	1	0	0	1
	Cukup	3	2	1	4	0,91829583
Kecepatan konfirmasi	Buruk	3	3	0	0	0,44881563
	Cukup	3	3	0	0	6
	Puas	1	1	1	1	

Pada tabel 3 didapatkan gain tertinggi pada atribut tampilan dan navigasi dari perhitungan atribut pengiriman produk dengan subset buruk, karena entropy pada masing-masing subset adalah 0 maka dapat disimpulkan bahwa semua subset memiliki nilai final.

Tabel 4. Node 1.3

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain
1.3	Total Kriteria = Cukup	39	21	18	0,99572745	2
	Mudah mencari produk					0,75729513
	Sangat Buruk	1	1	0	0	7
	Buruk	4	4	0	0	
	Cukup	11	8	3	7	0,84535093
	Puas	14	0	14	0	
	Sangat Puas	9	4	5	0,99107606	
	Kelancaran proses pembelian					0,45814701
	Buruk	1	1	0	0	3
	Cukup	10	10	0	0	
Kejelasan informasi produk	Puas	21	10	11	3	0,99836367
	Sangat Puas	7	0	0	0	
	Cukup	39	21	18	2	0,99572745
	Respon tim layanan					0,54356444
Buruk	8	7	1	3	0,34493427	
Cukup	16	12	4	4	1	
Puas	13	2	11	5	0,81127812	
Sangat Puas	2	0	2	0	0,61938219	

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain
Solusi yang diberikan	Buruk	3	3	0	0	0,217775317
	Cukup	22	15	7	0,902393283	
	Puas	9	2	7	0,764204507	
	Sangat Puas	5	1	4	0,721928095	
	Sangat Buruk	1	1	0	0	
Tampilan dan navigasi	Buruk	2	2	0	0	0,293819639
	Cukup	19	14	5	0,831474388	
	Puas	13	4	9	0,89049164	
	Sangat Puas	4	0	4	0	
	Sangat Buruk	1	1	0	0	
Sering menerima promosi	Buruk	6	4	2	0,918295834	0,233680057
	Cukup	17	13	4	0,787126586	
	Puas	10	2	8	0,721928095	
	Sangat Puas	5	1	4	0,721928095	
	Sangat Buruk	1	1	0	0	
Harga produk	Buruk	15	12	3	0,721928095	0,26064308
	Cukup	18	8	10	0,99107606	
	Puas	5	0	5	0	
	Sangat Puas	5	0	5	0	
	Sangat Buruk	1	1	0	0	
Kecepatan Konfirmasi	Buruk	5	5	0	0	0,416803822
	Cukup	14	11	3	0,749595257	
	Puas	14	4	10	0,863120569	
	Sangat Puas	5	0	5	0	
	Sangat Buruk	1	1	0	0	

Pada tabel 4 dilakukan perhitungan berdasarkan atribut “pengiriman produk dengan subset cukup” dan didapatkan gain tertinggi ada pada atribut mudah mencari produk. Dapat disimpulkan jika “mudah mencari produk = sangat buruk / buruk / puas” maka nilai sudah final, sedangkan untuk subset cukup dan sangat puas harus dicarinya.

Tabel 5. Node 1.4

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain
1.4 Total Kriteria = Cukup Kelancaran proses pembelian		11	8	3	0,845350937	0,481714573
	Cukup	6	6	0	0	
	Puas	4	2	2	1	
	Sangat Puas	1	0	1	0	
Kejelasan informasi produk	Cukup	11	8	3	0,845350937	0
Respon tim layanan	Cukup	5	5	0	0	0,55034071
	Buruk	2	2	0	0	

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain
Solusi yang diberikan	Puas	4	1	3	0,811278124	0,26821316
	Cukup	8	7	1	0,543564443	
Sering menerima promosi	Puas	2	1	1	1	0,413088436
	Sangat Puas	1	0	1	0	
	Cukup	4	4	0	0	
	Buruk	2	1	1	1	
	Puas	2	2	0	0	
Harga produk	Sangat Puas	3	1	2	0,918295834	0,049452073
	Puas	6	4	2	0,918295834	
	Cukup	4	3	1	0,811278124	
	Buruk	1	1	0	0	
Kecepatan konfirmasi	Puas	5	4	1	0,721928095	0,075860623
	Puas	5	3	2	0,970950594	
	Buruk	1	1	0	0	
	Cukup	5	4	1	0,721928095	
	Puas	5	3	2	0,970950594	

Pada tabel 5 dilakukan perhitungan untuk atribut mudah mencari produk dengan subset cukup, gain tertinggi terdapat pada atribut kecepatan konfirmasi. Dapat disimpulkan jika “kecepatan konfirmasi = buruk” maka nilai sudah final, sedangkan untuk subset cukup dan puas harus dicari akar nya.

Tabel 6. Node 1.5

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain
1.5 Total Kriteria = Cukup Kelancaran proses pembelian		5	4	1	0,721928095	0,721928095
	Cukup	4	4	0	0	
Kejelasan informasi produk	Sangat Puas	1	1	0	0	0
	Cukup	5	4	1	0,721928095	
Respon tim layanan	Cukup	3	3	0	0	0,321928095
	Puas	2	1	1	1	
Solusi yang diberikan	Cukup	4	3	1	0,811278124	0,072905595
	Puas	1	1	0	0	
Sering menerima promosi	Cukup	4	4	0	0	0,721928095
	Sangat Puas	1	1	1	0	
	Cukup	4	4	0	0	
Harga Produk	Sangat Puas	1	1	1	0	0,170950594
	Cukup	2	2	0	0	
	Puas	3	2	1	0,918295834	

Pada tabel 6 dilakukan perhitungan untuk atribut kecepatan konfirmasi dengan subset cukup, gain tertinggi terdapat pada atribut sering menerima promosi dengan entropy pada semua subset adalah 0. Dapat disimpulkan bahwa semua subset telah mencapai nilai final.

Tabel . Node 1.6

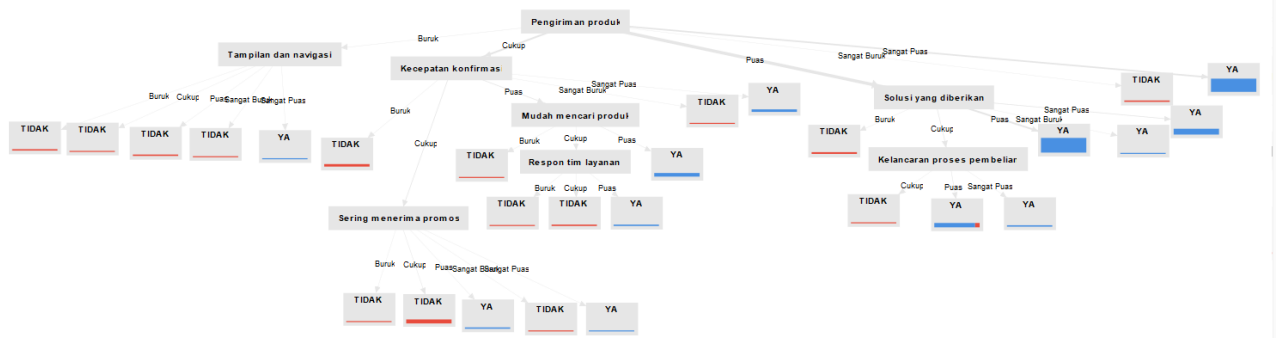
Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain
1.6 Total Kriteria = Puas Kelancaran proses pembelian		5	3	2	0,970950594	0,170950594
	Cukup	1	1	0	0	
Kejelasan informasi produk	Puas	4	2	2	1	0
	Cukup	5	3	2	0,970950594	
Respon tim layanan	Cukup	5	3	2	0,970950594	0,970950594
	Cukup	2	2	0	0	

Nod e	Keterangan	Jml kasus (S)	TIDAK (S1)	YA (S2)	Entropy	Gain	
Solusi yang diberikan	Buruk	1	1	0	0	0,970950594	
	Puas	2	0	2	0		
	Cukup	3	3	0	0		
	Sangat Puas	1	0	1	0		
Sering menerima promosi	Puas	1	0	1	0	0,170950594	
	Buruk	2	1	1	1		
	Puas	1	1	0	0		
	Sangat Puas	2	1	1	1		
Harga produk	Puas	3	2	1	0,918295834	0,419973094	
	Buruk	1	1	0	0		
	Cukup	1	0	1	0		

Pada tabel 7 dilakukan perhitungan untuk atribut kecepatan konfirmasi dengan subset puas, gain tertinggi terdapat pada atribut respon tim layanan dan atribut solusi yang diberikan dengan entropy pada semua subset adalah 0. Dapat disimpulkan bahwa semua subset memiliki nilai final.

3.3 Hasil

Setelah melakukan proses perhitungan dari atribut dan subset pada tabel di atas maka dapat terbentuk pohon keputusan yang dihasilkan oleh tools rapid miner seperti gambar berikut.



Gambar 3 Decision Tree

Dari pohon keputusan/decision tree yang dihasilkan oleh rapid miner, dihasilkan juga beberapa rules/aturan sesuai pohon keputusan yang telah didapat seperti gambar 3 di bawah ini

```

Pengiriman produk = Buruk
| | Tampilan dan navigasi = Buruk: TIDAK (YA=0, TIDAK=2)
| | Tampilan dan navigasi = Cukup: TIDAK (YA=0, TIDAK=1)
| | Tampilan dan navigasi = Puas: TIDAK (YA=0, TIDAK=2)
| | Tampilan dan navigasi = Sangat Buruk: TIDAK (YA=0, TIDAK=1)
| | Tampilan dan navigasi = Sangat Puas: YA (YA=1, TIDAK=0)
Pengiriman produk = Cukup
| | Kecepatan konfirmasi = Buruk: TIDAK (YA=0, TIDAK=5)
| | Kecepatan konfirmasi = Cukup
| | | Sering menerima promosi = Buruk: TIDAK (YA=0, TIDAK=1)
| | | Sering menerima promosi = Cukup: TIDAK (YA=0, TIDAK=9)
| | | Sering menerima promosi = Puas: YA (YA=2, TIDAK=0)
| | | Sering menerima promosi = Sangat Buruk: TIDAK (YA=0, TIDAK=1)
| | | Sering menerima promosi = Sangat Puas: YA (YA=1, TIDAK=0)
| | Kecepatan konfirmasi = Puas
| | | Mudah mencari produk = Buruk: TIDAK (YA=0, TIDAK=1)
| | | Mudah mencari produk = Cukup
| | | Respon tim layanan = Buruk: TIDAK (YA=0, TIDAK=1)
| | | Respon tim layanan = Cukup: TIDAK (YA=0, TIDAK=2)
| | | Respon tim layanan = Puas: YA (YA=2, TIDAK=0)
| | | Mudah mencari produk = Puas: YA (YA=0, TIDAK=0)
| | Kecepatan konfirmasi = Sangat Buruk: TIDAK (YA=0, TIDAK=1)
| | Kecepatan konfirmasi = Sangat Puas: YA (YA=0, TIDAK=0)
Pengiriman produk = Puas
| | Solusi yang diberikan = Buruk: TIDAK (YA=0, TIDAK=3)
| | Solusi yang diberikan = Cukup
| | Kelancaran proses pembelian = Cukup: TIDAK (YA=0, TIDAK=1)
| | Kelancaran proses pembelian = Puas: YA (YA=9, TIDAK=1)
| | Kelancaran proses pembelian = Sangat Puas: YA (YA=2, TIDAK=0)
| | Solusi yang diberikan = Puas: YA (YA=3, TIDAK=0)
| | Solusi yang diberikan = Sangat Buruk: YA (YA=1, TIDAK=0)
| | Solusi yang diberikan = sangat Puas: YA (YA=14, TIDAK=0)
Pengiriman produk = Sangat Buruk: TIDAK (YA=0, TIDAK=2)
Pengiriman produk = Sangat Puas: YA (YA=34, TIDAK=0)
    
```

Gambar 3. Rule decision tree

Dari perhitungan nilai entropy dan gain pada pembahasan sebelumnya didapatkan nilai gain tertinggi terdapat pada atribut pengiriman produk yang berarti atribut pengiriman produk merupakan atribut yang paling signifikan terhadap kepuasan pelanggan. Pada pengujian performa menggunakan tool rapid miner, model yang dihasilkan memiliki akurasi sebesar 93%, menunjukkan kemampuan yang baik dalam memprediksi kepuasan pelanggan. Detail dari akurasi dapat dilihat pada gambar 4 dibawah ini.

accuracy: 93.33%

	true YA	true TIDAK	class precision
pred. YA	22	1	95.65%
pred. TIDAK	1	6	85.71%
class recall	95.65%	85.71%	

Gambar 5 Akurasi model

4. KESIMPULAN

Dalam penelitian ini, kami menerapkan dan menguji metode perhitungan entropy serta information gain untuk membangun pohon keputusan dalam analisis atribut-atribut yang mempengaruhi kepuasan pelanggan. Sepuluh atribut utama yang dianalisis meliputi: Mudah mencari produk, Kelancaran proses pembelian, Pengiriman produk, Kejelasan informasi produk, Respon tim layanan, Solusi yang diberikan, Tampilan dan navigasi, Sering menerima promosi, Harga produk, dan Kecepatan konfirmasi. Analisis dimulai dengan menghitung entropy untuk setiap subset data berdasarkan atribut-atribut tersebut. Information gain digunakan untuk menentukan atribut paling signifikan dalam membagi data pada setiap node pohon keputusan. Hasil perhitungan menunjukkan atribut "Pengiriman produk" memiliki gain tertinggi pada langkah awal dengan entropy sebagai berikut: Sangat Buruk (0), Buruk (0,591672779), Cukup (0,999525689), Puas (0,319337387), dan Sangat Puas (0). Selanjutnya, atribut dengan gain tertinggi dalam subset "Pengiriman produk = Buruk" adalah "Tampilan dan navigasi". Proses serupa dilakukan untuk subset lainnya, termasuk "Pengiriman produk = Cukup" dan "Kecepatan konfirmasi". Setiap langkah menggunakan atribut dengan gain tertinggi untuk membagi data hingga entropi nol atau mendekati nol, menandakan homogenitas data pada node tersebut. Atribut seperti "Kecepatan konfirmasi" dan "Respon tim layanan" juga ditemukan berpengaruh signifikan terhadap kepuasan pelanggan. Namun, penelitian ini memiliki keterbatasan utama, yakni tidak dianalisisnya subset "Puas" dari atribut "Pengiriman produk". Adapun model yang dihasilkan memiliki akurasi sebesar 93%, menunjukkan kemampuan yang baik dalam memprediksi kepuasan pelanggan.

REFERENCES

- [1] A. S. S. o'g'li Shirin, Mirzaeva Nodirovna, "E-Commerce Trends Shaping The Future of Retail.pdf." 2024.
- [2] R. Setyawati, "Pengaruh Kualitas Pelayanan Terhadap Tingkat Kepuasan Konsumen," *Inov. J. Ekon. Keuang. dan Manaj.*, vol. 19, no. 1, pp. 57–63, 2023, [Online]. Available: <https://journal.feb.unmul.ac.id/index.php/INOVASI/article/view/12660/2345>
- [3] W. Yuliyanto, "Pengaruh Promosi Terhadap Kepuasan Pelanggan di Toko Aleca Shopid Kebumen," *J. Bus. Econ. Res.*, vol. 1, no. 2, pp. 168–172, 2020, doi: 10.47065/jbe.v1i2.244.
- [4] F. Anggraini and A. Budiarti, "PENGARUH HARGA, PROMOSI, DAN KUALITAS PELAYANAN TERHADAP LOYALITAS PELANGGAN DIMEDIASI KEPUASAN PELANGGAN PADA KONSUMEN GOJEK," *J. Pendidik. Ekon.*, vol. 08, pp. 86–94, 2020, [Online]. Available: <https://ejournal.unesa.ac.id/index.php/jupe/article/view/36354>
- [5] S. A. A. Wazir, Indra, "Analisis Kepuasan Konsumen Dengan Algoritma C4.5 Pada Umkm Minimall," *Comasie*, vol. 3, no. 3, pp. 21–30, 2020.
- [6] Nurul Farhana, Harly Okprana, and Rizky Khairunnisa Sormin, "Analisis Tingkat Kepuasan Pelanggan Pada Aplikasi Tiktok Shop Dengan Metode Algoritma C4.5," *SmartEDU J.*, vol. 1, no. 3, pp. 101–111, 2022.
- [7] W. T. Wu et al., "Data Mining in Clinical Big Data: The Frequently Used Databases, Steps, And Methodological Models," *Mil. Med. Res.*, vol. 8, no. 1, pp. 1–12, 2021, doi: 10.1186/s40779-021-00338-z.
- [8] C. R. Aditya Nugroho and T. Kristiana, "Penerapan Algoritma C4.5 Untuk Kepuasan Pelanggan Toko Online Parfume Chantik," *J. Algoritm.*, vol. 3, no. 1, pp. 10–21, 2022, doi: 10.35957/algoritm.v3i1.3169.
- [9] K. A. Saputra, J. T. Hardinata, M. R. Lubis, S. R. Andani, and I. S. Saragih, "Klasifikasi Algoritma C4.5 Dalam Penerapan Tingkat Kepuasan Siswa Terhadap Media Pembelajaran Online," *Kaji. Ilm. Inform. dan Komput.*, vol. 1, no. 3, pp. 113–118, 2020, [Online]. Available: <https://djournals.com/klik>
- [10] S. N. Bardab, T. M. Ahmed, and T. A. A. Mohammed, "Data mining classification algorithms: An overview," *Int. J. Adv. Appl. Sci.*, vol. 8, no. 2, pp. 1–5, 2021, doi: 10.21833/ijaas.2021.02.001.
- [11] A. Triayudi and W. O. Widyarto, "Educational Data Mining Analysis Using Classification Techniques," *J. Phys. Conf. Ser.*, vol. 1933, no. 1, 2021, doi: 10.1088/1742-6596/1933/1/012061.
- [12] R. C. Chen, C. Dewi, S. W. Huang, and R. E. Caraka, "Selecting critical features for data classification based on machine learning methods," *J. Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00327-4.
- [13] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 20–28, 2021, doi: 10.38094/jast20165.
- [14] S. W. Siahaan, K. D. R. Sianipar, and P. P. P. A. N. W. F. I. R. H. Zer, "Penerapan Algoritma C4.5 Dalam Meningkatkan Kemampuan Bahasa Inggris Pada Mahasiswa," vol. 13, no. 2, pp. 229–239, 2020.
- [15] M. Bansal, A. Goyal, and A. Choudhary, "A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning," *Decis. Anal. J.*, vol. 3, no. May, p. 100071, 2022, doi: 10.1016/j.dajour.2022.100071.
- [16] M. F. Rizqullah, N. T. Raihana, and M. I. Jambak, "Komparasi Penerapan Algoritma C4.5, K-Nearest Neighbor, dan Naïve Bayes untuk Keberlangsungan Pasien Galag Jantung," *Klik Kaji. Ilm. Inform. Dan Komput.*, vol. 4, no. 5, pp. 2580–2587,

- 2024, doi: 10.30865/klik.v4i5.1788.
- [17] R. Sovia, A. Muhammad, S. Arlis, Guslendra, and S. Defit, "Analysis of sales levels of pharmaceutical products by using data mining algorithm C45," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 22, no. 1, pp. 476-484, 2021, doi: 10.11591/ijeecs.v22.i1.pp476-484.
- [18] S. A. S. Ratih Nurdiani Sari, Novrina, "Penerapan algoritma c4.5 untuk memprediksi bencana gunung meletus di indonesia," *Tek. dan Sci.*, vol. 3, no. 2, pp. 1-9, 2024.
- [19] A. R. S. Nasution, "Identifikasi Permasalahan Penelitian," *ALACRITY J. Educ.*, vol. 1, no. 2, pp. 13-19, 2021, doi: 10.52121/alacrity.v1i2.21.
- [20] L. Thorndahl, Katrine and D. Stentoft, "Educating Students for a Complex Future – Why Integrating a Problem Analysis in Problem Based Learning has Something to Offer," *Interdiscip. J. Probl. Learn.*, vol. 14, no. 1, 2020.
- [21] A. Setiawan and D. Pasha, "Sistem Pengolahan Data Penilaian Berbasis Web Menggunakan Metode Picies," *J. Teknol. dan Sist. Inf.*, vol. 1, no. 1, pp. 97-104, 2020, doi: 10.33365/jtsi.v1i1.225.
- [22] I. Ahmad, S. Samsugi, and Y. Irawan, "Implementasi Data Mining Sebagai Pengolahan Data," *J. Teknoinfo*, vol. 16, no. 1, p. 46, 2022, [Online]. Available: <http://portaldata.org/index.php/portaldata/article/view/107>