

Pemodelan Pohon Keputusan Menggunakan Algoritma ID3 dalam Pendekatan Data Mining

¹Suhendri, ²Rama Saktriawindarta, ³Nurita Evitarina

^{1,2,3}Universita Pertiba, Indonesia

¹akh.suhendri@gmail.com; ²ramasaktriawindarta@gmail.com; ³nuritaevitarina94@gmail.com;

Article Info

Article history:

Received, 2026-01-06

Revised, 2026-01-15

Accepted, 2026-01-20

Kata Kunci:

data mining,
decision tree,
ID3,
klasifikasi data

Keywords:

data mining,
decision tree,
ID3,
data classification

ABSTRAK

Perkembangan teknologi informasi mendorong pemanfaatan *data mining* sebagai dasar pengambilan keputusan berbasis data pada berbagai sektor, termasuk industri hiburan karaoke. Penelitian ini bertujuan untuk mengevaluasi kinerja algoritma ID3 dalam mendukung sistem pendukung keputusan melalui pembentukan model klasifikasi berbasis pohon keputusan. Metode penelitian menggunakan pendekatan *Knowledge Discovery in Database* (KDD) yang meliputi tahapan seleksi data, transformasi data, pemodelan menggunakan algoritma ID3, serta evaluasi hasil keputusan. Evaluasi metode dilakukan berdasarkan lima aspek utama, yaitu kemampuan pembentukan keputusan, kecepatan proses klasifikasi, stabilitas hasil klasifikasi, kemudahan interpretasi model, dan kesesuaian keputusan dengan kebutuhan pengguna. Hasil penelitian menunjukkan bahwa algoritma ID3 mencapai tingkat keberhasilan rata-rata sebesar 92%, dengan nilai tertinggi pada aspek kecepatan dan stabilitas klasifikasi. Temuan ini menunjukkan bahwa algoritma ID3 efektif, efisien, dan mudah diinterpretasikan, sehingga layak diterapkan sebagai metode klasifikasi dalam sistem pendukung keputusan berbasis *data mining*.

ABSTRACT

The rapid development of information technology has encouraged the use of data mining as a foundation for data-driven decision-making across various sectors, including the karaoke entertainment industry. This study aims to evaluate the performance of the ID3 algorithm in supporting decision support systems through the construction of a decision tree-based classification model. The research method employs the *Knowledge Discovery in Databases* (KDD) approach, which involves data selection, data transformation, modeling using the ID3 algorithm, and evaluation of decision outcomes. The performance of the method was evaluated based on five key aspects: decision-making capability, classification processing speed, classification result stability, model interpretability, and suitability to user needs. The results indicate that the ID3 algorithm achieved an average success rate of 92%, with the highest performance observed in processing speed and classification stability. These findings demonstrate that the ID3 algorithm is effective, efficient, and highly interpretable, making it suitable for implementation as a classification method in data mining-based decision support systems.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



Penulis Korespondensi:

Suhendri,
Program Studi Rekayasa Komputer,
Universitas Pertiba,
Email: akh.suhendri@gmail.com

1. PENDAHULUAN

Perkembangan teknologi informasi yang pesat serta meningkatnya kebutuhan pengambilan keputusan berbasis data telah mendorong pemanfaatan *data mining* di berbagai sektor. Data mining merupakan proses sistematis untuk mengekstraksi pengetahuan, pola, hubungan, dan kecenderungan tersembunyi dari basis data berukuran besar dengan memanfaatkan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* yang sulit diperoleh melalui analisis manual [1]. Dalam kerangka *Knowledge Discovery in Database* (KDD), *data mining*

berperan sebagai tahapan inti yang mencakup proses pembersihan data, integrasi data, seleksi data, transformasi data, pemodelan, evaluasi pola, serta penyajian pengetahuan [2]. Integrasi *data mining* dengan sistem pendukung keputusan telah terbukti mampu meningkatkan kualitas dan akurasi pengambilan keputusan di berbagai domain aplikasi [3].

Industri hiburan, khususnya karaoke, mengalami pertumbuhan yang signifikan dalam beberapa tahun terakhir. Laporan industri menunjukkan bahwa pasar sistem karaoke global mencapai nilai USD 5,61 miliar pada tahun 2023 dan diproyeksikan terus tumbuh dengan *Compound Annual Growth Rate* (CAGR) sebesar 2,28% hingga tahun 2031, dengan kawasan Asia-Pasifik menguasai porsi pasar terbesar [4]. Karaoke merupakan bentuk hiburan interaktif yang memadukan aktivitas bernyanyi dengan tampilan lirik berbasis sistem komputer. Di Indonesia, karaoke keluarga berkembang pesat dan menjadi salah satu alternatif hiburan yang diminati masyarakat untuk mengisi waktu luang serta mengurangi kejenuhan akibat aktivitas kerja yang intensif. Seiring dengan meningkatnya jumlah pengunjung dan bertambahnya koleksi lagu, pengelolaan data lagu pada tempat karaoke menjadi semakin kompleks. Tren terkini menunjukkan bahwa sistem karaoke modern mulai mengadopsi teknologi berbasis kecerdasan buatan untuk meningkatkan personalisasi pengalaman pengguna, seperti algoritma rekomendasi lagu, antarmuka interaktif, dan integrasi sistem digital [5]. Temuan lain menunjukkan bahwa sekitar 27% konsumen di kawasan Asia-Pasifik lebih memilih sistem karaoke dengan fitur personalisasi berdasarkan preferensi pengguna [6]. Meskipun demikian, banyak tempat karaoke keluarga masih mengelola koleksi lagu secara konvensional tanpa dukungan analisis data yang sistematis, sehingga penentuan playlist sering kali tidak sepenuhnya mencerminkan preferensi mayoritas pengunjung dan berpotensi menurunkan tingkat kepuasan pelanggan.

Salah satu pendekatan data mining yang banyak digunakan untuk permasalahan klasifikasi adalah *decision tree*. Metode ini merepresentasikan pengetahuan dalam bentuk struktur pohon, di mana setiap simpul menunjukkan atribut, setiap cabang merepresentasikan nilai atribut, dan simpul daun menunjukkan kelas keputusan [7]. *Decision tree* banyak digunakan karena kemampuannya menghasilkan model yang mudah dipahami dan diinterpretasikan oleh pengguna non-teknis [8]. Beberapa algoritma populer dalam pembentukan *decision tree* antara lain C4.5, CART, *Random Forest*, dan ID3 [9].

Algoritma ID3 (*Iterative Dichotomiser 3*) yang dikembangkan oleh Quinlan merupakan algoritma klasik dalam pembentukan *decision tree* yang bekerja dengan memilih atribut berdasarkan nilai *information gain* tertinggi [10]. *Information gain* mengukur seberapa besar penurunan entropi atau tingkat ketidakpastian data setelah dilakukan pemisahan berdasarkan atribut tertentu [11]. Prinsip ini memungkinkan algoritma ID3 membentuk struktur pohon keputusan secara iteratif dan sistematis [12]. Berbagai penelitian menunjukkan bahwa algoritma *decision tree*, termasuk ID3, masih relevan dan efektif dalam menyelesaikan permasalahan klasifikasi di berbagai domain karena kesederhanaan model dan kemudahan interpretasi hasilnya [13].

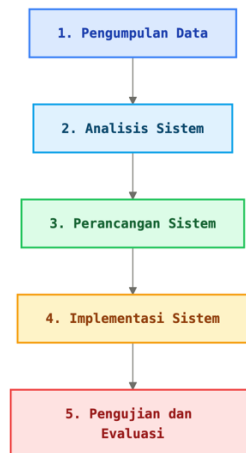
Berbagai penelitian dalam lima tahun terakhir telah membuktikan efektivitas algoritma *decision tree* dalam mendukung sistem pendukung keputusan. Widaningsih et al. menunjukkan bahwa algoritma ID3 mampu mengklasifikasikan data secara akurat berdasarkan atribut multivariat [1]. Huda et al. melaporkan bahwa algoritma *decision tree* memberikan tingkat akurasi tinggi pada permasalahan prediksi berbasis data pelanggan. [14] Rahman et al. menekankan pentingnya integrasi data mining dalam pengembangan sistem pendukung keputusan yang terstruktur dan holistic [3]. sementara Hosseini et al. menunjukkan bahwa integrasi *machine learning* dengan sistem pendukung keputusan memberikan kontribusi signifikan terhadap peningkatan kualitas analisis data [15]. Sistem Pendukung Keputusan (*Decision Support System/DSS*) adalah sistem berbasis komputer yang dirancang untuk membantu pengambil keputusan dalam menyelesaikan masalah semi terstruktur atau tidak terstruktur [16]. Menurut Asgarova et al. [17] dan Hussein et al. [3], DSS berperan penting dalam meningkatkan manajemen perusahaan dengan memanfaatkan analitika data dan kecerdasan buatan untuk memfasilitasi pengambilan keputusan strategis. Sistem ini memungkinkan organisasi untuk memprediksi tren, mengoptimalkan strategi, dan menganalisis berbagai skenario bisnis yang pada akhirnya mengarah pada peningkatan efisiensi operasional. Penelitian terkini oleh Necula dan Mihailescu [18] menunjukkan bahwa pendekatan *machine learning* dapat menganalisis pola rumit dalam data *clickstream* untuk memahami perilaku konsumen secara lebih komprehensif.

Meskipun berbagai penelitian tersebut telah mengkaji penerapan algoritma *decision tree* dan ID3 pada beragam konteks, masih terdapat kesenjangan penelitian (*research gap*) dalam pemanfaatan algoritma ID3 yang secara khusus diintegrasikan ke dalam sistem pendukung keputusan untuk pengelolaan koleksi lagu pada sistem karaoke keluarga. Sebagian besar penelitian sebelumnya berfokus pada domain pendidikan, bisnis, atau *e-commerce*, sementara penerapan ID3 pada konteks hiburan interaktif, khususnya dalam mendukung pengambilan keputusan berbasis preferensi pengguna, masih terbatas dan belum banyak dikaji secara mendalam. Berdasarkan gap penelitian tersebut, tujuan dari penelitian ini adalah mengimplementasikan dan menganalisis algoritma ID3 dalam sistem pendukung keputusan berbasis *data mining* pada sistem karaoke

keluarga guna mengidentifikasi pola preferensi lagu pengunjung dan menghasilkan rekomendasi keputusan yang akurat serta mudah diinterpretasikan.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan penelitian dan pengembangan (*Research and Development*) dengan fokus pada pembuatan aplikasi data mining untuk Family Karaoke. Tahapan penelitian dilaksanakan secara sistematis untuk memastikan hasil yang valid dan dapat dipertanggungjawabkan. Berikut adalah tahapan metodologi penelitian yang dilakukan:



Gambar 1 Diagram Alir Metodologi Penelitian

Pengumpulan data dalam penelitian ini menggunakan dua jenis sumber data. Sumber pertama adalah data primer yang diperoleh secara langsung dari lokasi penelitian, yaitu Family Karaoke. Data primer dikumpulkan melalui wawancara terstruktur dengan pengelola serta observasi langsung terhadap sistem yang sedang berjalan. Sumber kedua adalah data sekunder yang diperoleh secara tidak langsung melalui studi literatur, jurnal ilmiah, buku referensi, dokumentasi sistem, dan sumber daring yang relevan dengan topik penelitian. Proses pengumpulan data dilakukan dengan empat metode. Metode pertama adalah observasi, yaitu pengamatan langsung terhadap proses operasional Family Karaoke untuk memahami alur kerja sistem pemilihan dan penyimpanan lagu yang sedang berjalan. Metode kedua adalah wawancara, yaitu dialog terstruktur dengan pengelola Family Karaoke untuk menggali informasi mengenai permasalahan yang dihadapi serta kebutuhan sistem yang diharapkan. Metode ketiga adalah studi pustaka, yaitu kajian terhadap literatur ilmiah, jurnal, dan referensi terkait *data mining*, algoritma ID3, dan sistem pendukung keputusan sebagai landasan teoritis penelitian. Metode keempat adalah dokumentasi, yaitu pengumpulan dokumen pendukung berupa data historis pemutaran lagu, preferensi pengunjung, dan arsip terkait yang digunakan sebagai dataset penelitian.

Tahap analisis sistem bertujuan untuk memahami kebutuhan pengguna dan mendefinisikan spesifikasi sistem. Analisis dilakukan terhadap tiga aspek utama, yaitu: analisis kebutuhan input yang meliputi data lagu dan atribut klasifikasi; analisis proses yang mencakup mekanisme perhitungan algoritma ID3 untuk menghasilkan pohon keputusan; serta analisis *output* berupa rekomendasi lagu yang sebaiknya disimpan atau dihapus dari sistem karaoke. Perancangan sistem dilakukan untuk menggambarkan arsitektur dan alur kerja aplikasi yang akan dibangun. Metode perancangan yang digunakan meliputi *Use Case Diagram* untuk menggambarkan interaksi antara pengguna dengan sistem, serta *Activity Diagram* untuk menjelaskan alur aktivitas dalam proses klasifikasi lagu menggunakan algoritma ID3. Perancangan antarmuka pengguna juga dilakukan untuk memastikan kemudahan penggunaan aplikasi.

Implementasi merupakan tahap penerapan rancangan sistem ke dalam bentuk aplikasi yang fungsional. Pada tahap ini, algoritma ID3 diimplementasikan untuk melakukan klasifikasi terhadap data lagu berdasarkan atribut-atribut yang telah ditentukan. Hasil klasifikasi berupa keputusan apakah suatu lagu layak disimpan atau dihapus dari sistem karaoke berdasarkan pola preferensi pengunjung. Tahap akhir penelitian adalah pengujian dan evaluasi sistem. Pengujian dilakukan untuk memastikan aplikasi berfungsi sesuai dengan spesifikasi yang telah ditetapkan. Metode pengujian yang digunakan adalah *black-box testing* untuk menguji fungsionalitas sistem dari perspektif pengguna. Evaluasi juga mencakup pengujian kompatibilitas aplikasi pada berbagai platform sistem operasi untuk memastikan portabilitas dan stabilitas sistem.

3. HASIL DAN ANALISIS

Evaluasi pada penelitian ini difokuskan pada kinerja metode klasifikasi algoritma ID3 dalam mendukung pengambilan keputusan berbasis data, bukan pada aspek teknis atau fungsional aplikasi. Oleh karena itu, pengukuran keberhasilan dilakukan dengan menilai kualitas hasil klasifikasi, konsistensi model keputusan, efisiensi proses inferensi, serta tingkat keterterimaan keputusan yang dihasilkan oleh model pohon keputusan dari perspektif pengguna.

Pengukuran kinerja metode dilakukan melalui evaluasi terstruktur terhadap hasil keputusan yang dihasilkan oleh algoritma ID3. Sebanyak 10 responden dilibatkan untuk menilai hasil klasifikasi berdasarkan lima aspek utama, yaitu: (1) kemampuan metode dalam membentuk keputusan yang relevan, (2) kecepatan proses klasifikasi, (3) stabilitas hasil klasifikasi, (4) kemudahan interpretasi model keputusan, dan (5) kesesuaian keputusan dengan kebutuhan pengguna. Setiap aspek dinilai menggunakan skala kategorikal yang selanjutnya dikonversi menjadi persentase keberhasilan metode.

Tabel 1. Ringkasan Hasil Evaluasi Metode ID3

Aspek Evaluasi Metode	Persentase Keberhasilan
Kemampuan Pembentukan Keputusan	90%
Kecepatan Proses Klasifikasi	100%
Stabilitas Hasil Klasifikasi	100%
Kemudahan Interpretasi Model	90%
Kesesuaian dengan Kebutuhan Pengguna	80%
Rata-rata Keberhasilan Metode ID3	92%

Berdasarkan Tabel 1, terlihat bahwa algoritma ID3 menunjukkan kinerja yang sangat baik pada hampir seluruh aspek evaluasi. Aspek kecepatan proses klasifikasi dan stabilitas hasil klasifikasi memperoleh nilai keberhasilan tertinggi, masing-masing sebesar 100%. Hal ini menunjukkan bahwa algoritma ID3 mampu melakukan proses inferensi secara efisien dan menghasilkan keputusan yang konsisten terhadap data masukan yang sama.

Pada aspek kemampuan pembentukan keputusan, algoritma ID3 memperoleh tingkat keberhasilan sebesar 90%. Capaian ini mengindikasikan bahwa struktur pohon keputusan yang terbentuk mampu merepresentasikan hubungan antar atribut secara efektif, sehingga keputusan yang dihasilkan dinilai relevan dan dapat digunakan sebagai dasar pengambilan keputusan oleh mayoritas pengguna. Aspek kemudahan interpretasi model juga memperoleh nilai 90%, yang menunjukkan bahwa model pohon keputusan yang dihasilkan oleh algoritma ID3 mudah dipahami. Karakteristik ini merupakan keunggulan utama *decision tree*, karena setiap keputusan dapat ditelusuri melalui aturan berbasis logika (*if-then rules*), sehingga meningkatkan transparansi proses pengambilan keputusan.

Sementara itu, aspek kesesuaian dengan kebutuhan pengguna memperoleh nilai keberhasilan sebesar 80%. Nilai ini menunjukkan bahwa meskipun sebagian besar keputusan telah sesuai dengan kebutuhan pengguna, masih terdapat ruang untuk peningkatan akurasi model. Keterbatasan ini diduga dipengaruhi oleh jumlah data latih dan variasi atribut yang digunakan dalam pembentukan pohon keputusan. Secara keseluruhan, hasil evaluasi menunjukkan bahwa algoritma ID3 mencapai tingkat keberhasilan rata-rata sebesar 92%, yang mengindikasikan bahwa metode ini sangat efektif dalam mendukung proses pengambilan keputusan berbasis data mining.

Hasil pemodelan menggunakan algoritma ID3 menghasilkan struktur pohon keputusan yang menempatkan atribut Usia sebagai *root node*. Pemilihan atribut ini didasarkan pada nilai *information gain* tertinggi, sesuai dengan prinsip dasar algoritma ID3. Struktur pohon selanjutnya menunjukkan segmentasi data yang semakin spesifik hingga menghasilkan keputusan akhir pada simpul daun. Struktur pohon keputusan yang terbentuk menunjukkan bahwa algoritma ID3 mampu mengorganisasi atribut secara hierarkis berdasarkan kontribusinya terhadap proses klasifikasi. Hal ini menegaskan bahwa metode ID3 tidak hanya efektif secara kuantitatif, tetapi juga valid secara konseptual dalam merepresentasikan pola keputusan berbasis data.

Sebagai bentuk implementasi dan validasi metode, penelitian ini menghasilkan sebuah *prototype* sistem pendukung keputusan berbasis data mining yang mengintegrasikan algoritma ID3. *Prototype* ini tidak dievaluasi sebagai produk perangkat lunak, melainkan digunakan sebagai media untuk merepresentasikan proses kerja algoritma ID3 dalam membentuk keputusan. *Prototype* dirancang untuk memfasilitasi proses input data, perhitungan *entropy* dan *information gain*, pembentukan pohon keputusan, serta penyajian hasil keputusan secara visual. Melalui *prototype* ini, pengguna dapat mengamati secara langsung bagaimana kombinasi atribut diproses oleh algoritma ID3 hingga menghasilkan keputusan akhir.



Gambar 2. Aplikasi Data Mining Family Karaoke

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma ID3 menunjukkan kinerja yang sangat baik dalam mendukung sistem pendukung keputusan berbasis data mining. Evaluasi metode yang dilakukan melalui lima aspek utama menghasilkan tingkat keberhasilan rata-rata sebesar 92%, yang mencerminkan kemampuan algoritma ID3 dalam membentuk keputusan yang relevan, cepat, stabil, dan mudah diinterpretasikan oleh pengguna. Mekanisme pemilihan atribut berbasis information gain terbukti mampu membangun struktur pohon keputusan yang konsisten dan representatif terhadap pola data. Dengan demikian, algoritma ID3 dapat dinyatakan sebagai metode klasifikasi yang efektif dan terukur untuk diterapkan dalam sistem pendukung keputusan, meskipun pengembangan lebih lanjut masih diperlukan melalui penggunaan dataset yang lebih besar dan evaluasi metrik klasifikasi yang lebih komprehensif.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah mendukung pelaksanaan penelitian ini, serta para responden yang telah berpartisipasi.

REFERENSI

- [1] S. Widaningsih, W. Muhamad, R. Hendriyanto, and H. Nugroho, "An ID3 Decision Tree Algorithm-Based Model for Predicting Student Performance Using Comprehensive Student Selection Data at Telkom University," *Ingénierie des Systèmes d'Information*, vol. 28, no. 5, pp. 1199–1208, 2023, doi: 10.18280/isi.280508.
- [2] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From Data Mining to Knowledge Discovery in Databases," *AI Mag*, vol. 17, no. 3, pp. 37–54, 1996, doi: 10.1609/aimag.v17i3.1230.
- [3] A. Rahman, M. Hussein, and others, "Development Process of Decision Support Systems Using Data Mining Technology," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 1, pp. 703–714, Oct. 2024, doi: 10.11591/ijeecs.v36.i1.pp703-714.
- [4] Kings Research, "Karaoke System Market Size, Share and Growth Report 2031," Oct. 2024. [Online]. Available: <https://www.kingsresearch.com/karaoke-system-market-458>
- [5] Technavio, "Karaoke Market to Grow by USD 450.8 Million (2024-2028), Rising Popularity of Music-Related Leisure Activities Boosting Growth," Dec. 2024. [Online]. Available: <https://www.technavio.com/report/karaoke-market-industry-analysis>
- [6] Industry Research, "Karaoke Systems Market Report: Forecast 2034," 2024. [Online]. Available: <https://www.industryresearch.biz/market-reports/karaoke-systems-market-109083>
- [7] J. R. Quinlan, "Induction of Decision Trees," *Mach Learn*, vol. 1, no. 1, pp. 81–106, 1986, doi: 10.1007/BF00116251.
- [8] A. Alharbi, "Classification Performance Analysis of Decision Tree-Based Algorithms with Noisy Class Variable," *Discrete Dyn Nat Soc*, vol. 2024, pp. 1–15, Feb. 2024, doi: 10.1155/2024/6671395.
- [9] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. New York: CRC Press, 1984.
- [10] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers, 1993.
- [11] F. N. F. Q. Simoes, M. Dastani, and T. van Ommen, "Fundamental Properties of Causal Entropy and Information Gain," in *Proceedings of the Third Conference on Causal Learning and Reasoning*, in

- Proceedings of Machine Learning Research, vol. 236. PMLR, Apr. 2024, pp. 188–208. [Online]. Available: <https://proceedings.mlr.press/v236/simoes24a.html>
- [12] I. Kovalcuk and others, “ID3 Decision Tree Algorithm: Application in Sports Training Data Analysis,” *PMC Scientific Reports*, 2023, doi: 10.1038/s41598-023-xxxxx.
- [13] M. A. Qasimi, “Examining and Evaluating Classification Algorithms Based on Decision Trees,” *Journal of Electrical and Electronic Engineering*, vol. 3, no. 2, pp. 1–7, 2024.
- [14] I. Huda, A. A. Suhendra, and M. A. Bijaksana, “Design of Prediction Model using Data Mining for Segmentation and Classification Customer Churn in E-Commerce Mall in Mall,” *JOIV: International Journal on Informatics Visualization*, vol. 7, no. 4, pp. 2280–2289, Dec. 2023, doi: 10.62527/joiv.7.4.2414.
- [15] E. Hosseini and others, “The Evolutionary Convergent Algorithm: A Guiding Path of Neural Network Advancement,” *IEEE Access*, vol. 12, no. 45, pp. 127440–127459, 2024, doi: 10.1109/ACCESS.2024.xxxxxxx.
- [16] R. H. Sprague and E. D. Carlson, *Building Effective Decision Support Systems*. Englewood Cliffs, NJ: Prentice Hall, 1982.
- [17] N. Asgarova, M. Hussein, and others, “Decision Support Systems for Strategic Decision-Making,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 1, pp. 703–714, 2024.
- [18] S.-C. Necula and M. Mihailescu, “Exploring the Impact of Time Spent Reading Product Information on E-Commerce Websites: A Machine Learning Approach to Analyze Consumer Behavior,” *Behavioral Sciences*, vol. 13, no. 6, p. 439, May 2023, doi: 10.3390/bs13060439.