

IMPLEMENTASI DATA MINING PADA ANALISIS KARAKTERISTIK PELANGGAN : SYSTEMATIC LITERATURE REVIEW

Maria Marclyna Nau, Vita Nurul Fathya, Okky Pratama Martadireja

Manajemen Teknologi Keimigrasian, Politeknik Imigrasi

Jalan Satria Sudirman No. 1, Tangerang, Indonesia

marclynanauixa.151015@gmail.com

ABSTRAK

Pada masa ini, salah satu hal yang menjadi kunci dari transisi perkembangan zaman adalah pengolahan data. Data menjadi krusial karena berbagai pengaruh yang dimilikinya seperti pada prediksi, analisis, pengambilan keputusan, penetapan kebijakan, terciptanya inovasi, efisiensi, peningkatan mutu pelayanan, dan masih banyak lagi. Berangkat dari hal ini, salah satu metode pengolahan data yang umum digunakan adalah data mining yang bertujuan untuk menambang dan menggali informasi pada data di berbagai sektor baik pemerintahan, perusahaan, pendidikan, kesehatan, dan lain sebagainya. Tinjauan literatur sistematis ini akan mencari tau secara spesifik metode dan algoritma seperti apa yang digunakan dalam implementasi data mining pada analisis pelanggan. Pada penelitian ini, dilakukan tinjauan terhadap penelitian-penelitian terdahulu yang memiliki keterkaitan dengan implementasi data mining ini sendiri. Dalam tinjauan literatur sistematis ini akan dilakukan filterisasi pada referensi artikel terkait hingga menemukan referensi yang dapat menjadi acuan dalam membantu melakukan analisis pelanggan dengan metode dan algoritma yang tepat serta melihat seberapa efektif penggunaan data mining pada analisis karakteristik pelanggan. Hasil penelitian ini meninjau bahwasannya algoritma K-means adalah algoritma yang umum digunakan serta terbukti efektif dalam penggunaannya.

Kata kunci : Data mining, pelanggan potensial, literature review, karakteristik pelanggan, PRISMA

1. PENDAHULUAN

Teknologi informasi saat ini senantiasa mengalami perkembangan yang sangat pesat, transisi dari masa ke masa hingga saat ini menandakan bagaimana dinamika kehidupan manusia senantiasa bersinggungan dengan kemajuan teknologi. Manusia tengah berada pada era revolusi industri 4.0 yang mana otomatisasi, analisis data, serta kecerdasan buatan telah masuk kedalam kehidupan manusia [1].

Perkembangan ini membawa dampak signifikan terhadap bagaimana cara manusia berinteraksi dengan teknologi dan dunia digital. Salah satu akibat dari pesatnya perkembangan teknologi ini adalah pemanfaatan data dalam sistem informasi digital. Di era digital yang sarat akan perkembangan ini, jumlah data yang dihasilkan oleh berbagai bidang di dunia terbilang sangat banyak. Dalam segala bidang kehidupan di masa ini, berbagai data telah dihasilkan oleh teknologi informasi yang semakin futuristik. Berkembangnya teknologi informasi, mampu menghasilkan segudang data yang berisi informasi [2].

Data yang dihasilkan oleh aktivitas manusia dan sistem semakin besar baik dari segi volume, variasi, maupun kecepatan. Dengan demikian, diperlukan suatu metode yang efektif untuk mengekstraksi informasi dari data yang tersedia. Namun, walaupun sudah banyak penelitian terkait pemahaman dan penerapan data mining, masih belum ada pemahaman komprehensif mengenai tren, metode, dan aplikasi data mining yang paling efektif untuk digunakan.

Dalam sebuah instansi baik itu pemerintah ataupun swasta yang menyediakan pelayanan bagi masyarakat, perkembangan teknologi dan data juga menjadi salah satu instrumen penting dalam

menciptakan kepuasan masyarakat atau pelanggan dan juga pelayanan yang prima. Terutama dengan hadirnya internet yang dengan cepat dapat menyebarkan dan mendapatkan informasi. Hal ini menjadikan masyarakat atau pelanggan menjadi lebih kritis terhadap pelayanan yang mereka rasakan. Penggunaan teknologi informasi membuat perusahaan berkesempatan untuk meningkatkan efisiensi dan produktivitas, melalui sistem otomatisasi serta integrasi segala proses dalam bisnis mampu berjalan secara efektif dan efisien serta membantu perusahaan dalam memperluas jangkauan pasar [3].

Dari perkembangan ini perlu dilakukan data mining untuk melihat seperti apa karakteristik pelanggan dan pelayanan seperti apa yang cocok bagi mereka. Data mining merupakan proses dalam mengelola data dengan tujuan menemukan pola ataupun informasi tertentu. Data mining dapat didefinisikan sebagai sebuah kumpulan teknik-teknik yang bertujuan menemukan pola-pola tersembunyi pada data yang telah dikumpulkan.[4].

Hubungan antara penyedia layanan dan pelanggan sebagai penerima layanan merupakan hubungan dengan keterkaitan yang tinggi. Penting bagi penyedia pelayanan untuk memberikan layanan sebaik mungkin hingga mencapai suatu kepuasan pelanggan. Kepuasan pelanggan merupakan hasil dari suatu proses kinerja pada sebuah perusahaan yang dapat dirasakan oleh konsumen, dimana produknya sesuai dengan harapan konsumen itu sendiri [5].

Penelitian ini bertujuan untuk melakukan *Systematic Literature Review* (SLR) terkait implementasi data mining dalam menganalisis karakteristik pelanggan. Penelitian ini diharapkan

dapat memberikan kontribusi signifikan bagi akademisi dan praktisi dalam memahami perkembangan data mining serta optimalisasi data mining dalam analisis karakteristik pelanggan.

2. TINJAUAN PUSTAKA

Data mining ialah proses mengolah data demi menggali informasi yang penting dan pola tertentu yang tidak dapat dilihat secara langsung lewat data yang mentah. Konsep dasar data mining adalah menentukan informasi tak kasat mata pada basis data, dengan tujuan agar ditemukannya informasi dan pola yang dapat bermanfaat dalam data [6].

Proses data pada data mining merupakan serangkaian tahapan sistematis yang dimanfaatkan untuk mengekstrak informasi penting maupun pola yang tak terlihat dari kumpulan data yang besar dan juga kompleks [5].

Data mining merupakan salah satu teknik untuk menggali informasi penting yang terpendam atau tersembunyi dalam suatu kumpulan data (database) yang sangat besar hingga ditemukan pola menarik yang tidak diketahui sebelumnya [7].

Pada kenyataannya, data mining merupakan suatu fase menuju *Knowledge Discovery In Databases* (KDD) [8]. Penerapan data mining diharapkan mampu membantu percepatan pada pengambilan keputusan yang dapat memungkinkan perusahaan dalam mengelola informasi yang terdapat dalam data konsumen dan menjadi sebuah pengetahuan (knowledge) yang baru untuk mengambil keputusan yang diperlukan [9].

Pada data mining, terdapat beberapa jenis teknik, di antaranya: klasifikasi, regresi, pengelompokan, peringkasan, pemodelan dependensi, serta deteksi perubahan dan juga penyimpanan [10].

Klasifikasi merupakan tipe analisis data yang mampu membantu seseorang dalam menentukan kelas label dari sampel yang akan diklasifikasikan [11].

Proses klasifikasi sendiri meliputi pembelajaran dari data sampel dengan algoritma klasifikasi, diikuti dengan pembuatan estimasi ketepatan dari data sampel menggunakan aturan klasifikasi [12].

Pengelompokan merupakan tugas deskriptif dengan tujuan untuk mengidentifikasi kelompok-kelompok kategori yang terbatas dalam deskripsi data. Peringkasan sendiri melibatkan metode yang bertujuan untuk menemukan deskripsi singkat yang terdapat pada satu set data. Pemodelan dependensi merupakan proses untuk menemukan model lokal yang mendeskripsikan hubungan penting yang ada pada antar variabel maupun nilai fitur di dalam data. Deteksi perubahan serta penyimpangan ialah proses untuk menemukan perubahan yang substansial dalam data [10].

3. METODE PENELITIAN

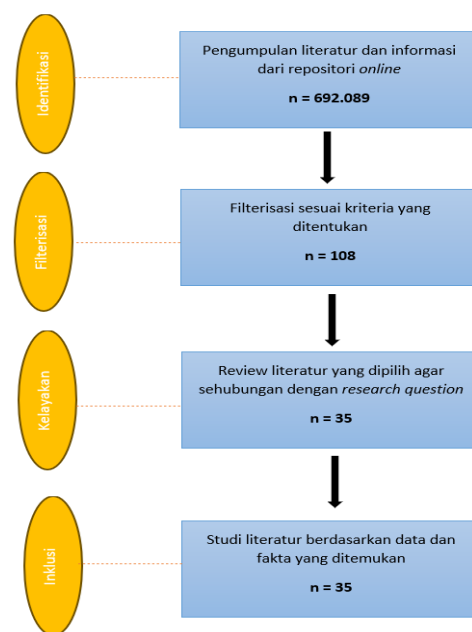
Penelitian systematic review ini dilakukan dengan menerapkan metode Preferred Reporting Items for Systematic Reviews and Meta-analyses

(PRISMA). Tahapan-tahapan dalam melakukan systematic review ini disusun dengan sedemikian rupa serta terencana dan terstruktur.

Adapun fase dari systematic review ini meliputi: Penyusunan Background and Purpose (Latar Belakang dan tujuan)

- Research Question
- Searching for the literature
- Selection Criteria
- Practical Screen
- Quality Checklist and Procedures
- Data Extraction Strategy
- Data Synthesis Strategy [13].

Proses ini sendiri terbagi menjadi 4 (empat) tahap, yakni: Identifikasi, Filterisasi, Kelayakan, dan Inklusi.



Gambar 1. Metode PRISMA

3.1. Tahap 1: Tolak Ukur Kelayakan Artikel

- IC1: Artikel merupakan penelitian asli yang telah dipelajari dan ditulis dalam bahasa Inggris atau Indonesia.
- IC2: Artikel diterbitkan antara tahun 2018 hingga 2025.
- IC3: Artikel bertujuan untuk mengimplementasikan data mining dengan analisis pelanggan

3.2. Tahap 2: Penetapan Sumber Literatur

- Literatur dicari pada basis data daring dengan repositori untuk studi akademis seperti Springopen, ScienceDirect, Elsevier dan Google Scholar.
- Pada jurnal atau yang memenuhi syarat untuk IC, dianalisis juga penelitian lain yang sehubungan dengan penelitian ini.

3.3. Tahap 3: Pemilihan Literatur

- Kata Kunci penentuan adalah “implementasi data mining” dan “analisis pelanggan”
- Untuk mengeksplorasi serta menentukan judul, abstrak serta artikel
- Kata kunci didapatkan dari hasil pencarian kelayakan kriteria yang sudah ditentukan sebelumnya.
- Baca artikel yang tidak dihilangkan dari tahap sebelumnya secara penuh maupun sebagian untuk menentukan apakah item tersebut memenuhi syarat untuk peninjauan berikutnya
- Artikel yang terpilih kemudian dinilai kembali untuk menemukan penelitian terkait

3.4. Tahap 4: Pengumpulan Data

Data dikumpulkan secara manual dengan membuat daftar yang berisi ekstraksi data. Pada penelitian ini telah difilter 242.874 artikel menurut kata kunci “Implementasi Data Mining” dan 449.215 artikel menurut kata kunci “Analisis Pelanggan” yang

berasal dari sumber dan kriteria, 108 jurnal ilmiah layak menjadi calon referensi menurut judul dan abstrak. Setelah dilakukan penelitian yang lebih lanjut, dengan membaca isi secara lebih mendetil hanya ditemukan 35 artikel terpilih yang memenuhi syarat untuk menjawab pertanyaan penelitian atau *research question* pada penelitian ini.

Tabel 1. Data repositori

Sumber	Implementasi Data Mining	Analisis Pelanggan	Kandidat	Terpilih
Spring Open	2.992	4.534	12	5
Science Direct	220.982	127.681	17	7
Google Scholar	18.900	317.000	79	23
Total	242.874	449.215	108	35

SLR ini memberikan jawaban atas beberapa Research Question seperti berikut:

Tabel 2. Pertanyaan penelitian

ID	Pertanyaan Penelitian	Latar Belakang
RQ1	Metode apa yang pernah diusulkan peneliti sebelumnya untuk melakukan data mining dalam analisis pelanggan?	Mengidentifikasi metode yang pernah digunakan oleh peneliti sebelumnya mengenai data mining pada analisis pelanggan
RQ2	Algoritma seperti apa yang sering dimanfaatkan untuk melakukan data mining?	Mengidentifikasi algoritma apa yang digunakan untuk melakukan data mining
RQ3	Bagaimana efektivitas implementasi data mining pada analisis karakteristik pelanggan?	Mengidentifikasi efektivitas implementasi data mining pada analisis karakteristik pelanggan

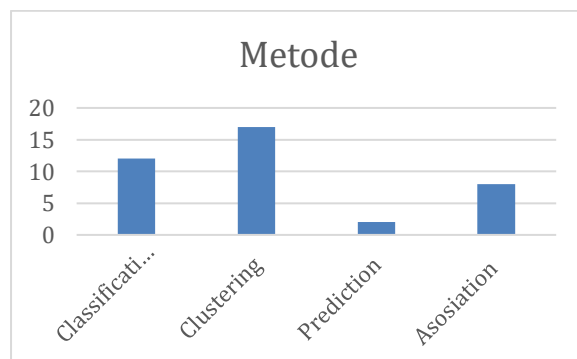
4. HASIL DAN PEMBAHASAN

Tabel 3. Daftar literatur review

No	Tahun	Judul	Metode	Algoritma	Tujuan Analisis
1.	2025	Discovering...[14]	Clustering, Classification	K-means, Neural Network, Decision Tree	Perilaku Pelanggan
2.	2022	Distributed...[15]	Clustering, Classification	Fuzzy C-Means	Perilaku Pelanggan
3.	2021	An Analytic...[6]	Clustering, Classification	K-means, C4.5	Tingkat Kepuasan Pelanggan
4.	2024	A Hybrid...[7]	Clustering	DBSCAN	Perilaku Pelanggan
5.	2024	An Approach...[8]	Clustering	K-means, Fuzzy C-means	Segmentasi Pelanggan
6.	2024	Predictive Model...[9]	Prediction	Random Forest, Gradient Boosting Machines, Deep Neural Network	Kepuasan Pelanggan
7.	2022	A Comprehensive[10]	Clustering, Asosiation	K-means, Apriori	Perilaku Pelanggan
8.	2024	Predictive Analytics...[11]	Classification	Decision Tree, Random Forest, Support Vector Machine	Perilaku Pelanggan
9.	2022	Implementasi data...[12]	Clustering	K-means	Pelanggan Potensial
10.	2022	Analisis...[16]	Classification	C4.5	Kepuasan Pelanggan
11.	2023	Segmentasi...[17]	Clustering	K-means	Pelanggan Potensial
12.	2024	Penerapan...[18]	Asosiation	Apriori	Pola Pembelian
13.	2021	Implementasi data...[19]	Clustering	K-means	Pelanggan Potensial
14.	2023	Penerapan Data...[20]	Asosiation	FP-Growth	Pola Pembelian
15.	2023	Penerapan Data...[21]	Classification	Linear Regression	Perilaku Pelanggan
16.	2021	Analisis Tingkat...[22]	Classification	C4.5	Tingkat Kepuasan
17.	2021	Penerapan Data...[23]	Clustering	K-means	Segmentasi Pelanggan
18.	2023	Penerapan Data...[24]	Asosiation	Apriori	Kebiasaan Pelanggan
19.	2021	Analisis Pola...[25]	Asosiation	Apriori	Transaksi Pelanggan

No	Tahun	Judul	Metode	Algoritma	Tujuan Analisis
20.	2024	Penerapan Data...[26]	Prediction	Linear Regression	Prediksi Pelanggan
21.	2021	Penerapan Data...[27]	Classification	Naive Bayes	Prediksi Pelanggan
22.	2021	Implementasi Data...[28]	Clustering	K-means	Tingkat Kepuasan Pelanggan
23.	2024	Pengembangan...[29]	Asosiation	Apriori	Pola Pembelian
24.	2022	Penggunaan Algoritma...[30]	Asosiation	Apriori	Pola Pembelian
25.	2021	The Use...[31]	Clustering	K-means	Segmentasi Pelanggan
26.	2023	Penerapan...[32]	Classification	C4.5	Kepuasan Pelanggan
27.	2024	Analisis Segmentasi...[33]	Clustering	K-means	Segmentasi Pelanggan
28.	2022	Analisis Segmentasi...[34]	Clustering	K-means	Segmentasi Pelanggan
29.	2023	Implementasi Data...[35]	Asosiation	FP-Growth	Pola Pembelian
30.	2023	Analisis Segmentasi...[36]	Clustering	K-means	Segmentasi Pelanggan
31.	2023	Analisis Kepuasan...[37]	Classification	C4.5	Kepuasan Pelanggan
32.	2025	Discovering...[38]	Clustering	K-means	Segmentasi Pelanggan
33.	2022	Analisa Efisiensi...[39]	Classification	Random Forest, Decision Tree	Segmentasi Pelanggan
34.	2025	Tailoring...[40]	Clustering	K-means	Segmentasi Pelanggan
35.	2025	Penggunaan...[41]	Classification	Decision Tree, C4.5	Perilaku Pelanggan

Dari data diatas, jawaban atas *research question* dapat ditentukan yakni metode penelitian seperti apa yang pernah digunakan oleh peneliti terdahulu dan menjadi referensi terkait implementasi data mining untuk analisis pelanggan. Beberapa metode tersebut adalah sebagai berikut.



Gambar 2. Grafik penggunaan metode data mining

Dari grafik di atas, dapat terlihat bahwa ada empat metode yang pernah digunakan oleh peneliti sebelumnya terkait dengan implementasi data mining untuk analisis pelanggan. Metode ini yakni classification, clustering, prediction, dan association. Dari hasil tinjauan terhadap artikel-artikel diatas ditemukan bahwa metode clustering adalah metode yang paling umum digunakan, disusul dengan metode classification, lalu association, dan yang paling sedikit penggunaannya yakni metode prediction seperti yang sangat jelas terlihat pada grafik.

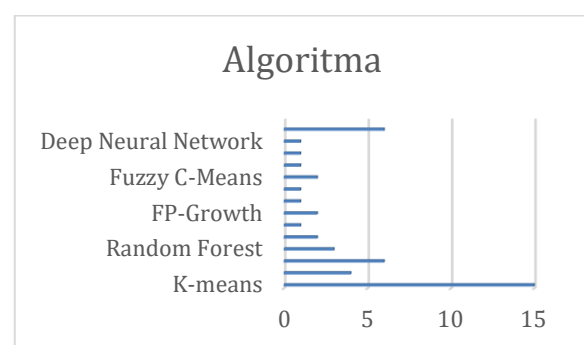
Peneliti terdahulu pada artikel [14], [15], [16], [17], [18], [20], [22], [24], [26], [30], [35], [38], [40],

[41], [43], [45], [47] menggunakan metode clustering pada penelitian mereka dikarenakan metode ini memudahkan untuk menemukan struktur data melalui kelompok-kelompok data itu sendiri. Hal ini sangat bermanfaat untuk analisis pelanggan karena dapat menemukan pola-pola tertentu dari pengelompokan pelanggan. Berikut adalah tabel terkait tujuan analisis pelanggan dengan menggunakan metode clustering.

Tabel 2. Tujuan Analisis

No	Tujuan Analisis	Artikel
1.	Perilaku Pelanggan	[14], [15], [17], [20]
2.	Kepuasan Pelanggan	[16], [35]
3.	Segmentasi Pelanggan	[18], [30], [38], [40], [41], [43], [45], [47]
4.	Pelanggan Potensial	[22], [24], [26]

Selanjutnya, menjawab pertanyaan dari *research question* lainnya, terkait algoritma apa yang relevan digunakan sesuai dengan penelitian terdahulu. Beberapa algoritma itu yakni:



Gambar 3. Algoritma yang digunakan

Berdasarkan grafik diatas algoritma K-means menjadi yang sangat populer digunakan pada penelitian-penelitian terdahulu dikarenakan berbagai alasan. Salah satunya karena algoritma ini memiliki kemudahan dalam pengimplementasiannya dibandingkan algoritma lain yang terdapat pada metode clustering seperti hierarchical dan DBSCAN. K-means juga sangat cocok digunakan untuk analisis pelanggan karena mampu dengan mudah mengelompokkan pelanggan berdasarkan kemiripan karakteristiknya. Berikut merupakan 3 alasan utama mengapa peneliti terdahulu banyak menggunakan algoritma ini.

- a. Efisiensi dan kesederhanaan dalam pengimplementasiannya. Algoritma K-means dapat membagi data dalam waktu yang relatif cepat meskipun jumlah data yang dikelola bersifat besar.
- b. K-means relevan digunakan untuk mengelola data kuantitatif seperti jumlah transaksi, usia, dan juga frekuensi sehingga sangat cocok digunakan dalam penelitian-penelitian tersebut
- c. Kemudahan evaluasi dan penerapannya dalam analisis pelanggan. Hasil clustering dengan algoritma K-means dapat dengan mudah dievaluasi dengan metode seperti Davis-Bouldien Index (DBI) untuk memastikan kualitas dari clustering yang telah dilakukan.

Hal ini menunjukkan bahwa pada analisis karakteristik pelanggan, metode yang sering kali digunakan adalah clustering atau pengelompokan daripada sekedar memprediksi perilaku pelanggan secara individual. Temuan ini sejalan dengan penelitian-penelitian diatas, contohnya pada penelitian [24] dan [47], yang menunjukkan bahwa metode clustering mampu menghasilkan segmentasi pelanggan yang lebih terstruktur dibandingkan classification, terutama dalam skenario di mana tidak ada label data eksplisit yang dapat digunakan. Namun, beberapa penelitian termasuk [14] menemukan bahwa dalam beberapa kasus, kombinasi clustering dengan classification dapat meningkatkan akurasi analisis pelanggan, misalnya dengan menggunakan hasil clustering sebagai input dalam model prediksi perilaku pelanggan. Implementasi data mining dalam analisis karakteristik pelanggan terbukti efektif, terutama melalui metode clustering yang memungkinkan pengelompokan pelanggan berdasarkan pola tertentu. Dari sudut pandang teoritis, penelitian ini menunjukkan bahwa clustering, khususnya dengan K-Means, tetap menjadi metode dominan dalam analisis pelanggan. Temuan ini mendukung teori bahwa clustering adalah metode yang sangat cocok untuk menganalisis data tidak berlabel, di mana tujuan utama adalah menemukan struktur tersembunyi dalam dataset. Secara praktis, temuan ini memiliki implikasi penting bagi dunia bisnis dan industri. Dengan mengetahui bahwa metode clustering (terutama K-Means) paling sering digunakan dalam analisis pelanggan, perusahaan

dapat mengoptimalkan strategi pemasaran dan layanan berbasis analisis perilaku pelanggan.

5. KESIMPULAN DAN SARAN

Penelitian ini menemukan bahwa metode clustering, khususnya dengan algoritma K-Means, adalah teknik data mining yang paling banyak digunakan dalam analisis pelanggan karena efisiensinya dalam mengelompokkan pola perilaku. Clustering terbukti efektif dalam segmentasi pelanggan di berbagai sektor, seperti e-commerce, perbankan, dan sebagainya. Namun, penelitian ini memiliki keterbatasan, seperti kurangnya evaluasi terhadap algoritma lain dan kombinasi metode yang lebih kompleks. Oleh karena itu, penelitian lanjutan disarankan untuk mengeksplorasi metode hybrid clustering-classification, membandingkan K-Means dengan algoritma lain seperti DBSCAN, serta menguji model pada dataset industri untuk validasi lebih lanjut.

DAFTAR PUSTAKA

- [1] A. Tahar, P. B. Setiadi, S. Rahayu, M. M. Stie, and M. Surabaya, "Strategi Pengembangan Sumber Daya Manusia dalam Menghadapi Era Revolusi Industri 4.0 Menuju Era Society 5.0," 2022.
- [2] O. Solihin, "IMPLEMENTASI BIG DATA PADA SOSIAL MEDIA SEBAGAI STRATEGI KOMUNIKASI KRISIS PEMERINTAH," 2021.
- [3] R. Anggraeni and I. E. Maulani, "PENGARUH TEKNOLOGI INFORMASI TERHADAP PERKEMBANGAN BISNIS MODERN," vol. 3, no. 2, 2023.
- [4] C. Zai, "IMPLEMENTASI DATA MINING SEBAGAI PENGOLAHAN DATA," 2022.
- [5] P. Chairy, "ANALISIS LOYALITAS PELANGGAN DITINJAU DARI FAKTOR KEPERCAYAAN DAN KEPUASAN PELANGGAN PADA BENGGEL SEPEDA MOTOR AHASS 2192," 2021.
- [6] M. Vahidi Farashah, A. Etebarian, R. Azmi, and R. Ebrahimzadeh Dastjerdi, "An analytics model for TelecoVAS customers' basket clustering using ensemble learning approach," *J Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00421-1.
- [7] F. Akhavan and E. Hassannayebi, "A hybrid machine learning with process analytics for predicting customer experience in online insurance services industry," *Decision Analytics Journal*, vol. 11, Jun. 2024, doi: 10.1016/j.dajour.2024.100452.
- [8] H. (Hojatollah) Hamidi and B. Haghi, "An approach based on data mining and genetic algorithm to optimizing time series clustering for efficient segmentation of customer behavior," *Computers in Human Behavior Reports*, vol. 16, Dec. 2024, doi: 10.1016/j.chbr.2024.100520.

- [9] Hoanh-Su Le, Thao-Vy Huynh Do, Minh Hoang Nguyen, Thanh-Thuy Thi Pham, Nhung Thi Nguye, and Van-Ho Nguyen, "Predictive model for customer satisfaction analytics in E-commerce sector using machine learning and deep learning," 2024.
- [10] K. Purna Prakash *et al.*, "A comprehensive analytical exploration and customer behaviour analysis of smart home energy consumption data with a practical case study," *Energy Reports*, vol. 8, pp. 9081–9093, Nov. 2022, doi: 10.1016/j.egy.2022.07.043.
- [11] H. GhorbanTanhaei, P. Boozary, S. Sheykhan, M. Rabiee, F. Rahmani, and I. Hosseini, "Predictive analytics in customer behavior: Anticipating trends and preferences," *Results in Control and Optimization*, vol. 17, p. 100462, Dec. 2024, doi: 10.1016/j.rico.2024.100462.
- [12] W. Romadhona, B. Indarmawan Nugroho, and A. Alim Murtopo, "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means," *Jurnal Minfo Polgan*, vol. 11, no. 2, pp. 100–104, Sep. 2022, doi: 10.33395/jmp.v11i2.11797.
- [13] N. I. Fitriyani, "Metode PRISMA Untuk Memprediksi Penyakit Kanker Payudara," 2021.
- [14] Y. Suh, "Discovering customer segments through interaction behaviors for home appliance business," *J Big Data*, vol. 12, no. 1, Dec. 2025, doi: 10.1186/s40537-025-01111-y.
- [15] A. W. Akram and Z. Alamgir, "Distributed fuzzy clustering algorithm for mixed-mode data in Apache SPARK," *J Big Data*, vol. 9, no. 1, Dec. 2022, doi: 10.1186/s40537-022-00671-7.
- [16] M. Adriansa, L. Yulianti, L. Elfianty, U. Dehasen Bengkulu, and J. Meranti Raya, "Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5," 2022.
- [17] Diotivano, I. Ruslianto, and D. Prawira, "SISTEM SEGMENTASI PELANGGAN POTENSIAL MENGGUNAKAN METODE K-MEANS DAN ANALISIS RFM," 2023.
- [18] H. D. Oktory and T. Y. Hadiwandura, "Penerapan Algoritma Apriori untuk Penentuan Pola Pembelian Kacamata pada Optik Indah Optik," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 4, pp. 1275–1281, Jul. 2024, doi: 10.57152/malcom.v4i4.1353.
- [19] Y. N. A. Cuhwanto and D. Agushinta, "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means," *PETIR*, vol. 15, no. 1, pp. 48–56, Dec. 2021, doi: 10.33322/petir.v15i1.1358.
- [20] Firmansyah and Odi Nurdiawan, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA FREQUENT PATTERN-GROWTH UNTUK MENENTUKAN POLA PEMBELIAN PRODUK CHEMICALS," 2023.
- [21] N. Riskya and S. Yuliana, "PENERAPAN DATA MINING UNTUK PREDIKSI PERILAKU PELANGGAN MENGGUNAKAN MULTIPLE LINEAR REGRESSION," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 11, no. 3, Aug. 2023, doi: 10.23960/jitet.v11i3.3194.
- [22] M. Sari, "ANALISIS TINGKAT KEPUASAN KONSUMEN RUMAH NUTRISI DI BANJARMASIN SEBAGAI IMPLEMENTASI PENERAPAN DATA MINING ALGORITMA C.45," 2021.
- [23] A. Pangestu and D. T. Ridwan, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS PENGELOMPOKAN PELANGGAN BERDASARKAN KUBIKASI AIR TERJUAL MENGGUNAKAN WEKA," 2021. [Online]. Available: <https://jurnal.umj.ac.id/index.php/just-it/index>
- [24] O. Pratama and J. Haerul Jaman, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA APRIORI UNTUK MENGETAHUI KEBIASAAN KONSUMEN DAN PREDIKSI STOK PRODUK (STUDI KASUS TOKO ELEKTRONIK WK)," 2023.
- [25] Styawati, Andi Nurkholis, and Krisma Nur Anjumi, "Analisis Pola Transaksi Pelanggan Menggunakan Algoritme Apriori," 2021.
- [26] Kiki Setiawan and Fauzi Ramdhani, "Penerapan Data Mining untuk Prediksi Pelanggan di PT. XYZ Menggunakan Algoritma Linear Regression," 2024.
- [27] F. Harahap, N. E. Saragih, E. T. Siregar, and H. Sariangisah, "Penerapan Data Mining Dengan Penerapan Data Mining Dengan Algoritma Naive Bayes Classifier Dalam Memprediksi Pembelian Cat," 2021.
- [28] S. Maria Sinaga, J. Tata Hardinata, and Fauzan M, "Implementasi Data Mining Clustering Tingkat Kepuasan Konsumen Terhadap Pelayanan Go-Jek," 2021.
- [29] A. R. Mangunsong, A. Rizky Simangunsong, V. Sihombing, and I. R. Munthe, "Pengembangan Sistem Rekomendasi Produk Berdasarkan Pola Pembelian dengan Pendekatan Algoritma Apriori," *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, vol. 7, no. 1, pp. 82–86, 2024.
- [30] N. Safitri and C. Bella, "PENGUNAAN ALGORITMA APRIORI DALAM PENERAPAN DATA MINING UNTUK ANALISIS POLA PEMBELIAN PELANGGAN (STUDI KASUS: TOKO DIENGVA BANDAR JAYA)," 2022.
- [31] A. A. Jamjoom, "The use of knowledge extraction in predicting customer churn in B2B," *J Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00500-3.
- [32] E. Jannah, V. Sihombing, and Masrizal, "Penerapan Data Mining Klasifikasi Kepuasan

- Pelanggan Transportasi Online Menggunakan Algoritma C4.5,” vol. 8, no. 1, 2023, [Online]. Available: http://ejournal.ust.ac.id/index.php/Jurnal_Means/
- [33] S. F. Djun, I. G. A. Gunadi, and S. Sariyasa, “Analisis Segmentasi Pelanggan pada Bisnis dengan Menggunakan Metode K-Means Clustering pada Model Data RFM,” *JTIM: Jurnal Teknologi Informasi dan Multimedia*, vol. 5, no. 4, pp. 354–364, Feb. 2024, doi: 10.35746/jtim.v5i4.434.
- [34] N. Huda Ahsina, F. Fatimah, F. Rachmawati, U. Ibn Khaldun Bogor JIKH Sholeh Iskandar Km, and K. Bogor, “ANALISIS SEGMENTASI PELANGGAN BANK BERDASARKAN PENGAMBILAN KREDIT DENGAN MENGGUNAKAN METODE K-MEANS CLUSTERING,” 2022.
- [35] S. Komariyah, S. Anwar, and B. Nurhakim, “Implementasi Data Mining FP-Growth Untuk Analisis Pola Pembelian Pada Transaksi Penjualan,” vol. 1, no. 2, 2023.
- [36] P. I. Pangestu, I. Hermanto, and D. Irmayanti, “ANALISIS SEGMENTASI PELANGGAN BERBASIS RECENCY FREQUENCY MONETARY (RFM) MENGGUNAKAN ALGORITMA K-MEANS,” 2023.
- [37] A. Surahmat and M. Sutrisno, “Analisis Kepuasan Pelanggan Dalam Industri Teknologi Menggunakan Algoritma C4.5,” 2023. [Online]. Available: <https://jurnal.umj.ac.id/index.php/just-it/index>
- [38] Y. Suh, “Discovering customer segments through interaction behaviors for home appliance business,” *J Big Data*, vol. 12, no. 1, p. 57, Mar. 2025, doi: 10.1186/s40537-025-01111-y.
- [39] Y. Christian, K. O. Yap, and R. Qi, “ANALISA EFISIENSI PENDEKATAN DATA-DRIVEN DALAM PROSES SEGMENTASI PASAR DENGAN STUDI KASUS STARTUP,” 2022, doi: 10.31949/infotech.v8i2.3732.
- [40] Q. Chen, B. J. Choi, and S. J. Lee, “Tailoring customer segmentation strategies for luxury brands in the NFT market – The case of SUPERGUCCI,” *Journal of Retailing and Consumer Services*, vol. 82, Jan. 2025, doi: 10.1016/j.jretconser.2024.104121.
- [41] Z. Nur Saputra and Zaehol Fatah, “PENGUNAAN DATA MINIG UNTUK MENGIDENTIFIKASI PELANGGAN BERESIKO TINGGI DALAM PENJUALAN MENGGUNAKAN ALGORITMA DECISION TREE C4.5,” *Jurnal Ilmiah Multidisiplin Ilmu*, vol. 2, no. 1, pp. 46–51, Feb. 2025, doi: 10.69714/s91z1k09.