

Literatur Review: Implementasi Metode dan Dataset pada Data Mining

Sabrina Fauzi Prameswari^{*1}, Nur Latifah Dwi Mutiara Sari²

^{1,2}Informatika, Universitas PGRI Semarang, Semarang

*Email: ¹sabrinafauzi44@gmail.com ²nurlatifah@upgris.ac.id

Abstract

Data mining is a rapidly growing field with wide applications across various sectors. Through an SLR approach, this research identifies, analyzes, and synthesizes relevant studies to provide a comprehensive overview of the techniques used and the datasets frequently applied. The research findings reveal the latest trends in algorithm usage, the advantages and disadvantages of applied algorithms, and the types of datasets most commonly used in data mining research. These findings are expected to guide researchers and practitioners in selecting appropriate methods, algorithms, and datasets for future data mining projects. The results of this literature review address the research questions posed, identifying eight algorithms used in the reviewed studies: Convolutional Neural Networks (CNN), K-Means, FP-Growth, K-Nearest Neighbors (KNN), Naive Bayes, Apriori Algorithm, Hierarchical, and C4.5. The most frequently used method/algorithm in the reviewed journals is K-Means, and the most commonly used dataset in the studied journals is the transaction dataset.

Keyword: *Data mining, Data mining methods, Dataset*

Abstrak

Data mining adalah bidang yang berkembang pesat dengan aplikasi yang luas di berbagai sektor. Melalui pendekatan SLR, penelitian ini mengidentifikasi, menganalisis, dan mensintesis studi-studi relevan untuk memberikan gambaran menyeluruh tentang teknik-teknik yang digunakan serta dataset yang sering diaplikasikan. Hasil penelitian mengungkap tren terbaru dalam penggunaan algoritma, kelebihan dan kekurangan algoritma yang diterapkan, serta jenis dataset yang paling umum digunakan dalam penelitian data mining. Temuan ini diharapkan dapat memberikan panduan kepada para peneliti dan praktisi dalam memilih metode atau algoritma dan dataset yang tepat untuk proyek-proyek data mining di masa depan. Hasil pada literature review ini adalah pembahasan dari research question yang dibuat, yaitu algoritma yang digunakan dalam penelitian yang direview ada 8 antara lain Convolutional Neural Networks (CNN), K-Means, FP-Growth, K-Nearest Neighbors (KNN), Naive Bayes, Algoritma Apriori, Hirarki, C4.5. metode/ algoritma yang sering digunakan dalam jurnal yang direview adalah k-means, serta data set yang sering digunakan pada jurnal yang diteliti adalah dataset transaksi.

Kata Kunci: *Data mining, Metode data mining, Dataset*

1. Pendahuluan

Di era informasi ini, perkembangan teknologi informasi yang pesat telah menghasilkan peningkatan eksponensial dalam jumlah data yang dihasilkan dan disimpan dari berbagai sumber seperti media sosial, transaksi e-commerce, sensor industri, dan rekam medis. Ledakan data ini membawa tantangan baru dalam menganalisis dan menginterpretasikan informasi yang terkandung di dalamnya, di mana data mining menjadi sangat penting. Data mining, atau penambangan data adalah proses untuk mengungkap pola, tren, dan pengetahuan tersembunyi dari kumpulan data besar melalui teknik analitis, matematika, dan komputasi. Salah satu kemajuan teknologi saat ini yang dapat digunakan untuk membedakan

jenis daun berdasarkan bentuk daun adalah data mining dan machine learning. Pembelajaran mesin merupakan salah satu teknologi terkini yang berkembang pesat di era di mana mesin diprogram untuk belajar sendiri melalui pengembangan algoritma [1].

Sejalan dengan itu, penambangan data adalah penemuan korelasi, pola, dan tren baru yang berguna dengan menambang kumpulan data besar[2]. Berbagai metode dan algoritma data mining telah dikembangkan, termasuk clustering, klasifikasi, regresi, dan asosiasi, masing-masing dengan kekuatan dan kelemahan unik. Clustering mengelompokkan data berdasarkan kesamaan, sementara klasifikasi mengkategorikan data ke dalam kelas yang telah ditentukan sebelumnya. Regresi memodelkan hubungan antar variabel untuk memprediksi nilai numerik, dan asosiasi mencari hubungan antara item dalam dataset. Pemilihan metode yang tepat sangat penting karena efektivitasnya bervariasi tergantung pada karakteristik data dan tujuan analisis.

Di era digital, bisnis dan teknologi terus berkembang pesat. Para pebisnis menghadapi berbagai tantangan ketika mengembangkan bisnisnya. Tantangan terbesar bagi pebisnis adalah ketidakmampuan mengelola data dalam jumlah besar yang dihasilkan dari seluruh transaksi nasabah[3]. Tinjauan literatur dalam data mining bertujuan mengkaji dan membandingkan berbagai metode yang telah diterapkan serta mengevaluasi efektivitasnya berdasarkan hasil dari berbagai dataset. Kami akan mengidentifikasi tren terkini, tantangan, dan peluang dalam pemilihan metode dan dataset, serta mengevaluasi bagaimana karakteristik dataset mempengaruhi hasil analisis. Dengan pemahaman mendalam tentang keterbatasan dan kelebihan dari berbagai metode dan dataset, kami berharap dapat mendorong pengembangan metode baru yang lebih efektif dan efisien serta penggunaan dataset yang lebih representatif dan relevan. Penelitian ini bertujuan memberikan panduan komprehensif bagi peneliti dan praktisi dalam memilih metode dan dataset yang paling sesuai untuk aplikasi spesifik mereka dalam data mining.

2. Metode

Penelitian ini menggunakan pendekatan kualitatif kajian literatur sistematis (systematic literature review). Kajian literatur sistematis dipilih karena memungkinkan untuk mengidentifikasi, mengevaluasi, dan menyintesis penelitian yang relevan terkait dengan implementasi metode dan dataset pada data mining. Pendekatan ini memastikan bahwa data yang dikumpulkan bersifat komprehensif, terkini, dan relevan. Langkah Langkah SLR:

1. Planning

Untuk memastikan bahwa SLR dilakukan dengan sistematis dan terstruktur, sejumlah langkah penting dilakukan selama tahap perencanaan. Ini termasuk:

- a. Identifikasi Tujuan dan Ruang Lingkup Penelitian: Menentukan tujuan utama SLR adalah untuk mengidentifikasi tren, masalah, dan peluang terbaru dalam memilih metode dan dataset untuk data mining.
- b. Pengembangan Protokol Review: Menciptakan protokol SLR yang komprehensif yang mencakup persyaratan inklusi dan eksklusi, sumber data, strategi pencarian, dan metode analisis.
- c. Formulasi Pertanyaan Penelitian (RQ): Membuat pertanyaan penelitian yang akan berfungsi sebagai pedoman utama selama proses peninjauan.

2. Conducting

Pada tahap ini mengumpulkan dan menganalisis literatur yang relevan adalah bagian dari tahap pelaksanaan:

- a. Pencarian Literatur: Gunakan database akademik seperti Google Scholar, IEEE Xplore, ACM Digital Library, dan ScienceDirect untuk melakukan pencarian literatur yang efektif. "Teknik data mining", "dataset data mining", dan "trend dalam data mining" adalah beberapa kata kunci yang digunakan.
- b. Studi dipilih berdasarkan kriteria inklusi dan eksklusi yang ditetapkan. Proses ini dilakukan dalam beberapa tahap, mulai dari pemilihan berdasarkan judul dan abstrak hingga pemilihan berdasarkan teks penuh.

- c. Ekstraksi Data: Mengumpulkan informasi penting dari studi yang terpilih. Ini termasuk teknik pengolahan data yang digunakan, jenis dataset, hasil penelitian, dan ulasan tentang keuntungan dan kekurangan.
 - d. Analisis dan Sintesis Data: Data yang telah diekstraksi dianalisis untuk menemukan pola dan tema yang muncul.
3. Redacting
- Tahap redacting atau Tahap penyusunan dan pelaporan hasil meliputi langkah-langkah berikut:
- a. Interpretasi Hasil: Menginterpretasikan temuan berdasarkan analisis yang telah dilakukan. Ini termasuk identifikasi tren, tantangan, dan peluang dalam penggunaan metode dan dataset data mining.
 - b. Pelaporan Hasil: Menyusun laporan SLR yang komprehensif, mencakup metodologi yang digunakan, temuan utama, diskusi, serta implikasi dari temuan tersebut. Laporan ini juga mencakup rekomendasi untuk penelitian lebih lanjut.
 - c. Validasi Temuan: Memastikan validitas dan reliabilitas temuan dengan melakukan review oleh rekan sejawat atau ahli dalam bidang terkait.

Pertanyaan penelitian ini atau Research Question (RQ) dibuat berdasarkan kebutuhan dari topik yang dipilih. Berikut ini adalah pertanyaan penelitian yang digunakan dalam penelitian ini:

RQ1: Jenis Metode/Algoritma Apa Yang Digunakan? Membahas tentang algoritma yang digunakan dalam 30 jurnal yang direview, ditemukan beberapa metode dan algoritma seperti CNN, naive bayes, k-means dan lain sebagainya.

RQ2: Algoritma Apa Yang Terbukti Terbaik dan Paling Sering Digunakan?Setelah dilakukan beberapa klasifikasi dan analisis tentang algoritma dan metode yang digunakan ditemukan beberapa algoritma dan metode yang sering digunakan didalam jurnal yang direview,

RQ3: Jenis Dataset Apa Yang Sering Digunakan Dalam Penelitian Data Mining?Mencari dan membuat rekapitulasi penggunaan data set yang digunakan untuk perancangan sistem yang ada didalam jurnal yang direview.

Daftar Jurnal yang digunakan untuk penelitian ini

No	Judul	Tahun	Penulis
1	Implementasi Algoritma Convolutional Neural Network (Cnn) Untuk Klasifikasi Daun Dengan Metode Data Mining Semma Menggunakan Keras[1]	2022	Yosia Amado Suwitono
2	Implementasi Metode Data Mining K-Means Clustering Terhadap Data Pembayaran Transaksi Menggunakan Bahasa Pemrograman Python Pada Cv Digital Dimensi[2]	2022	Dodi Alexsander Manalu Goldie Gunadi
3	Menganalisa Pola Pembelian Pelanggan Menggunakan Metode Data Mining: Customer Segementation Dan Market Basket Analysis Dengan Algoritme K-Means Dan Eclat[3]	2022	Dewi Kumala Sari, I Kadek Dwi Nuryana
4	Penerapan Data Mining Pola Penjualan Barang Pada Koperasi Dengan Menggunakan Metode Algoritma Fp-Growth[4]	2024	Destari Magdalena Br. Sitorus,
5	Penerapan Metode Data Mining Terhadap Data Transaksi Penjualan Menggunakan Algoritma Apriori[5]	2020	Adhitia Erfina Nusa Putra
6	Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa[6]	2024	Nirwana Hendrastuty
7	Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor (Studi Kasus: Rakuni Bakery, Pastry And Cake)[7]	2020	Sandy Aldy Pradana Mochamad Adhari Adiguna
8	Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal[8]	2020	Haris Kurniawan, Sarjon Defit, Sumijan

9	Penerapan Data Mining Untuk Clustering Penduduk Miskin Di Kota Tanjungbalai Menggunakan Metode Algoritma K-Means[9]	2024	Zunaida Sitorus, Suhartika Suhartika
10	Pemetaan Penduduk Penerima Bantuan Sosial Desa Waru Jaya Menggunakan Algoritma K-Means Clustering[10]	2023	Deni Bahtiar
11	Penerapan Algoritma K-Means Clustering Untuk Data Obat[11]	2023	Tis Asy Aria, Yuliadi Yuliadi
12	Penerapan Algoritma K-Means Pada Pengelompokan Data Set Bahan Pangan Indonesia Tahun 2022-2023[12]	2023	Alfian Adiyanto Yudhistira Arie Wijaya
13	Implementasi Data Mining Untuk Memprediksi Penyakit Diabetes Menggunakan Algoritma Naives Bayes Dan K-Nearest Neighbor[13]	2023	Fitrokh Nur Ikhromr Ipin Sugiyarto
14	Penerapan Algoritma K-Means Pada Dataset Pengangguran Terbuka Berdasarkan Pendidikan Di Provinsi Jawa Barat[14]	2024	Sintia Kusuma Arum, Rini Astuti
15	Implementation Of Data Mining To Predict Graduation Of Smk Al Huda Kedungwungu Students Using The Naïve Bayes Classifier Algorithm[15]	2023	Putriyana Putriyana
16	Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering Dengan Metode K-Means[16]	2023	Nur Afiasari Nana Suarna
17	Analisis Dataset Status Gizi Pada Balita Menggunakan Algoritma K-Means Clustering[17]	2023	Fina Raudotul Janah Rudi Kurniawan
18	Implementasi Fp-Growth Pada Data Tracer Study Universitas Wiraraja Untuk Menemukan Pola Asosiasi[18]	2024	Iddrus Iddrus, Dewi Wulansari, Johan Dharmawan
19	Implementasi Data Mining Metode Asosiasi Algoritma Fp-Growth Pada Perusahaan Ritel[19]	2020	Alexander Radityo Wibowo, Arief Jananto
20	Penerapan Data Mining Dengan Metode Klasifikasi Menggunakan Algoritma C4.5[20]	2022	Iyan Sunandar, Fauji Faisal Nugraha
21	Penerapan Data Mining Untuk Evaluasi Data Penjualan Menggunakan Metode Clustering Dan Algoritma Hirarki Divisive Di Perusahaan Media World Pekanbaru[21]	2019	Yuda Irawan
22	Penerapan Data Mining Denganmetode Algoritma Apriori Untukmenentukan Pola Pembelian Ikan[22]	2019	Saefudin, Septian
23	Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes[23]	2021	Fauzan Alghifari
24	Implementasi Data Mining Untuk Analisis Data Penjualan Dengan Menggunakan Algoritma Apriori (Studi Kasus Dapoerin's)[24]	2020	Arfhan Prasetyo, Ricki Sastra, Numan Musyaffa
25	Metode Asosiasi Data Mining Untuk Analisa Persediaan Fiber Optik Menggunakan Algoritma Apriori[25]	2020	Novalia Barkah,Entin Sutinah
26	Implementasi Data Mining Untuk Klasifikasi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor[26]	2019	Inna Alvi Nikmatun
27	Penerapan Data Mining Untuk Menentukan Pola Penjualan Sepatu Menggunakan Metode Algoritma Apriori[27]	2022	Syahriani
28	Aplikasi Algoritma Apriori Dalam Data Mining Penjualan Tanaman Hias[28]	2024	Hana Suci Aninda Geofani
29	Implementasi Algoritma Apriori Dalam Perencanaan Persediaan Alat Kesehatan Pada Apotek[29]	2022	Sabilla M, Lumbanbatu K, Prahmana I
30	Implementasi Metode Wrapper Sequential Feature Selection (WSFS) pada Dataset Stroke Menggunakan Metode Naïve Bayes Multinomial[30]	2023	Talidah N, Yulitania E

3. Hasil dan Pembahasan

A. RQ1: Metode atau algoritma yang digunakan pada penelitian

Selama lima tahun terakhir, berbagai algoritma dan metode telah digunakan dalam penelitian data mining untuk menemukan pola, tren, dan pengetahuan tersembunyi dari dataset yang besar. Beberapa algoritma yang paling sering disebutkan dalam literatur meliputi:

- a. Convolutional Neural Networks (CNN): Digunakan terutama untuk pengenalan pola dan analisis gambar.
- b. K-Means: Algoritma clustering yang populer yang digunakan untuk mengelompokkan data ke dalam kelompok yang memiliki banyak kesamaan.
- c. FP-Growth: Digunakan untuk mengidentifikasi pola asosiasi dalam data transaksi, terutama dalam analisis keranjang belanja.
- d. K-Nearest Neighbors (KNN): Algoritma klasifikasi yang sederhana tetapi berguna untuk banyak aplikasi.
- e. Naive Bayes: Algoritma klasifikasi yang berbasis probabilitas yang digunakan dalam berbagai aplikasi, seperti diagnosis medis dan penyaringan email spam.
- f. Algoritma Apriori: Mirip dengan FP-Growth, digunakan untuk menemukan aturan asosiatif dalam data transaksi. Ini menggunakan metode yang berbeda.
- g. Algoritma hirarki: Digunakan untuk mengelompokkan data ke dalam struktur bertingkat (hierarchical).
- h. C4.5: Algoritma pembelajaran mesin yang digunakan untuk membangun pohon keputusan.

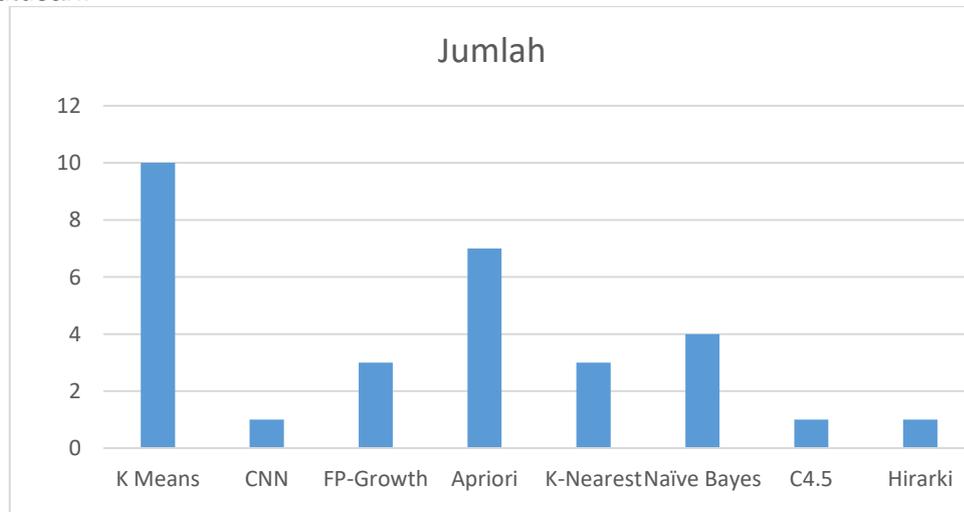


Diagram 1. Metode atau algoritma yang digunakan

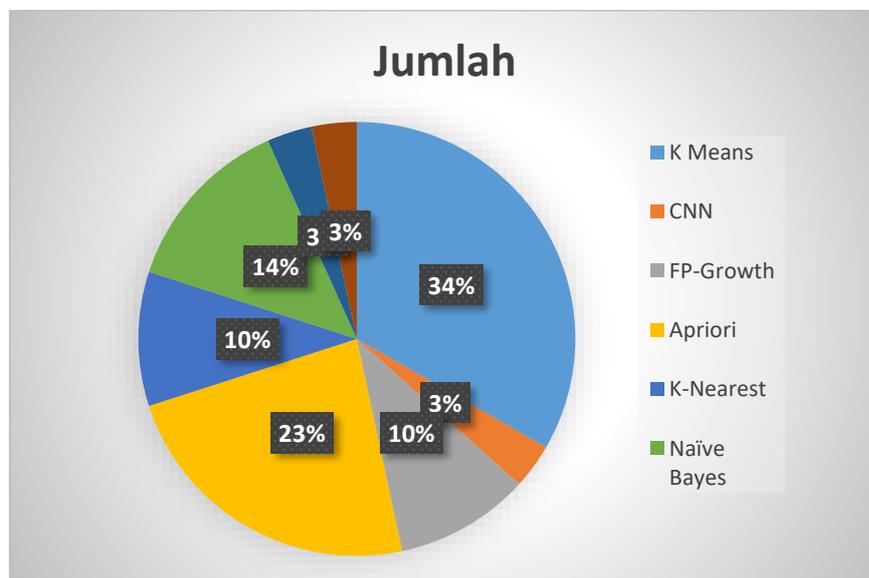


Diagram 2. Metode atau algoritma yang digunakan

Dari diagram diatas dapat diketahui bahwa persentase K-means sebesar 34%, Apriori 23%, Naïve Bayes 14%, FP-Growth dan K-Nearest 10%, sedangkan CNN, Hirarki, dan C4.5 sebanyak 3%.

B. RQ2: Metode/Algoritma terbaik atau yang paling sering digunakan

Beberapa alasan popularitas dan keberhasilan K-Means meliputi:

- Ketika berbagai metode data mining dievaluasi, K-Means muncul sebagai yang paling sering digunakan dan dianggap terbaik dalam banyak studi.
- Dari 30 jurnal yang direview k-means keluar sebanyak 10 kali dalam penelitian.
- Kesederhanaan dan Kecepatan: K-Means sangat mudah digunakan dan cepat dalam konvergensi, terutama untuk dataset yang besar.
- Kemampuan Clustering: K-Means sangat cocok untuk analisis pelanggan, segmentasi pasar, dan banyak aplikasi lainnya.
- Fleksibilitas: K-Means dapat digunakan untuk berbagai jenis data dan domain, termasuk data berbasis waktu dan numerik.

Namun, penting untuk diingat bahwa efektivitas metode bergantung pada tujuan analisis dan jenis dataset yang digunakan. Algoritma lain seperti CNN mungkin lebih cocok dalam beberapa situasi, terutama dalam analisis citra atau data yang kompleks.

C. RQ3: Dataset yang paling sering digunakan

Dalam literatur data mining, jenis dataset yang sering digunakan adalah Dataset Transaksi dan dataset pribadi. Dataset transaksi adalah data yang berasal dari transaksi, seperti data pembelian dalam e-commerce atau data penggunaan layanan. Dataset ini sering digunakan dalam analisis keranjang belanja, untuk menemukan pola pembelian pelanggan dan untuk melakukan segmentasi pasar. Dataset pribadi data yang mengandung informasi pribadi yang dapat digunakan untuk mengidentifikasi individu.

Penggunaan dataset numerik mencerminkan kebutuhan untuk menganalisis data yang berbeda dalam konteks yang berbeda. Dataset numerik biasanya lebih mudah dianalisis dengan algoritma seperti K-Means dan KNN, sementara dataset transaksi lebih cocok untuk algoritma asosiatif seperti FP-Growth dan Apriori.

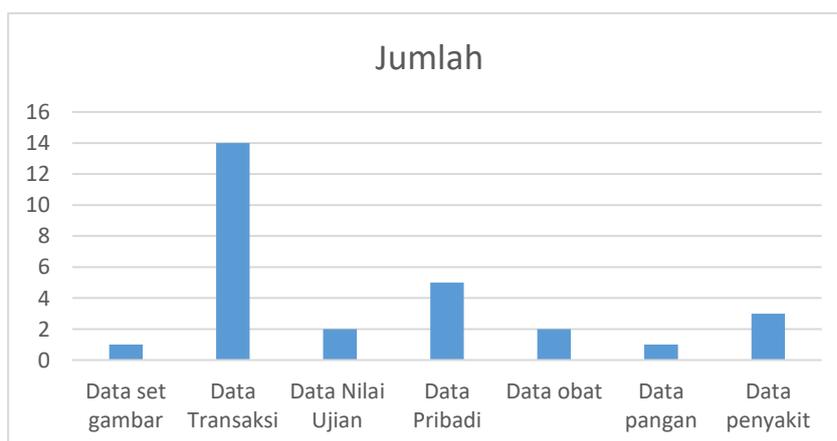


Diagram 3. Dataset yang digunakan yang digunakan

4. Kesimpulan

Beberapa alasan popularitas dan keberhasilan K-Means meliputi ketika berbagai metode data mining dievaluasi, K-Means muncul sebagai yang paling sering digunakan dan dianggap terbaik dalam banyak studi. Kesederhanaan dan Kecepatan K-Means sangat mudah digunakan dan cepat dalam konvergensi, terutama untuk dataset yang besar. Kemampuan Clustering K-Means sangat cocok untuk analisis pelanggan, segmentasi pasar, dan banyak aplikasi lainnya. Fleksibilitas K-Means dapat digunakan untuk berbagai jenis data dan domain, termasuk data berbasis waktu dan numerik. Namun, penting untuk diingat bahwa efektivitas metode bergantung pada tujuan analisis dan jenis dataset yang digunakan. Algoritma lain

seperti CNN mungkin lebih cocok dalam beberapa situasi, terutama dalam berbagai kasus yang melibatkan analisis data visual dan spasial, karena kemampuan mereka dalam menangkap fitur-fitur lokal dari data melalui operasi konvolusi. Dalam literatur data mining, jenis dataset yang sering digunakan adalah Dataset Transaksi dan dataset pribadi. Dataset transaksi adalah data yang berasal dari transaksi, seperti data pembelian dalam e-commerce atau data penggunaan layanan. Dataset pribadi data yang mengandung informasi pribadi yang dapat digunakan untuk mengidentifikasi individu.

5. Referensi

- [1] Y. A. Suwitono and F. J. Kaunang, "Implementasi Algoritma Convolutional Neural Network (CNN) Untuk Klasifikasi Daun Dengan Metode Data Mining SEMMA Menggunakan Keras," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 6, no. 2, pp. 109–121, Nov. 2022, doi: 10.31603/komtika.v6i2.8054.
- [2] D. A. Manalu and G. Gunadi, "IMPLEMENTASI METODE DATA MINING K-MEANS CLUSTERING TERHADAP DATA PEMBAYARAN TRANSAKSI MENGGUNAKAN BAHASA PEMROGRAMAN PYTHON PADA CV DIGITAL DIMENSI," *Infotech: Journal of Technology Information*, vol. 8, no. 1, pp. 43–54, Jun. 2022, doi: 10.37365/jti.v8i1.131.
- [3] D. K. Sari and I. K. D. Nuryana, "Menganalisa Pola Pembelian Pelanggan Menggunakan Metode Data Mining: Customer Segementation dan Market Basket Analysis Dengan Algoritme K-Means dan Eclat," *Journal of Emerging Information System and Business Intelligence (JEISBI)*, vol. 5, no. 2, pp. 73–82, 2024.
- [4] D. M. B. Sitorus, T. Syaputra, and M. Hutasuhut, "Penerapan Data Mining Pola Penjualan Barang Pada Koperasi Dengan Menggunakan Metode Algoritma FP-Growth," *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, vol. 3, no. 2, pp. 101–110, 2024.
- [5] A. Erfina, M. Melawati, and N. D. Arianti, "Penerapan Metode Data Mining Terhadap Data Transaksi Penjualan Menggunakan Algoritma Apriori (Studi Kasus: Toko Fasentro Fancy)," *SANTIKA is a scientific journal of science and technology*, vol. 10, no. 1, pp. 11–17, 2020.
- [6] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, vol. 3, no. 1, pp. 46–56, 2024.
- [7] S. A. Pradana and M. A. Adiguna, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-NEAREST NEIGHBOR (STUDI KASUS: RAKUNI BAKERY, PASTRY AND CAKE)," *OKTAL: Jurnal Ilmu Komputer dan Sains*, vol. 3, no. 06, pp. 1596–1610, 2024.
- [8] Haris Kurniawan, Sarjon Defit, and Sumijan, "Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal," *Journal of Applied Computer Science and Technology*, vol. 1, no. 2, pp. 80–89, Dec. 2020, doi: 10.52158/jacost.v1i2.102.
- [9] Z. Sitorus, "PENERAPAN DATA MINING UNTUK CLUSTERING PENDUDUK MISKIN DI KOTA TANJUNGBALAI MENGGUNAKAN METODE ALGORITMA K-MEANS," 2024. [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR>
- [10] D. Bahtiar *et al.*, "PEMETAAN PENDUDUK PENERIMA BANTUAN SOSIAL DESA WARU JAYA MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING," 2023. [Online]. Available: <http://pijarpemikiran.com/index.php/Scientia>

-
- [11] T. Asy Aria, M. Julkarnain, and F. Hamdani, "KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Algoritma K-Means Clustering Untuk Data Obat," *Media Online*, vol. 4, no. 1, pp. 649–657, 2023, doi: 10.30865/klik.v4i1.1117.
- [12] A. Adiyanto and Y. A. Wijaya, "PENERAPAN ALGORITMA K-MEANS PADA PENGELOMPOKAN DATA SET BAHAN PANGAN INDONESIA TAHUN 2022-2023," 2023. [Online]. Available: <https://hargapangan.id/>.
- [13] F. Nur Ikhromr, I. Sugiyarto, U. Faddillah, and B. Sudarsono, "IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI PENYAKIT DIABETES MENGGUNAKAN ALGORITMA NAIVES BAYES DAN K-NEAREST NEIGHBOR IMPLEMENTATION OF DATA MINING TO PREDICT DIABETES DISEASE USING NAIVES BAYES AND K-NEAREST NEIGHBOR ALGORITHMS," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 6, no. 1, 2023.
- [14] S. Kusuma Arum, R. Astuti, and F. Muhammad Basysyar, "PENERAPAN ALGORITMA K-MEANS PADA DATASET PENGANGGURAN TERBUKA BERDASARKAN PENDIDIKAN DI PROVINSI JAWA BARAT," 2024.
- [15] P. *1 and O. Nurdiawan, "Experimental Student Experiences Implementasi Data Mining Untuk Memprediksi Kelulusan Siswa SMK Al Huda Kedungwungu Dengan Menggunakan Algoritma Naïve Bayes Classifier Implementasi Data Mining Untuk Memprediksi Kelulusan Siswa SMK Al Huda Kedungwungu Dengan Menggunakan Algoritma Naïve Bayes Classifier," pp. 2985–3877, doi: 10.58330/ese.v1i4.202.
- [16] N. Afiasari, N. Suarna, and N. Rahaningsi, "Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means," *Jurnal SAINTEKOM*, vol. 13, no. 1, pp. 100–110, Mar. 2023, doi: 10.33020/saintekom.v13i1.402.
- [17] F. R. Janah, R. Kurniawan, and T. Suprpti, "ANALISIS DATASET STATUS GIZI PADA BALITA MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING," 2023.
- [18] I. Iddrus, D. Wulansari, and J. Dharmawan, "PENERAPAN DATA MINING DALAM Mencari Pola Asosiasi Data Tracer Study Menggunakan Algoritma FP-GROWTH (STUDY KASUS UPT. PUSAT LAYANAN KARIR UNIVERSITAS WIRARAJA)," *Jurnal Teknika*, vol. 16, no. 1, pp. 1–6, Mar. 2024, doi: 10.30736/jt.v16i1.1173.
- [19] A. R. Wibowo and A. Jananto, "IMPLEMENTASI DATA MINING METODE ASOSIASI ALGORITMA FP-GROWTH PADA PERUSAHAAN RITEL," vol. 10, no. 2, pp. 200–212, 2020.
- [20] F. Faisal Nugraha, I. Sunandar, and C. Juliane, "Penerapan Data Mining Dengan Metode Klasifikasi Menggunakan Algoritma C4.5," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 4, 2022, [Online]. Available: <http://jurnal.mdp.ac.id>
- [21] Yuda Irawan, "PENERAPAN DATA MINING UNTUK EVALUASI DATA PENJUALAN MENGGUNAKAN METODE CLUSTERING DAN ALGORITMA HIRARKI DIVISIVE DI PERUSAHAAN MEDIA WORLD PEKANBARU," *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM)*, vol. 4, no. 1, pp. 13–20, Apr. 2019, doi: 10.20527/jtiulm.v4i1.34.
- [22] S. Saefudin and S. DN, "Penerapan Data Mining Dengan Metode Algoritma Apriori Untuk Menentukan Pola Pembelian Ikan," *JSil (Jurnal Sist. Informasi)*, vol. 6, no. 2, p. 36, 2019.
- [23] F. Alghifari and D. Juardi, "PENERAPAN DATA MINING PADA PENJUALAN MAKANAN DAN MINUMAN MENGGUNAKAN METODE ALGORITMA NAÏVE

- BAYES,” *JURNAL ILMIAH INFORMATIKA*, vol. 9, no. 02, pp. 75–81, Sep. 2021, doi: 10.33884/jif.v9i02.3755.
- [24] A. Prasetyo, R. Sastra, and N. Musyaffa, “IMPLEMENTASI DATA MINING UNTUK ANALISIS DATA PENJUALAN DENGAN MENGGUNAKAN ALGORITMA APRIORI (STUDI KASUS DAPOERIN’S),” *Jurnal Khatulistiwa Informatika*, vol. 8, no. 2, Dec. 2020, doi: 10.31294/jki.v8i2.8994.
- [25] N. Barkah, E. Sutinah, and N. Agustina, “Metode Asosiasi Data Mining Untuk Analisa Persediaan Fiber Optik Menggunakan Algoritma Apriori,” *Jurnal Kajian Ilmiah*, vol. 20, no. 3, pp. 237–248, Sep. 2020, doi: 10.31599/jki.v20i3.288.
- [26] I. A. Nikmatun and I. Waspada, “Implementasi data mining untuk klasifikasi masa studi mahasiswa menggunakan algoritma K-Nearest Neighbor,” *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 10, no. 2, pp. 421–432, 2019.
- [27] S. Syahriani, “Penerapan Data Mining Untuk Menentukan Pola Penjualan Sepatu Menggunakan Metode Algoritma Apriori,” *BINA INSANI ICT JOURNAL*, vol. 9, no. 1, p. 43, Jun. 2022, doi: 10.51211/biict.v9i1.1758.
- [28] H. S. A. Geofani, D. G. Sumudra, F. Sufah, S. Alfareza, and J. Haerul, “Aplikasi Algoritma Apriori dalam Data Mining Penjualan Tanaman Hias”.
- [29] M. Y. Sabilla, K. Lumbanbatu, and I. G. Prahmana, “Implementasi Algoritma Apriori Dalam Perencanaan Persediaan Alat Kesehatan Pada Apotek,” *JTIK (Jurnal Teknik Informatika Kaputama)*, vol. 6, no. 2, pp. 434–447, 2022.
- [30] N. Talidah *et al.*, “SNESTIK Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika Implementasi Metode Wrapper Sequential Feature Selection (WSFS) pada Dataset Stroke Menggunakan Metode Naïve Bayes Multinomial”, doi: 10.31284/p.snestik.2023.4059.