

## Data Mining dan Big Data Dalam Dunia Industri: *Literature Review*

Sri Rahmawati<sup>1\*</sup>, Randy Andhika Putra<sup>2</sup>, Raihan Dara Lufika<sup>3</sup>

<sup>1,3</sup>Program Studi Teknik Industri, Fakultas Teknik, Universitas Syiah Kuala

<sup>2</sup>Program Studi Teknik Industri, Fakultas Teknik, Universitas Teuku Umar

\*Email Korespondensi: [sirahmawati@usk.ac.id](mailto:sirahmawati@usk.ac.id)

Abstrak - Perkembangan teknologi informasi telah membawa transformasi signifikan dalam dunia industri, dengan data mining dan big data menjadi dua konsep utama yang mendorong inovasi dan efisiensi. Penelitian ini menganalisis metode-metode data mining dan big data serta penerapannya di berbagai sektor industri melalui pendekatan Systematic Literature Review (SLR). Sebanyak 25 artikel jurnal relevan dikaji untuk mengidentifikasi tren, tantangan, dan dampak penerapan teknologi ini. Hasil penelitian menunjukkan bahwa data mining dan big data berperan penting dalam meningkatkan pengambilan keputusan berbasis data, efisiensi operasional, dan pengalaman pelanggan. Namun, tantangan seperti kualitas data, privasi, keamanan, dan kompetensi SDM masih menjadi hambatan utama. Penelitian ini juga menyoroti peluang pengembangan ke depan, termasuk integrasi dengan Internet of Things (IoT) dan machine learning. Temuan ini memberikan wawasan berharga bagi peneliti dan praktisi dalam mengoptimalkan penerapan data mining dan big data di lingkungan industri.

**Kata kunci:** data mining; big data; dunia industri

**Abstract** - The advancement of information technology has led to significant transformations in the industrial sector, with data mining and big data being two key concepts driving innovation and efficiency. This study analyzes data mining and big data methods and their applications across various industries using a Systematic Literature Review (SLR) approach. A total of 25 relevant journal articles were reviewed to identify trends, challenges, and impacts of these technologies.

The findings indicate that data mining and big data play a crucial role in improving data-driven decision-making, operational efficiency, and customer experience. However, challenges such as data quality, privacy, security, and human resource competence remain significant obstacles. This study also highlights future development opportunities, including integration with technologies such as the Internet of Things (IoT) and machine learning. These findings provide valuable insights for researchers and practitioners to optimize the implementation of data mining and big data in industrial environments.

**Keywords:** production machine; risk of damage; risk analysis

### PENDAHULUAN

Data mining merupakan hasil dari data yang berkembang pesat. Analisis data komputer telah digunakan sejak tahun 1960-an. Frasa "Data Mining" awalnya muncul ditahun 1990-an. Ini merupakan teknologi yang efektif untuk menambang informasi berharga atau mengekstraksi data tertentu dari berbagai kumpulan data besar. Ini disebut sebagai KDD, atau Penemuan Pengetahuan dalam Database. Basis data lintas aplikasi, statistik, pembelajaran mesin, dan kecerdasan buatan semuanya termasuk dalam bidang ilmu komputer multidisiplin (Sun, 2022). Data mining merupakan proses menemukan pola dan informasi berguna dari kumpulan data besar, sedangkan big data merujuk pada volume, kecepatan, dan variasi data yang sangat besar yang tidak dapat dikelola dengan metode tradisional. Perkembangan teknologi informasi telah membawa perubahan signifikan dalam dunia industri. Data mining dan big data menjadi dua konsep utama yang mendorong inovasi dan efisiensi di berbagai sektor industri. Data mining merujuk pada proses mengekstrak pengetahuan dari data dalam jumlah besar menggunakan algoritma statistik dan *machine learning*, sementara big data mencakup pengelolaan dan analisis data yang terlalu kompleks untuk diproses dengan alat tradisional.

Transformasi digital yang signifikan telah dipicu oleh revolusi industri keempat di sejumlah sektor industri, menjadikan data sebagai aset strategis yang bernilai tinggi. Untuk mendapatkan wawasan yang

dapat meningkatkan efisiensi operasional, merampingkan proses manufaktur, dan meningkatkan pengalaman pelanggan, bisnis kini berlomba-lomba untuk mendapatkan perhatian pelanggan dengan menerapkan big data dan teknologi penambangan data. Namun, proses memasukkan teknologi ini ke dalam proses perusahaan tidak mudah dan memerlukan pemahaman menyeluruh tentang hambatan implementasi dan janji teknologi. Tinjauan literatur ini memberikan ringkasan menyeluruh tentang penggunaan dan efek big data dan penambangan data di seluruh sektor ekonomi Indonesia.

Sejumlah penelitian telah menunjukkan bahwa bisnis dapat memperoleh keunggulan kompetitif yang substansial dengan menerapkan teknologi big data dan data mining. Dari sudut pandang manajerial, kemampuan analitik data memfasilitasi pengambilan keputusan berbasis bukti yang lebih tepat dan tepat waktu. Namun, secara teknis, perkembangan dalam platform analitik, infrastruktur komputer, dan algoritme pembelajaran mesin telah meningkatkan kapasitas organisasi untuk mengelola jumlah, kecepatan, dan keragaman data yang terus bertambah. Tren utama, praktik terbaik, dan kesenjangan penelitian dalam penerapan big data dan penambangan data di lingkungan industri Indonesia diidentifikasi dalam tinjauan literatur lingkungan ini. Teknologi big data dan data mining memiliki potensi manfaat yang sangat besar, tetapi ada sejumlah kendala yang harus diatasi sebelum dapat diperlakukan. Saat menilai kelayakan penerapan teknologi ini, faktor-faktor seperti kualitas data, privasi, keamanan, dan kebutuhan akan tenaga kerja dengan kemampuan analitis yang memadai sangat penting. Selain membahas keberhasilan implementasi, tinjauan literatur ini melihat variabel penting yang memengaruhi kemanjuran big data dan inisiatif penambangan data dalam konteks industri yang beragam. Ini juga menyoroti konsekuensi teoretis dan praktis bagi peneliti dan praktisi di bidang ini. Untuk mengetahui hal tersebut, maka akan dilakukan *literature review*. *Literature review* merupakan salah satu metode yang dapat digunakan untuk melakukan kegiatan penelitian dan sebagai pembuktian secara hierarki (*hierarchy of evidence*) yang popular (Setiawan dkk, 2020). *Literature review* menunjukkan bahwa teknik untuk melakukan pembuktian melalui pendekatan masalah tertentu secara proses ilmiah sehingga menghasilkan luaran berupa artikel yang dimaksudkan untuk melakukan pencarian kembali secara ilmiah atau dalam sebuah studi (Kurnia, 2021). Tujuan dari penelitian ini yaitu untuk menganalisis metode-metode data mining dan big data dalam dunia industri.

## METODE PENELITIAN

Penelitian ini menggunakan metode *Systematic Literature Review* yang dilakukan dalam lima tahapan. Penelitian ini seluruhnya menggunakan data sekunder dalam pelaksanaannya. Tahap pertama melakukan pencarian artikel, dimana artikel dicari menggunakan bantuan mesin pencarian artikel yaitu *Google Scholar*. Artikel dikumpulkan dari tahun 2018 hingga tahun 2025 dengan topik spesifik terkait data mining dan big data dalam dunia industri. Pencarian menggunakan kata kunci "Data Mining" dan kata kunci lainnya terkait data mining dan big data dalam dunia industri. Tahap kedua yaitu melakukan seleksi artikel, dimana tahap seleksi ini berguna untuk memisahkan artikel yang tidak terkait dengan data mining dan big data dalam dunia industri. Tahap ketiga adalah merangkum artikel, dimana artikel dirangkum berdasarkan metode yang digunakan dan hasil penelitian yang diperoleh pada setiap artikel. Tahap keempat yaitu pengelompokan artikel. Artikel dikelompokkan berdasarkan tahun terbit dan metode yang digunakan. Tahap terakhir adalah melakukan analisis artikel berdasarkan tahun terbit dan metode yang digunakan. Selain itu, metode risiko yang digunakan akan dibahas secara rinci dan arah pengembangannya dimasa depan sesuai dengan kemajuan ilmu pengetahuan dan teknologi.

## HASIL DAN PEMBAHASAN

Bagian ini membahas hasil dan diskusi tentang pengumpulan, pengolahan, dan analisis data. Selain itu, bagian ini juga membahas artikel yang dihasilkan dari temuan penyaringan. Tahap awal pengumpulan artikel didasarkan pada judul yaitu data mining dan big data dalam dunia industri. Setelah melalui tahap pengumpulan data, diperoleh 30 artikel yang telah dikumpulkan, namun yang sesuai topik penelitian yaitu sebanyak 25 artikel. Adapun hasil pengumpulan artikel ditampilkan pada Tabel 1.

---

Tabel 1. Klasifikasi artikel berdasarkan tahun penelitian, metode yang digunakan, dan hasil dari penelitian

No.	Pengarang	Metode	Hasil Penelitian
1	(Septa & Hoirul, 2022)	Literature sistematis ( <i>Systematic Literature Review</i> ) dengan pendekatan kualitatif.	Ditemukan 10 artikel yang membahas mengenai peran big data dalam sektor industri perdagangan, dengan 5 artikel berisi terkait peran big data salah satunya sebagai strategi marketing serta 5 artikel lain juga membahas tantangan dalam penerapan big data salah satunya kurangnya kompetensi SDM. Berdasarkan keseluruhan temuan artikel jurnal ilmiah bahwa big data memiliki berbagai peran dalam sektor industri perdagangan di Indonesia salah satunya sebagai alat strategi bagi perusahaan bidang perkantoran, namun dalam penerapannya memiliki tantangan paling besar dalam kompetensi SDM dalam mengelola big data.
2	(Eka Mayasari & Agussalim, Agussalim, 2023)	Pendekatan <i>literature review</i>	big data dan data analisis telah menjadi elemen penting dalam perusahaan modern dalam membantu perusahaan mengambil keputusan yang lebih tepat, meningkatkan efisiensi operasional perusahaan, dan mencapai keunggulan kompetitif. Namun, perusahaan juga harus memperhatikan masalah privasi dan keamanan data untuk memastikan bahwa penggunaan Big Data mereka sesuai dengan regulasi yang berlaku.
3	(Puspita et al., 2023)	Metode deskriptif dengan pendekatan kualitatif	Big data mampu membuat kerja pemerintah daerah sebagai regulator menjadi lebih efisien dan tepat sasaran pada bidang perpajakan. Big data mampu menyediakan data yang dibutuhkan pemerintah dalam merumuskan dan menjalankan kebijakan, seperti penagihan dan pengawasan pajak berbasis teknologi.
4	(Siregar, 2018)	<i>K-Means</i>	Data mining diterapkan menggunakan model proses <i>K-Means</i> yang menyediakan proses standar penggunaan data mining pada berbagai bidang digunakan dalam klasifikasi karena hasil metode ini mudah dipahami dan diinterpretasikan.
5	(Haryanto, 2024)	<i>Kagle</i> dan AI dengan <i>machine learning</i> dengan <i>phyton</i>	Hasil dari studi ini menunjukkan bahwa sekitar 68% pasien mengalami penyakit jantung, menurut analisis data <i>Science</i> dan <i>python</i> . Dari empat belas pasien baru dan tiga belas parameter yang telah diidentifikasi sebelumnya, analisis ini cukup akurat dalam memprediksi penyakit jantung.
6	(Marjo Sumantri, 2023)	Data algoritma <i>linear regression</i>	Data produksi dianggap telah mencapai target jika mesin mampu menghasilkan minimal 40 shot dalam satu jam kerja normal, dan minimal 25 shot pada hari Sabtu selama setengah hari. Sebagai hasilnya, produksi dianggap telah sesuai dengan standar yang ditetapkan jika perolehan shot mesin memenuhi kriteria ini. Kedua, apabila data produksi menunjukkan bahwa mesin tidak mencapai standar yang telah ditetapkan, maka produksi dianggap belum memenuhi target. Hal ini terjadi jika jumlah shot yang dihasilkan oleh mesin berada di bawah standar yang telah ditentukan.
7	(Nur & Hura, 2024)	Metode kualitatif dengan metode pengumpulan data menggunakan studi literatur	Hasil dari penelitian ini diharapkan dapat memberikan wawasan praktis dan rekomendasi strategis bagi perusahaan-perusahaan logistik dalam memaksimalkan penggunaan teknologi big data, sehingga meningkatkan daya saing dan efisiensi operasional di era digital.

8	(Windani, 2023)	Pendekatan penelitian kualitatif dengan integrative literature review	Penerapan pemolisian prediktif berbasis big data di era modern memiliki potensi besar dalam mencegah kejahatan dan meningkatkan keamanan masyarakat. Melalui analisis data yang komprehensif, kepolisian dapat memetakan pola kejahatan dan mengidentifikasi area rawan kejahatan secara lebih efektif. Namun, penelitian ini juga menyoroti tantangan yang dihadapi, seperti risiko stigmatisasi lingkungan dan masalah perlindungan data pribadi. Meskipun teknologi informasi memberikan alat yang kuat untuk analisis dan pengambilan keputusan, penting bagi lembaga kepolisian untuk menjaga keseimbangan antara efektivitas operasional dan perlindungan hak privasi individu.
9	(Agung Laksono, Imam Taufik, Erie Kresna Andana, Heri Aji Setiawan, 2022)	Metode kualitatif dengan pendekatan studi literatur ( <i>library research</i> ).	Analisis Big Data telah menjadi fokus utama dalam dunia bisnis, dengan munculnya berbagai tren yang signifikan. Penelitian mengidentifikasi penggunaan teknik analisis data canggih seperti <i>machine learning</i> dan data mining, serta integrasi sumber data yang beragam sebagai tren utama dalam analisis Big Data. Selain itu, terdapat peningkatan fokus pada analisis <i>real-time</i> yang memungkinkan perusahaan untuk meramalkan tren pasar, mendeteksi anomali, dan meningkatkan pengambilan keputusan berbasis data.
10	(Supriyanto et al., 2021)	Metode kualitatif dengan pendekatan studi kepustakaan ( <i>literature study</i> ).	Penggunaan big data dalam pembelajaran daring di Indonesia masih tergolong rendah, meskipun potensi dan manfaatnya sangat besar. Penelitian ini mengungkapkan bahwa big data dapat menyimpan dan mengintegrasikan berbagai jenis data yang diperlukan dalam proses pembelajaran, seperti data siswa, data pengajar, dan hasil evaluasi. Namun, terdapat beberapa kendala yang menghambat penerapan big data, termasuk ketidakmerataan kecepatan akses data di berbagai daerah di Indonesia.
11	(Tannady et al., 2024)	Metode data mining dengan algoritma Naive Bayes.	Analisis big data dari platform musik streaming Spotify menggunakan metode data mining <i>Naive Bayes</i> dapat memberikan wawasan yang mendalam mengenai preferensi musik pengguna. Dengan menganalisis data besar yang dihasilkan oleh pengguna, termasuk genre, durasi, popularitas lagu, dan karakteristik demografis, penelitian ini berhasil mengklasifikasikan preferensi musik pengguna dengan akurasi yang cukup baik. Meskipun akurasi yang diperoleh mencapai 42.40%, penelitian ini menekankan bahwa masih ada ruang untuk perbaikan agar model dapat mencapai tingkat akurasi yang lebih tinggi, idealnya di atas 70%. Penggunaan metode <i>Naive Bayes</i> memungkinkan Spotify untuk memberikan rekomendasi musik yang lebih personal dan relevan bagi pengguna, sehingga meningkatkan pengalaman mendengarkan musik. Penelitian ini juga menyoroti pentingnya eksplorasi lebih lanjut terhadap potensi data besar di platform streaming untuk mendukung strategi pemasaran dan pengembangan produk musik yang lebih efektif.
12	(Dio et al., 2023)	Analisis Data Mining menggunakan Association Rule Market Basket	Penerapan analisis Market Basket menggunakan algoritma <i>FP-Growth</i> pada data transaksi retail berhasil mengidentifikasi delapan aturan asosiasi yang signifikan, di mana setiap hubungan antar produk

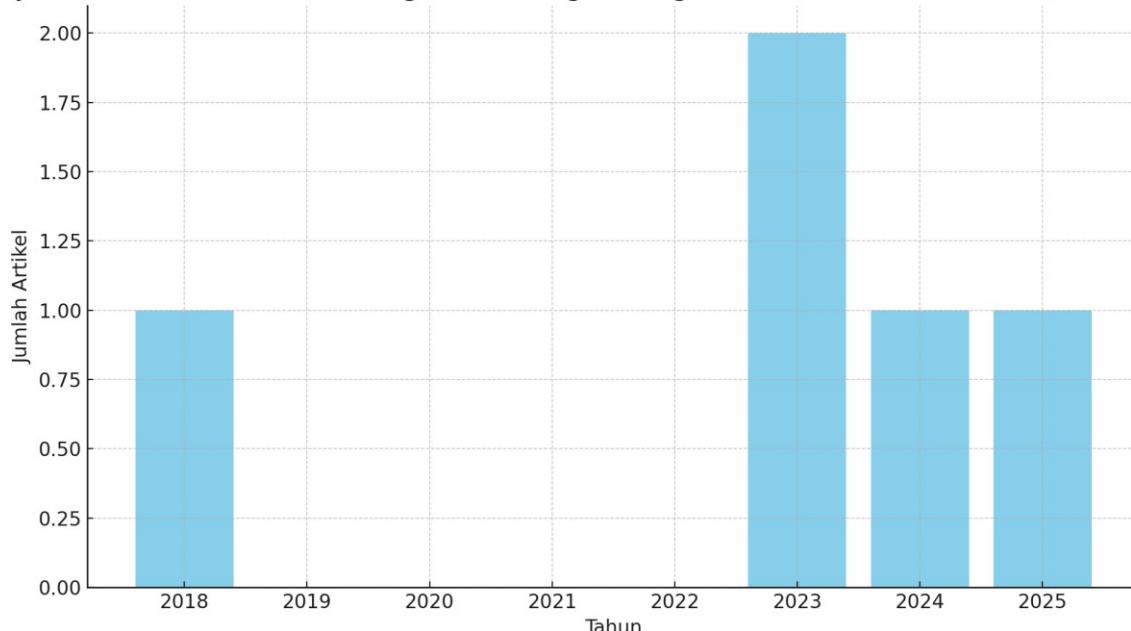
			<i>Analysis (AR-MBA) dengan algoritma FP-Growth.</i>	memiliki nilai lift ratio lebih dari 1 dan tingkat kepercayaan minimal 30%. Penelitian ini menggunakan 450 data transaksi dari sebuah retail di Kota Batam untuk menganalisis pola pembelian produk yang sering dibeli secara bersamaan oleh konsumen. Temuan ini memberikan wawasan berharga bagi pengelola retail dalam merancang tata letak fasilitas yang lebih efisien, dengan tujuan meningkatkan pengalaman berbelanja pelanggan.
13	(Funco Tanipu, Ni Putu Rai Yuliartini, Lalu Sibuan Amir, 2024)	Metode Systematic Literature Review (SLR).	Penerapan Big Data dalam sistem pendidikan di Indonesia memiliki dampak positif yang signifikan terhadap peningkatan inklusivitas dan aksesibilitas pendidikan. Melalui analisis literatur yang mencakup sumber-sumber yang terindeks di Scopus, DOAJ, dan Google Scholar dalam dekade terakhir, penelitian ini mengidentifikasi berbagai strategi efektif, seperti pengenalan kursus Data Science dan program perencanaan pendidikan nasional yang memanfaatkan data untuk memprediksi kinerja siswa.	
14	(Domaro & Dewayanto, 2023)	Metode Systematic Literature Review (SLR) dengan analisis artikel dari database Scopus.	Adopsi teknologi Big Data Analytics (BDA) memiliki dampak positif yang signifikan dalam mendeteksi fraud di bidang audit. Melalui analisis sistematis terhadap 20 artikel yang diterbitkan antara tahun 2019 hingga 2023, penelitian ini menemukan bahwa integrasi BDA dalam praktik audit tidak hanya meningkatkan efisiensi dan efektivitas proses audit, tetapi juga membantu auditor dalam mengidentifikasi dan mencegah tindakan kecurangan dengan lebih baik. Penelitian ini mengungkapkan bahwa faktor-faktor seperti inovasi teknologi, ukuran perusahaan, dan kualitas sumber daya manusia berperan penting dalam tingkat adopsi BDA oleh firma audit.	
15	(Bambang lareno, 2019)	Metode penelitian kualitatif dengan studi literatur yang berkaitan dengan teknik sipil dan data mining.	Penerapan algoritma data mining dalam bidang teknik sipil memiliki potensi yang signifikan untuk meningkatkan efisiensi dan efektivitas dalam perencanaan dan pelaksanaan proyek. Melalui analisis literatur yang dilakukan, penelitian ini mengidentifikasi berbagai aplikasi data mining, seperti optimasi desain struktural, analisis kelayakan investasi, dan prediksi kekuatan material.	
16	(Prameswari et al., 2024)	Metode Systematic Literature Review (SLR).	Data mining merupakan bidang yang berkembang pesat dengan aplikasi yang luas di berbagai sektor. Melalui pendekatan SLR, penelitian ini berhasil mengidentifikasi, menganalisis, dan mensintesis studi-studi relevan untuk memberikan gambaran menyeluruh tentang teknik-teknik yang digunakan serta dataset yang sering diaplikasikan dalam penelitian data mining. Temuan penelitian mengungkapkan tren terbaru dalam penggunaan algoritma, kelebihan dan kekurangan dari algoritma yang diterapkan, serta jenis dataset yang paling umum digunakan.	
17	(Isnaen et al., 2024)	Metode Systematic Literature Review (SLR).	Penerapan teknik data mining dalam sistem pengambilan keputusan pemasaran di perusahaan memiliki dampak yang signifikan terhadap efektivitas strategi pemasaran. Melalui analisis literatur yang sistematis, penelitian ini mengidentifikasi berbagai metode data mining yang sering digunakan, termasuk K-Means, Eclat Algorithm, dan DBSCAN, serta menilai efektivitas masing-masing metode dalam konteks	

			pemasaran. Penelitian ini menemukan bahwa data mining tidak hanya membantu dalam mengidentifikasi pola perilaku konsumen dan preferensi pasar, tetapi juga dalam meningkatkan akurasi prediksi dan pengambilan keputusan berbasis data. Selain itu, penelitian ini mengungkapkan tantangan yang dihadapi dalam persiapan data dan kualitas data yang diperlukan untuk analisis yang efektif. Dengan memanfaatkan dataset yang beragam, seperti data transaksi dan ulasan pelanggan, perusahaan dapat mengoptimalkan kampanye pemasaran mereka, mengurangi biaya akuisisi pelanggan, dan meningkatkan retensi pelanggan. Temuan ini memberikan wawasan berharga bagi praktisi pemasaran untuk mengadopsi teknik data mining dalam strategi mereka guna mencapai hasil yang lebih baik.
18	(Rohman et al., 2025)	Metode data mining dengan teknik <i>clustering</i> menggunakan algoritma <i>K-Means</i> .	Analisis pola prestasi siswa di Madrasah I'dadiyah Sukorejo menggunakan metode data mining, khususnya teknik <i>clustering</i> dengan algoritma <i>K-Means</i> , dapat memberikan wawasan yang signifikan mengenai kinerja akademik siswa. Penelitian ini mengelompokkan siswa ke dalam tiga kategori utama: siswa berprestasi tinggi, sedang, dan rendah, berdasarkan nilai ujian dan partisipasi dalam kegiatan ekstrakurikuler. Temuan menunjukkan bahwa siswa dengan prestasi tinggi memiliki keterlibatan aktif dalam kegiatan ekstrakurikuler, yang berkontribusi pada pengembangan keterampilan sosial dan motivasi belajar. Sementara itu, siswa berprestasi rendah diidentifikasi sebagai kelompok yang memerlukan perhatian lebih dalam bentuk bimbingan tambahan dan pelatihan motivasi.
19	(Nurhayati, 2019)	Metode survei dengan analisis literatur mengenai penggunaan Big Data dalam industri media dan hiburan.	Penggunaan Big Data dalam industri media dan hiburan telah membawa perubahan signifikan dalam cara perusahaan memahami dan melayani pelanggan. Melalui analisis literatur yang komprehensif, penelitian ini mengidentifikasi berbagai tantangan dan peluang yang dihadapi oleh industri dalam mengadopsi teknologi Big Data. Temuan menunjukkan bahwa Big Data memungkinkan perusahaan untuk menganalisis perilaku konsumen secara <i>real-time</i> , yang membantu dalam penyesuaian konten dan strategi pemasaran yang lebih efektif. Selain itu, penelitian ini menemukan bahwa perusahaan yang memanfaatkan Big Data dapat meningkatkan pengalaman pengguna dengan memberikan rekomendasi konten yang lebih relevan dan personal.
20	(Yanti, 2020)	Metode survei dengan analisis literatur mengenai penerapan Big Data dalam industri perjalanan dan pariwisata.	Penerapan Big Data dalam industri perjalanan dan pariwisata memiliki potensi yang sangat besar untuk meningkatkan efisiensi dan efektivitas operasional. Melalui analisis literatur yang dilakukan, penelitian ini mengidentifikasi berbagai aplikasi Big Data, termasuk dalam prediksi harga tiket, analisis permintaan, dan pengembangan sistem rekomendasi perjalanan yang personal dan adaptif. Temuan menunjukkan bahwa penggunaan Big Data memungkinkan perusahaan untuk memahami perilaku konsumen dengan lebih baik, sehingga dapat menawarkan produk dan layanan yang lebih sesuai dengan kebutuhan pelanggan.

			Selain itu, kombinasi antara Big Data dan teknologi <i>Internet of Things</i> (IoT) dapat membantu dalam penetapan harga yang lebih dinamis dan responsif terhadap perubahan permintaan pasar.
21	(Husni et al., 2022)	Metode <i>Exploratory Data Analysis</i> (EDA) menggunakan <i>Google Colaboratory</i> untuk menganalisis data penjualan video game.	Analisis Big Data penjualan video game menggunakan metode <i>Exploratory Data Analysis</i> (EDA) dapat memberikan wawasan yang mendalam mengenai tren penjualan di pasar global. Penelitian ini mengidentifikasi genre video game dengan penjualan tertinggi dan terendah, di mana genre <i>action</i> mencatat penjualan tertinggi sebesar 1745.27, sedangkan genre <i>strategy</i> mencatat penjualan terendah sebesar 174.5. Selain itu, analisis berdasarkan platform menunjukkan bahwa PS2 mendominasi penjualan dengan total 1255.64, sementara PCFX memiliki penjualan terendah sebesar 0.03. Dengan menggunakan visualisasi data, penelitian ini memudahkan perusahaan dalam menentukan game mana yang perlu dipromosikan dan diproduksi, serta memberikan informasi yang valid untuk pengambilan keputusan strategis. Temuan ini menegaskan pentingnya pemanfaatan Big Data dalam industri video game untuk meningkatkan efektivitas pemasaran dan pengembangan produk. Penelitian ini juga merekomendasikan agar perusahaan lebih fokus pada genre dan platform yang menunjukkan potensi penjualan yang lebih tinggi untuk memaksimalkan keuntungan.
22	(Hidayani et al., 2022)	Metode tinjauan literatur terstruktur dengan analisis terhadap 90 artikel jurnal internasional dari tahun 2012 hingga 2021.	Penerapan analitik Big Data (ABD) dalam rantai pasok memiliki dampak yang signifikan terhadap kemampuan perusahaan dalam melakukan analisis dan pengambilan keputusan. Melalui tinjauan literatur terstruktur, penelitian ini mengidentifikasi delapan faktor kunci yang mempengaruhi keberhasilan penerapan ABD, yaitu kerjasama antar divisi, transparansi di antara mitra rantai pasok, dukungan manajemen, pengembangan dan penyelarasan strategi, perubahan efisiensi operasi dan pemeliharaan, budaya pengambilan keputusan, dukungan dana, dan studi kelayakan untuk adopsi Big Data. Temuan menunjukkan bahwa kerjasama yang baik antar divisi dan transparansi informasi di antara mitra rantai pasok sangat penting untuk meningkatkan kepercayaan dan efisiensi operasional.
23	(Lubis Hayadi, 2022)	& Pendekatan kualitatif dan teknik pengumpulan data melalui studi literatur.	Pemerintah perlu mendefinisikan krisis dengan jelas untuk memahami karakteristiknya. Selain itu, pemanfaatan big data memungkinkan pemerintah untuk mempercepat pelaksanaan program dan mengidentifikasi masalah melalui analisis media sosial. Penelitian ini menekankan pentingnya pengumpulan umpan balik dari masyarakat untuk meningkatkan transparansi dan partisipasi dalam pengambilan keputusan. Dengan memanfaatkan big data, pemerintah dapat mengelola informasi secara lebih efektif, yang pada gilirannya dapat membantu dalam mengatasi krisis dan memperbaiki reputasi lembaga di mata publik.
24	(Unus, 2024)	Metode Deskriptif, Kualitatif,	Penerapan teknologi data mining berbasis metode <i>clustering</i> dapat secara signifikan meningkatkan

		<p>pengumpulan data melalui <i>desk study</i> dan analisis literatur.</p> <p>efisiensi pengolahan data di kalangan Usaha Mikro, Kecil, dan Menengah (UMKM) di Indonesia. Dengan menggunakan teknik <i>clustering</i>, pelaku UMKM dapat mengelompokkan data penjualan, pemasaran, dan produksi secara lebih terstruktur, yang memungkinkan mereka untuk mengelola informasi dengan lebih baik. Penelitian ini juga mengungkapkan bahwa meskipun banyak UMKM belum sepenuhnya memanfaatkan teknologi digital, penerapan sistem data mining dapat membantu mereka dalam menghadapi tantangan di era Revolusi Industri 4.0. Dengan demikian, teknologi ini diharapkan dapat menjadi ujung tombak dalam pengembangan UMKM, meningkatkan daya saing, dan mendukung pertumbuhan ekonomi nasional.</p>
25	(Ansori, 2018)	<p>Pengembangan sistem informasi berbasis <i>Internet of Things</i> (IoT). <i>Simple Moving Average</i>.</p> <p>Penerapan sistem informasi pengawasan bahan bakar minyak berbasis <i>Internet of Things</i> (IoT) dapat mengurangi biaya pengawasan hingga 80% per tahun untuk setiap kapal. Dengan menggunakan sensor yang terpasang pada tangki kapal dan metode Simple Moving Average untuk mengolah data ketinggian permukaan bahan bakar, sistem ini mampu memberikan informasi <i>real-time</i> yang akurat mengenai volume bahan bakar. Penelitian ini juga mengindikasikan bahwa investasi awal sebesar Rp 59.500.000 untuk pengembangan sistem dapat memberikan penghematan biaya operasional yang signifikan, dengan <i>payback</i> period hanya dalam 2,38 bulan. Dengan demikian, sistem ini tidak hanya meningkatkan efisiensi pengawasan, tetapi juga memberikan nilai tambah bagi perusahaan pelayaran dalam mengelola sumber daya mereka.</p>

Jumlah Terbitan Artikel tentang Data Mining dan Big Data dalam Dunia Industri (2018-2025)



Gambar 2. Terbitan artikel berdasarkan Tahun

Berdasarkan Gambar 2, data mining dan big data dalam dunia industri sedikit dilakukan, seperti pada tahun 2018 hingga tahun 2019, 2020, 2021, dan 2022 mengalami penurunan yang sangat signifikan, dan pada tahun 2023 mengalami kenaikan. selain itu, pada tahun 2024 hingga 2025 terhitung sama ditemukan artikelnya dikarenakan mulai menjadi salah satu perhatian bagi peneliti maupun industri atau perusahaan. Setelah diketahui beberapa metode atau pendekatan pada data mining dan big data dalam dunia industri, maka perlu dilakukan pembahasan atau rangkuman dari data mining dan big data dalam dunia industri yang diperoleh, antara lain sebagai berikut.

1. Literature sistematis (*Systematic Literature Review*) dengan pendekatan kualitatif merupakan metode penelitian yang sistematis dan terstruktur untuk mengumpulkan, menganalisis, dan mensintesis hasil penelitian yang relevan dengan topik yang diteliti (Septa & Hoirul, 2022). Sedangkan pendekatan kualitatif merupakan penelitian yang berfokus pada pemahaman mendalam tentang fenomena sosial, perilaku manusia, dan pengalaman subjektif.
2. Metode deskriptif digunakan untuk menggambarkan peran Big Data dalam sektor industri perdagangan di Indonesia, dengan fokus pada bagaimana Big Data mempengaruhi strategi pemasaran dan pengambilan keputusan. Sedangkan pendekatan data kualitatif adalah mengumpulkan data dari 10 artikel jurnal yang relevan, yang dianalisis secara kualitatif untuk memberikan gambaran yang mendalam tentang penerapan Big Data dalam konteks industri perdagangan (Puspita et al., 2023).
3. Metode *K-Means* dalam konteks penelitian tentang peran Big Data dalam sektor industri perdagangan, metode *K-Means* dapat digunakan untuk menganalisis data pelanggan, mengidentifikasi segmen pasar, dan memahami perilaku konsumen. Misalnya, perusahaan dapat menggunakan *K-Means* untuk mengelompokkan pelanggan berdasarkan pola pembelian mereka (Siregar, 2018).
4. *Linear regression* adalah metode statistik yang digunakan untuk memodelkan hubungan antara satu variabel dependen (target) dan satu atau lebih variabel independen (prediktor) dengan cara menemukan garis terbaik yang meminimalkan kesalahan prediksi. Dalam konteks penelitian tentang peran Big Data dalam sektor industri perdagangan, algoritma *linear regression* dapat digunakan untuk menganalisis hubungan antara variabel-variabel yang mempengaruhi penjualan atau kinerja bisnis. Misalnya, perusahaan dapat menggunakan *linear regression* untuk memprediksi penjualan berdasarkan faktor-faktor seperti harga, promosi, dan jumlah pelanggan (Marjo & Sumantri, 2023).
5. Algoritma Navo Bayes digunakan untuk mengklasifikasikan preferensi musik pengguna berdasarkan data besar (big data) seperti genre, durasi lagu, popularitas dan karakteristik demografis. Akurasi klasifikasi yang diperoleh mencapai 32.40%, dengan potensi peningkatan untuk mencapai tingkat akurasi yang lebih tinggi (ideal di atas 70%). Metode ini membantu Spotify dan memberikan rekomendasi musik yang lebih personal dan relevan, sekaligus mendukung strategi pemasaran dan pengembangan produk (Tannady et al., 2024).
6. IoT merupakan sistem menggunakan sensor yang dipasang pada tangki kapal untuk memantau ketinggian permukaan bahan bakar secara *real-time*, data dikirim melalui jaringan IoT ke pusat pengawasan. *Simple Moving Average* digunakan untuk mengolah data ketinggian bahan bakar dengan menghitung rata-rata bergerak dari sejumlah data terbaru (Ansori, 2018).
7. *Exploratory Data Analysis* (EDA) merupakan teknik analisis untuk mengeksplorasi dataset secara visual dan statistik tanpa model prediktif. Metode EDA digunakan untuk mengeksplorasi karakteristik dataset penjualan video game secara statistik dan visual (Husni et al., 2022).
8. *Association Rule Mining* (ARM) merupakan teknik *unsupervised learning* untuk menemukan hubungan tersembunyi antar *item* dalam dataset transaksional. Sedangkan *Market Basket Analysis* (MBA) adalah aplikasi spesifik ARM di sektor retail digunakan untuk menganalisis pola pembelian konsumen. Selain itu, Algoritma FP-Growth merupakan metode alternatif apriori yang mengatasi kelemahan komputasi dengan membangun *Frequent-Pattern Tree* (FP-Tree) dengan menggunakan pendekatan *divide-and-conquer* (Dio et al., 2023).

## KESIMPULAN

Penelitian ini mengkaji penerapan data mining dan big data dalam dunia industri melalui tinjauan literatur sistematis terhadap 25 artikel. Hasil analisis menunjukkan bahwa kedua teknologi ini telah menjadi tulang punggung inovasi dan efisiensi di berbagai sektor industri, mulai dari perdagangan, logistik, ingga pendidikan dan pemerintahan. Metode seperti *K-Means*, *linear regression*, *Naive Bayes*, dan *FP-Growth* terbukti efektif dalam mengekstrak pola dan wawasan dari data besar. Namun, implementasinya tidak lepas dari tantangan, termasuk masalah kualitas data, privasi, keamanan, serta kesenjangan

kompetensi SDM. Di sisi lain, integrasi dengan teknologi terkini seperti IoT dan machine learning membuka peluang besar untuk pengembangan di masa depan. Untuk mengoptimalkan manfaat data mining dan big data, diperlukan kolaborasi antara pemerintah, industri, dan akademisi dalam mengatasi hambatan dan meningkatkan adopsi teknologi ini. Penelitian ini diharapkan dapat menjadi referensi bagi peneliti dan praktisi dalam merancang strategi penerapan yang lebih efektif dan berkelanjutan.

## DAFTAR PUSTAKA

- Agung Laksono, Imam Taufik, Erie Kresna Andana, Heri Aji Setiawan, C. H. H. (2022). *TREN YANG MUNCUL DALAM ANALISIS BIG DATA : PELUANG UNTUK INTELIJEN BISNIS* Agung Laksono , Imam Taufik , Erie Kresna Andana , Heri Aji Setiawan , Catur Hendro dan kompleks ini untuk mendukung pengambilan keputusan bisnis yang lebih baik . telah berfokus pa. 1524–1535.
- Ansori, A. (2018). Studi Permanaan Internet of Things dan Data Mining untuk Pengawasan Bahan Bakar Minyak (Studi Kasus: Perusahaan Pelayaran Penumpang Nasional). *Wave: Jurnal Ilmiah Teknologi Maritim*, 12(1), 31–42. <https://doi.org/10.29122/jurnalwave.v12i1.2915>
- Bambang Iaren. (2019). Aplikasi Algoritma Data Mining pada Bidang Teknik Sipil. *Buletin Profesi Insinyur*, 2(2), 89–90. file:///C:/Users/ACER/Downloads/47-95-1-SM.pdf
- Dio, R., Hidayati, J., Arifin, R., Putera, D. A., & Dermawan, A. A. (2023). Analisis Data Mining Pembelian dengan Association Rule Market Basket Analysis menggunakan algoritma FP-Growth. *Jurnal PASTI (Penelitian Dan Aplikasi Sistem Dan Teknik Industri)*, 17(2), 187. <https://doi.org/10.22441/pasti.2023.v17i2.005>
- Domaro, A., & Dewayanto, S. T. (2023). Systematic Literature Review : Dampak Teknologi Big Data Analytics Dalam Mendeteksi Fraud Pada Bidang Audit. *Diponegoro Journal of Accounting*, 12(3), 1–9. <http://ejournal-s1.undip.ac.id/index.php/accounting>
- Eka Mayasari, & Agussalim Agussalim. (2023). Literature Review: Big Data dan Data Analys pada Perusahaan. *Jurnal Ilmiah Sistem Informasi Dan Ilmu Komputer*, 3(3), 171–187. <https://doi.org/10.55606/juisik.v3i3.680>
- Funco Tanipu, Ni Putu Rai Yuliartini, Lalu Sibuan Amir, K. R. A. (2024). The Role of Big Data in Optimizing the Education System in Indonesia: A Systematic Literature Review. *Journal of Telecommunications and the Digital Economy*, 12(1), 39–64. <https://doi.org/10.18080/jtde.v12n1.783>
- Haryanto, H. (2024). Analisis Big Data dan Artificial Intelligence (AI): dalam Industri Khususnya Prediksi Penyakit Jantung dengan Phyton. *Go Infotech: Jurnal Ilmiah STMIK AUB*, 30(1), 76–86. <https://doi.org/10.36309/goi.v30i1.262>
- Hidayani, N., Tri Wahyuni, E., & Kardina Prima Amrania, G. (2022). Peran Big Data Terhadap Kemampuan Perusahaan Melakukan Analisis Rantai Pasok : Studi Reviu Literatur Terstruktur. *Jurnal IlmiahMEA(Manajemen,Ekonomi,DanAkuntansi*, 6(2), 1513–1530.
- Husni, D. T., Sitompul, D. R. H., Sinurat, S. H., Ruben, R., Situmorang, A., Ziegel, D. J., Rahmad, J., & Indra, E. (2022). Analisis Big Data Penjualan Video Games Menggunakan Eda. *Jurnal Teknik Informasi Dan Komputer (Tekinkom)*, 5(1), 43. <https://doi.org/10.37600/tekinkom.v5i1.517>
- Isnaen, R., Zaeni, P., & Renaldy, R. (2024). *IN-FEST 2024 Systematic Literature Review : Penerapan dan Teknik Data Mining pada Sistem Pengambilan Keputusan Pemasaran di Perusahaan IN-FEST 2024*. 2, 578–584.
- Kurnia, H. (2021). A Systematic Literature Review of Performance Pyramids System Implementation in the Manufacture Industries. *Indonesian Journal of Industrial Engineering and Management (IJIEM)*, 2(2), 115–126.
- Lubis, M. Z., & Hayadi, B. H. (2022). Strategi Komunikasi Krisis Pemerintah Menggunakan Big Data Pada Media Sosial. *Journal of Computer and Engineering Science*, 1(3), 1–14.
- Marjo, & Sumantri, E. (2023). Implementasi Data Mining Dalam Prediksi Target Produksi Pada Proses Kerja Mesin Molding Menggunakan Algoritma Linear regression (Studi Kasus : Pt. Aim Karawang). *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 4(3), 1694–1703. <https://doi.org/10.35870/jimik.v4i3.397>
- Nur, A., & Hura, B. K. (2024). *Revolusi Logistik di Era Digital : Evaluasi Penggunaan Big Data di Industri Logistik*. 02(03), 443–453.
- Nurhayati, A. (2019). A Survey on Big Data in the Media and Entertainment Industry. *ITEJ (Information Technology Engineering Journals)*, 4(2), 75–88. <https://doi.org/10.24235/itej.v4i2.50>
- Prameswari, S. F., Latifah, N., & Mutiara, D. (2024). *IN-FEST 2024 Literatur Review : Implementasi Metode dan Dataset pada Data Mining IN-FEST 2024*. 2, 531–539.
- Puspita, A. F., Rusydi, M. K., & Purwanti, L. (2023). Peran Big Data dalam Membangun Smart City untuk Sistem Perpajakan Daerah. *Jurnal Riset Dan Aplikasi: Akuntansi Dan Manajemen*, 6(2), 171–184. <https://doi.org/10.33795/jraam.v6i2.004>

- Rohman, A., Toyibin, N., Fatah, Z., Sukorejo, U. I., Timur, S. J., Sukorejo, U. I., & Timur, S. J. (2025). ANALISIS DATA MINING MENGGUNAKAN METODE CLUSTERING TERHADAP PRESTASI SISWA I 'DADIYAH SUKOREJO. 2(1), 96–105.
- Septa, S., & Hoirul, H. (2022). Peran Big Data pada Sektor Industri Perdagangan: Tinjauan Literatur pada Perusahaan Bidang Perkantoran. *Journal of Office Administration : Education and Practice*, 2(3), 198–210. <https://doi.org/10.26740/joaep.v2n3.p198-210>
- Setiawan, I., & Purba, H. H. (2020). A Systematic Literature Review of Key Performance Indicators (KPIs) Implementation. *Journal of Industrial Engineering & Management Research*, 1(3), 200-208. <https://doi.org/10.7777/jiemar.v1i3.79>
- Siregar, M. H. (2018). Data Mining Klasterisasi Penjualan Alat-Alat Bangunan Menggunakan Metode K-Means (Studi Kasus Di Toko Adi Bangunan). *Jurnal Teknologi Dan Open Source*, 1(2), 83–91. <https://doi.org/10.36378/jtos.v1i2.24>
- Sun, W. (2022). Data Mining in the Big Data Era. *Proceedings of the 2022 8th International Conference on Humanities and Social Science Research (ICHSSR 2022)*, 664(Ichssr), 2107–2111. <https://doi.org/10.2991/assehr.k.220504.381>
- Supriyanto, E. E., Bakti, I. S., & Furqon, M. (2021). The Role of Big Data in the Implementation of Distance Learning. *Paedagoria: Jurnal Kajian, Penelitian Dan Pengembangan Pendidikan*, 6356(4), 61–68. <https://doi.org/10.31764>
- Tannady, H., Fernandes, J., Honni, A., & Lee, S. (2024). *Analisis Big Data Spotify dengan Metode Data Mining Analysis of Spotify Big Data with Data Mining Method*. 7(2), 52–59.
- Unus, E. G. (2024). Teknologi Data Mining Berbasis Metode Clustering Sebagai Ujung Tombak Perkembangan UMKM Di Indonesia Dalam Era Revolusi Industri 4.0. *Jurnal Reppositor*, 3(3), 315–320. <https://doi.org/10.22219/repositor.v3i3.31068>
- Windani, C. A. (2023). Strategi dan Tantangan Predictive Policing di Era Big Data bagi Masyarakat Modern. *Deviance Jurnal Kriminologi*, 7(2), 101. <https://doi.org/10.36080/djk.2385>
- Yanti, P. P. (2020). A Survey: Application of Big Data in the Travel and Tourism Industry. *ITEJ (Information Technology Engineering Journals)*, 5(1), 1–13. <https://doi.org/10.24235/itej.v5i1.38>