



Penerapan Algoritma XGBoost dalam Klasifikasi Jumlah Korban Kecelakaan Kereta Api di Indonesia

Selphia Nur Azzahra¹, Irwansyah^{2*}, Tupan Tri M³

^{1,2}Teknik Informatika, Fakultas Teknologi Industri dan Informatika, Universitas Muhammadiyah Prof. Dr. HAMKA, Jakarta, Indonesia

³Sistem Informasi, Fakultas Teknologi, Institut Sosial dan Teknologi (ISTEK) Widuri, Jakarta, Indonesia

*✉ Corresponding Author: irwansyah@uhamka.ac.id

ARTICLE INFO

Vol. 1, No. 2, April 2026, pp. 74-84.

DOI: <https://doi.org/10.66217/digintel-ai.v1i2.10>

Article history:

Revised: 13 April 2026

Accepted: 18 April 2026

Published: 30 April 2026

Abstract

This study aims to classify the number of vehicle accident casualties caused by railway accidents in Indonesia into low, medium, and high-risk categories using the XGBoost algorithm, as well as to evaluate the model performance based on accuracy, precision, and recall metrics. The employed methodology is CRISP-DM, consisting of stages such as business understanding, data understanding, data preparation, modeling, evaluation, and deployment stages. The dataset was obtained from official reports of the National Transportation Safety Committee (KNKT) and online news articles from 1991 to early 2025, resulting in 112 valid records after preprocessing, including data labeling, transformation of nominal attributes, and conversion of date data into numerical form. The classification process was carried out using RapidMiner. The results show that the XGBoost model achieved an accuracy of 88.39%, with the highest precision and recall values in the low-risk class (0.91 and 0.94) and high-risk class (0.88 and 0.87), while the performance for the medium-risk class remains relatively low (precision 0.75 and recall 0.68), indicating potential data imbalance or insufficient discriminative features. Based on these findings, it can be concluded that the XGBoost algorithm is effective in classifying railway accident risk levels; however, improvements in data quality and feature selection are still needed to achieve more optimal performance.

Keywords: Train accidents; Risk classification; XGBoost; Machine learning; CRISP-DM.

Abstrak

Penelitian ini bertujuan untuk mengklasifikasikan jumlah korban kecelakaan kendaraan yang disebabkan oleh kecelakaan kereta api di Indonesia ke dalam kategori risiko rendah, sedang, dan tinggi menggunakan algoritma XGBoost, serta mengevaluasi kinerja model berdasarkan metrik *accuracy*, *precision*, dan *recall*. Metodologi yang digunakan adalah CRISP-DM, yang terdiri dari tahapan-tahapan seperti *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Data penelitian diperoleh dari laporan resmi Komite Nasional Keselamatan Transportasi (KNKT) dan artikel media daring dari tahun 1991 hingga awal 2025, dengan total 112 data valid setelah melalui proses *preprocessing*, termasuk pelabelan data, transformasi atribut nominal, dan konversi data tanggal ke bentuk numerik. Proses klasifikasi dilakukan menggunakan RapidMiner. Hasil penelitian menunjukkan bahwa model *XGBoost* mampu mencapai tingkat akurasi sebesar 88,39%, dengan nilai *precision* dan *recall* tertinggi pada kelas risiko rendah



Copyright: © 2026 by the authors. Licensee DIGINTEL AI, Jakarta,

Indonesia. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0)

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

(0,91 dan 0,94) serta risiko tinggi (0,88 dan 0,87), sementara performa pada kelas risiko sedang masih relatif rendah (*precision* 0,75 dan *recall* 0,68), yang mengindikasikan adanya ketidakseimbangan data atau kurangnya fitur pembeda yang optimal. Berdasarkan hasil tersebut, dapat disimpulkan bahwa algoritma *XGBoost* efektif dalam mengklasifikasikan tingkat risiko kecelakaan kereta api, namun masih memerlukan peningkatan pada kualitas data dan fitur untuk memperoleh performa yang lebih optimal.

Kata Kunci: Kecelakaan kereta api; Klasifikasi risiko; *XGBoost*; *Machine learning*; CRISP-DM.

1. Pendahuluan

Transportasi memiliki peranan penting dalam menunjang mobilitas masyarakat sehari-hari. Di antara berbagai moda transportasi, kereta api menjadi salah satu pilihan yang banyak diminati karena efisiensi dan kapasitas angkutnya yang besar (Nafisah Nurul Hakim, 2020). Salah satu aspek krusial dalam sistem perkeretaapian adalah perlintasan sebidang, yaitu titik pertemuan antara jalur kereta api dan jalan umum yang berpotensi menimbulkan risiko kecelakaan. Berbagai insiden, seperti kendaraan yang melintasi rel saat kereta melintas atau aktivitas masyarakat di sekitar jalur rel, menjadi penyebab utama terjadinya kecelakaan yang mengakibatkan korban jiwa maupun kerugian material (Sinai et al., 2024). Oleh karena itu, keselamatan transportasi kereta api menjadi isu yang sangat penting dan memerlukan perhatian serius, terutama mengingat tingginya minat masyarakat Indonesia terhadap moda transportasi ini.

Berdasarkan data investigasi Komite Nasional Keselamatan Transportasi (KNKT), sebagian besar kecelakaan kereta api disebabkan oleh permasalahan prasarana, seperti kerusakan atau patahnya rel, yang berdampak signifikan terhadap operasional dan keselamatan (X. Liu et al., 2018). Meskipun PT Kereta Api Indonesia (KAI) sebagai operator utama terus berupaya meningkatkan kualitas layanan dan keselamatan, kejadian kecelakaan masih sering terjadi dan memerlukan pendekatan yang lebih sistematis berbasis data untuk mitigasi risiko (Permatasari et al., 2024). Selain itu, peningkatan volume perjalanan kereta api yang seiring dengan pertumbuhan jumlah penduduk dan kebutuhan mobilitas turut berkontribusi terhadap meningkatnya potensi risiko kecelakaan di perlintasan sebidang. Kondisi ini diperparah oleh masih banyaknya perlintasan tanpa pengawasan serta rendahnya tingkat kedisiplinan pengguna jalan, sehingga diperlukan pendekatan yang tidak hanya bersifat reaktif, tetapi juga prediktif dan preventif.

Seiring dengan perkembangan teknologi, pemanfaatan *Machine Learning* telah banyak digunakan dalam berbagai bidang untuk menganalisis data dan mendukung pengambilan keputusan, termasuk dalam sektor transportasi. Penelitian terbaru oleh Senkondo et al., (2025) mengimplementasikan algoritma *Random Forest*, *XGBoost*, dan *Transformer-CNN* untuk memprediksi kecelakaan pada perlintasan kereta api, dengan hasil menunjukkan bahwa model berbasis machine learning mampu mengungguli metode statistik tradisional serta menangkap hubungan *non-linear* antar variabel. Selain itu, Wang et al., (2025) mengembangkan pendekatan *Explainable Artificial Intelligence* (XAI) dengan menggabungkan *XGBoost* dan SHAP untuk menganalisis tingkat keparahan kecelakaan, yang berhasil mengidentifikasi faktor utama seperti kecepatan kendaraan, kondisi jalan, dan pencahayaan sebagai penentu risiko. Penelitian lain juga menunjukkan bahwa algoritma *XGBoost* memiliki performa unggul dibandingkan metode *machine learning* lainnya dengan nilai AUC sekitar 89% dalam klasifikasi kecelakaan kereta api (Bridgelall & Tolliver, 2021).

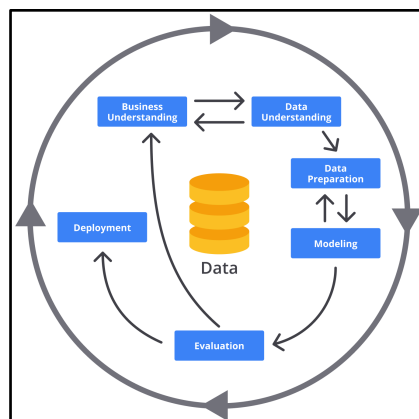
Meskipun demikian, sebagian besar penelitian masih berfokus pada prediksi kejadian kecelakaan atau analisis tingkat keparahan secara umum, dan belum banyak yang secara spesifik mengklasifikasikan jumlah korban kecelakaan ke dalam kategori tingkat risiko (J. Liu et al., 2025). Selain itu, penelitian yang mengintegrasikan data historis jangka panjang serta melakukan evaluasi model secara komprehensif menggunakan metrik klasifikasi seperti *accuracy*, *precision*, dan *recall* masih terbatas, khususnya dalam konteks kecelakaan kereta api di Indonesia (Zhou et al., 2020).

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengklasifikasikan jumlah korban kecelakaan kendaraan yang disebabkan oleh kereta api di Indonesia ke dalam kategori risiko rendah, sedang, dan tinggi menggunakan algoritma *XGBoost*, serta mengevaluasi kinerja model berdasarkan metrik *accuracy*, *precision*, dan *recall*. Data yang digunakan merupakan data historis kecelakaan kereta api di Indonesia dari tahun 1991 hingga awal 2025 yang diperoleh dari KNKT dan sumber berita daring. Dengan pendekatan berbasis CRISP-DM dan implementasi menggunakan *RapidMiner*, penelitian ini menghasilkan model klasifikasi yang akurat dan andal. Dengan demikian, penelitian ini berkontribusi pada pengembangan sistem klasifikasi risiko berbasis data sebagai dasar pengambilan keputusan dalam upaya mitigasi dan peningkatan keselamatan transportasi kereta api di Indonesia.

2. Metodologi

2.1 Metodologi *CRISP-DM*

Dalam penelitian ini digunakan pendekatan *Cross-Industry Standard Process for Data Mining* (CRISP-DM), yang merupakan kerangka kerja standar dalam proses data mining (Budyantara et al., 2020). Metodologi ini terdiri dari enam tahapan utama, yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Pendekatan CRISP-DM dipilih karena mampu memberikan alur kerja yang sistematis dan terstruktur dalam pengolahan data, mulai dari pemahaman masalah hingga implementasi model, sehingga sesuai untuk mendukung proses klasifikasi jumlah korban kecelakaan kereta api dalam penelitian ini. Tahapan CRISP-DM dalam penelitian ini ditunjukkan pada Gambar 1.



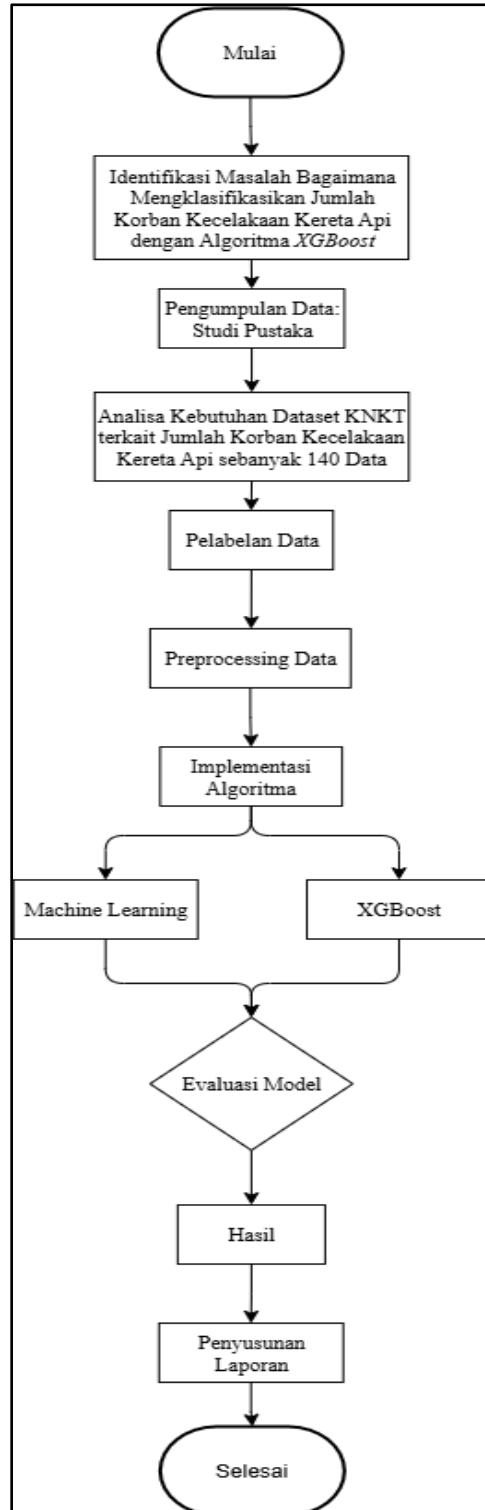
Gambar 1. Tahapan CRISP-DM
(Sumber: <https://www.dicoding.com>, 2025)

Tahapan dalam metodologi CRISP-DM pada penelitian ini dimulai dari *business understanding*, yaitu mengidentifikasi permasalahan kecelakaan kereta api serta pentingnya klasifikasi tingkat risiko korban sebagai dasar analisis. Selanjutnya, pada tahap *data understanding* dilakukan eksplorasi dan pemahaman terhadap data kecelakaan yang diperoleh dari KNKT dan sumber berita, termasuk identifikasi variabel dan pola awal. Tahap *data preparation* meliputi proses filtering data, pelabelan kategori risiko, serta *preprocessing* yang mencakup pembersihan data, transformasi, dan seleksi fitur (Novryadi et al., 2026). Pada tahap modeling, algoritma *XGBoost* diterapkan untuk melakukan klasifikasi multikelas dengan bantuan perangkat lunak *RapidMiner*. Kemudian, pada tahap *evaluation*, kinerja model dievaluasi menggunakan *confusion matrix* dengan metrik *accuracy*, *precision*, dan *recall* (Kurniawan et al., 2026). Tahap terakhir, yaitu *deployment*, dilakukan dengan

menginterpretasikan hasil model dan menyajikannya dalam bentuk laporan sebagai dasar pendukung pengambilan keputusan (Bagaskara et al., 2024).

2.2 Tahapan Penelitian

Proses pelaksanaan penelitian secara keseluruhan disajikan dalam Gambar 2.



Gambar 2. Tahapan Penelitian

Penelitian ini diawali dengan analisis masalah menggunakan pendekatan observasi non-partisipatif terhadap data kecelakaan kereta api di Indonesia untuk mengidentifikasi pola risiko dan variabel yang relevan. Data dikumpulkan dari sumber terpercaya, yaitu laporan resmi Komite Nasional Keselamatan Transportasi (KNKT) dan artikel berita daring, yang mencakup informasi jenis kecelakaan, lokasi, waktu, penyebab, serta jumlah korban. Selanjutnya, dilakukan proses filtering untuk memilih data yang relevan dalam rentang waktu 1991 hingga awal 2025, sehingga diperoleh dataset yang valid dan representatif. Tahap berikutnya adalah pelabelan data untuk keperluan *supervised learning*, di mana jumlah korban kecelakaan diklasifikasikan ke dalam tiga kategori risiko, yaitu rendah (0 korban), sedang (1–5 korban), dan tinggi (>5 korban). Data kemudian melalui tahap preprocessing yang meliputi pembersihan data (missing values dan duplikasi), transformasi data kategorikal ke numerik (*encoding*), serta seleksi fitur yang relevan (Putri et al., 2026). Dataset selanjutnya dibagi menjadi data latih (80%) dan data uji (20%) untuk keperluan pemodelan dan evaluasi (Irwansyah et al., 2024). Implementasi model dilakukan menggunakan algoritma XGBoost untuk klasifikasi multikelas dengan bantuan perangkat lunak RapidMiner. Untuk mengatasi potensi ketidakseimbangan data digunakan teknik pembagian data yang proporsional. Evaluasi model dilakukan menggunakan *confusion matrix* dengan metrik accuracy, precision, dan recall untuk mengukur kinerja klasifikasi (Anisah & Irwansyah, 2025). Hasil dari proses ini kemudian dianalisis dan disajikan dalam bentuk laporan sebagai dasar untuk menarik kesimpulan serta memberikan rekomendasi terkait mitigasi risiko kecelakaan kereta api berbasis data.

3. Hasil dan Pembahasan

3.1 Deskripsi Data

Data yang digunakan dalam penelitian ini merupakan data historis kecelakaan kereta api di Indonesia pada periode tahun 1991 hingga 2025. Data diperoleh dari laporan resmi Komite Nasional Keselamatan Transportasi (KNKT) serta berbagai sumber media daring yang relevan. Informasi yang dikumpulkan meliputi tahun kejadian, lokasi kecelakaan, jenis insiden, penyebab utama, jumlah korban, serta klasifikasi tingkat risiko. Pada tahap awal, jumlah data yang berhasil dikumpulkan sebanyak 140 baris. Namun, setelah dilakukan proses penyaringan untuk menghilangkan data yang tidak lengkap, duplikat, atau tidak relevan, jumlah data yang layak digunakan dalam analisis berkurang menjadi 112 baris.

3.2 Pelabelan Data

Sebelum dilakukan proses pelatihan model, dilakukan analisis eksplorasi data (*Exploratory Data Analysis*) untuk memahami distribusi jumlah korban, frekuensi jenis kecelakaan, serta mendeteksi adanya nilai ekstrem (*outlier*) yang berpotensi memengaruhi kinerja model. Selanjutnya, dilakukan proses pelabelan data berdasarkan jumlah total korban kecelakaan yang terdiri dari korban meninggal dunia, luka berat, dan luka ringan. Klasifikasi dilakukan ke dalam tiga kategori, yaitu:

- a. LOW: tidak terdapat korban (0 korban)
- b. MEDIUM: jumlah korban antara 1 hingga 5
- c. HIGH: jumlah korban lebih dari 5

Pelabelan ini bertujuan untuk membangun model klasifikasi multikelas menggunakan algoritma XGBoost.

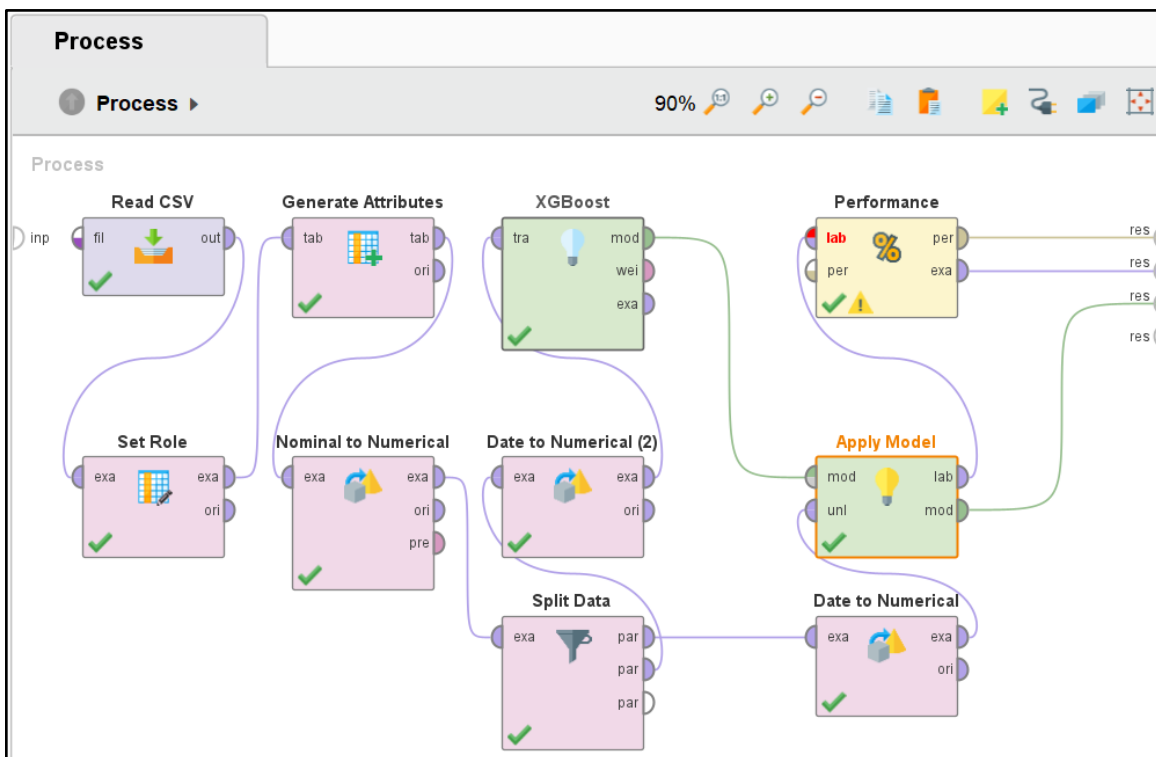
3.3 Preprocessing Data

Data yang telah dikumpulkan kemudian melalui tahap pra-pemrosesan untuk memastikan kualitas data sebelum digunakan dalam pelatihan model. Tahapan ini meliputi pembersihan data dari nilai kosong, penghapusan data duplikat, serta transformasi data ke dalam format yang sesuai. Atribut bertipe tanggal diubah menjadi bentuk numerik agar dapat diproses oleh algoritma *XGBoost*. Selain itu, atribut kategorikal dikonversi menjadi bentuk numerik menggunakan teknik encoding.

3.4 Implementasi Algoritma XGBoost

Dalam penelitian ini, beberapa *hyperparameter* utama yang digunakan meliputi *learning rate*, *max depth*, *number of estimators*, serta parameter regularisasi L1 dan L2. Nilai *learning rate* ditetapkan sebesar 0,3 untuk menjaga keseimbangan antara kecepatan konvergensi dan stabilitas model. Parameter *max depth* sebesar 6 digunakan untuk mengontrol kompleksitas pohon keputusan agar tidak terlalu dalam dan berpotensi menyebabkan *overfitting*. Jumlah estimator (*n_estimators*) ditentukan sebanyak 100, yang dianggap cukup untuk menghasilkan performa model yang stabil tanpa meningkatkan beban komputasi secara signifikan. Untuk mengurangi risiko *overfitting*, digunakan regularisasi L1 (*alpha*) sebesar 1,0 dan regularisasi L2 (*lambda*) sebesar 2,0. Selain itu, parameter *gamma* ditetapkan sebesar 0,1 untuk membatasi pemisahan cabang yang tidak memberikan penurunan loss yang signifikan.

Setelah proses pelatihan model XGBoost selesai, dilakukan evaluasi performa menggunakan tiga metrik utama, yaitu *accuracy*, *precision*, dan *recall*. Metrik ini digunakan untuk mengukur kemampuan model dalam mengklasifikasikan jumlah korban kecelakaan ke dalam tiga kategori, yaitu *Low*, *Medium*, dan *High*.



Gambar 3. Implementasi Algoritma XGBoost

Proses pengolahan data dalam penelitian ini dilakukan menggunakan beberapa operator pada RapidMiner yang tersusun secara sistematis. Tahap awal dimulai dengan operator *Read CSV* yang digunakan untuk membaca data kecelakaan kereta api dalam format CSV yang telah disiapkan sebelumnya. Selanjutnya, operator *Set Role* digunakan untuk menentukan atribut target (label), yaitu kategori jumlah korban yang terdiri dari *Low*, *Medium*, dan *High*. Operator *Generate Attributes* kemudian dimanfaatkan untuk membentuk atribut baru berdasarkan transformasi atau perhitungan dari atribut yang sudah ada. Untuk memastikan data dapat diproses oleh algoritma, dilakukan konversi atribut menggunakan operator *Nominal to Numerical* guna mengubah data kategorikal menjadi numerik, serta *Date to Numerical* untuk mengonversi atribut bertipe tanggal menjadi bentuk numerik. Setelah itu, data dibagi menjadi dua bagian menggunakan operator *Split Data* dengan rasio 80:20, di mana 80% data digunakan sebagai data pelatihan dan 20% sebagai data pengujian. Model kemudian dibangun

menggunakan algoritma *XGBoost* sebagai metode utama, dengan parameter yang telah ditentukan sebelumnya seperti *max_depth*, *learning_rate*, *lambda*, dan *alpha*. Model yang telah dilatih selanjutnya diterapkan pada data pengujian menggunakan operator *Apply Model* untuk menghasilkan prediksi kelas jumlah korban. Tahap akhir dilakukan evaluasi menggunakan operator *Performance*, yang menghasilkan metrik evaluasi seperti akurasi, *precision*, dan *recall* untuk menilai kinerja model secara keseluruhan.

3.5 Evaluasi dan Hasil Implementasi Algoritma *XGBoost*

Confusion Matrix untuk pemodelan menggunakan algoritma *XGBoost*, seperti yang ditunjukkan pada Gambar 4, menunjukkan bahwa terdapat 99 prediksi yang benar dari total 112 data uji. Berdasarkan hasil tersebut, diperoleh nilai *True Positive* (TP) untuk masing-masing kelas, yaitu 11 pada kelas *Medium*, 39 pada kelas *High*, dan 49 pada kelas *Low*. Selain itu, terdapat kesalahan prediksi yang ditunjukkan oleh nilai *False Positive* (FP), yaitu 10 pada kelas *Medium* dan 3 pada kelas *Low*. Sementara itu, nilai *False Negative* (FN) tercatat sebanyak 13 pada kelas *High*, yang merupakan data yang seharusnya termasuk kelas *High* namun diprediksi sebagai kelas *Medium* dan *Low*. Hasil evaluasi *confusion matrix* tersebut disajikan pada Gambar 4.

accuracy: 88.39%				
	true MEDIUM	true HIGH	true LOW	class precision
pred. MEDIUM	11	10	0	52.38%
pred. HIGH	0	39	0	100.00%
pred. LOW	0	3	49	94.23%
class recall	100.00%	75.00%	100.00%	

Gambar 4. Hasil Evaluasi *Confusion Matrix*

Perhitungan *Accuracy*

$$Accuracy = \frac{\text{Jumlah prediksi benar}}{\text{Total seluruh data}} = \frac{(11+39+49)}{112} = \frac{99}{112} = 0.8839 \times 100 = 88.39\% \quad (1)$$

Perhitungan *Precision* per Kelas

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

a. *Precision Medium*

TP = 11 (prediksi *Medium* dan benar)

FP = 10 (prediksi *Medium* tapi sebenarnya *High*)

$$Precision = \frac{TP}{TP+FP} = \frac{11}{(11+10)} = \frac{11}{21} = 0.5238 \times 100 = 52.38\% \quad (3)$$

b. *Precision High*

TP = 39

FP = 0

$$Precision = \frac{TP}{TP+FP} = \frac{39}{(39+0)} = \frac{39}{39} = 1 \times 100 = 100\% \quad (4)$$

c. *Precision Low*

$$TP = 49$$

FP = 3 (prediksi *LOW* tapi seharusnya *High*)

$$Precision = \frac{TP}{TP+FP} = \frac{49}{(49+3)} = \frac{49}{52} = 0.9423 \times 100 = 94.23\% \quad (5)$$

Recall per Kelas

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

a. *Recall Medium*

$$TP = 11$$

$$FN = 0$$

$$Recall = \frac{TP}{TP+FN} = \frac{11}{(11+0)} = \frac{11}{11} = 1 \times 100 = 100\% \quad (7)$$

b. *Recall High*

$$TP = 39$$

FN = 10 (prediksi *Medium*) + 3 (prediksi *Low*) = 13

$$Recall = \frac{TP}{TP+FN} = \frac{39}{(39+13)} = \frac{39}{52} = 0.75 \times 100 = 75\% \quad (8)$$

c. *Recall Low*

$$TP = 49$$

$$FN = 0$$

$$Recall = \frac{TP}{TP+FN} = \frac{49}{(49+0)} = \frac{49}{49} = 1 \times 100 = 100\% \quad (9)$$

Hasil pengujian menunjukkan bahwa algoritma *XGBoost* memiliki kemampuan yang baik dalam membedakan dan mengklasifikasikan jumlah korban kecelakaan kereta api berdasarkan data historis. Nilai akurasi yang mencapai 88,39% menunjukkan bahwa model memiliki keandalan tinggi. Meskipun akurasi mencapai 88,39% ketidakseimbangan data antar kelas menyebabkan *precision* kelas *Medium* rendah.

Pembahasan

Hasil penelitian menunjukkan bahwa model klasifikasi berbasis algoritma *XGBoost* mampu mengklasifikasikan jumlah korban kecelakaan kereta api ke dalam tiga kategori risiko, yaitu rendah (*Low*), sedang (*Medium*), dan tinggi (*High*), dengan tingkat akurasi sebesar 88,39%. Nilai tersebut mengindikasikan bahwa model memiliki kemampuan yang baik dalam memetakan pola dari data historis ke dalam kelas yang sesuai. Secara lebih rinci, performa model pada kelas *Low* dan *High* tergolong tinggi, dengan nilai *precision* dan *recall* masing-masing sebesar 0,91 dan 0,94 untuk kelas *Low*, serta 0,88 dan 0,87 untuk kelas *High*. Hal ini menunjukkan bahwa model mampu mengidentifikasi kedua kategori tersebut secara konsisten. Sebaliknya, performa pada kelas *Medium* masih relatif lebih rendah, dengan nilai *precision* sebesar 0,75 dan *recall* sebesar 0,68, yang mengindikasikan adanya kesulitan model dalam membedakan kelas ini dari kelas lainnya. Kondisi tersebut mengarah pada kemungkinan adanya tumpang tindih (*overlap*) karakteristik antar kelas serta keterbatasan fitur dalam merepresentasikan perbedaan secara diskriminatif.

Dari perspektif evaluasi model, hasil confusion matrix menunjukkan bahwa kesalahan klasifikasi didominasi oleh kasus pada kelas *High* yang diprediksi sebagai *Medium* atau *Low*, sebagaimana tercermin dari

nilai *false negative* yang relatif tinggi. Fenomena ini menunjukkan bahwa model belum sepenuhnya sensitif dalam mengidentifikasi kejadian dengan tingkat risiko tinggi. Di sisi lain, tingkat kesalahan pada kelas *Low* relatif rendah, yang mengindikasikan bahwa pola pada kelas tersebut lebih stabil dan mudah dikenali. Meskipun nilai akurasi secara keseluruhan tergolong tinggi, distribusi performa yang tidak merata antar kelas menegaskan pentingnya penggunaan metrik evaluasi tambahan seperti *precision* dan *recall* dalam menilai kinerja model secara lebih komprehensif, khususnya pada kasus klasifikasi multikelas dengan distribusi data yang tidak seimbang.

Temuan dalam penelitian ini konsisten dengan studi sebelumnya yang menyatakan bahwa *XGBoost* merupakan salah satu algoritma yang unggul dalam menangani permasalahan klasifikasi pada data kecelakaan transportasi. Bridgelall dan Tolliver (2021) menunjukkan bahwa *XGBoost* memiliki kemampuan tinggi dalam menangkap hubungan non-linear antar variabel dengan performa evaluasi yang kuat. Selain itu, Senkondo et al. (2025) juga melaporkan bahwa pendekatan berbasis *machine learning*, termasuk *XGBoost*, mampu mengungguli metode statistik konvensional dalam analisis keselamatan pada perlintasan kereta api. Dengan demikian, hasil penelitian ini memperkuat posisi *XGBoost* sebagai algoritma yang efektif dalam konteks klasifikasi risiko kecelakaan. Namun demikian, dibandingkan dengan penelitian Wang et al. (2025) yang mengintegrasikan pendekatan *Explainable Artificial Intelligence* (XAI), penelitian ini masih memiliki keterbatasan dalam aspek interpretabilitas model, karena belum mengeksplorasi kontribusi masing-masing fitur terhadap hasil prediksi secara mendalam.

Performa model dalam penelitian ini dipengaruhi oleh beberapa faktor utama, di antaranya kualitas dan kuantitas data, distribusi kelas, serta representasi fitur. Jumlah dataset yang relatif terbatas, yaitu sebanyak 112 data setelah proses *preprocessing*, berpotensi membatasi kemampuan model dalam menangkap pola yang lebih kompleks. Selain itu, ketidakseimbangan distribusi data antar kelas menyebabkan model cenderung lebih optimal dalam mempelajari kelas mayoritas dibandingkan kelas minoritas, khususnya pada kategori *Medium*. Faktor lain yang turut berkontribusi adalah kemungkinan adanya kemiripan karakteristik antara kelas *Medium* dan *High*, sehingga memperbesar peluang terjadinya misklasifikasi. Di samping itu, pemilihan fitur yang belum sepenuhnya optimal juga dapat memengaruhi kemampuan model dalam membedakan antar kelas secara signifikan.

Penelitian ini memiliki beberapa keunggulan, antara lain penggunaan pendekatan CRISP-DM yang sistematis dalam proses pengolahan data, penerapan algoritma *XGBoost* yang dikenal memiliki performa tinggi, serta pemanfaatan data historis dalam rentang waktu yang panjang sehingga memberikan konteks analisis yang lebih komprehensif. Selain itu, evaluasi model dilakukan menggunakan beberapa metrik kinerja, sehingga mampu memberikan gambaran yang lebih menyeluruh terhadap performa model. Namun demikian, penelitian ini juga memiliki sejumlah keterbatasan, di antaranya keterbatasan jumlah data, belum dilakukannya analisis *feature importance* atau pendekatan *explainability*, adanya ketidakseimbangan data antar kelas, serta belum adanya komparasi dengan algoritma lain sebagai baseline untuk mengukur keunggulan relatif model yang digunakan.

Secara keseluruhan, penelitian ini memberikan implikasi yang signifikan baik secara teoritis maupun praktis. Dari sisi teoritis, penelitian ini berkontribusi dalam pengembangan model klasifikasi risiko kecelakaan berbasis *machine learning*, khususnya pada domain transportasi kereta api di Indonesia. Dari sisi praktis, model yang dihasilkan berpotensi digunakan sebagai alat bantu dalam pengambilan keputusan oleh pemangku kepentingan, seperti pemerintah dan operator transportasi, dalam upaya mitigasi risiko kecelakaan. Dengan adanya klasifikasi tingkat risiko yang akurat, proses penentuan prioritas penanganan dapat dilakukan secara lebih efektif dan berbasis data. Selain itu, penelitian ini juga membuka peluang pengembangan sistem prediktif yang bersifat preventif, sehingga dapat mendukung upaya peningkatan keselamatan transportasi secara berkelanjutan.

4. Kesimpulan

Penelitian ini bertujuan untuk mengklasifikasikan jumlah korban kecelakaan kereta api di Indonesia ke dalam kategori risiko rendah, sedang, dan tinggi menggunakan algoritma XGBoost, serta mengevaluasi kinerja model berdasarkan metrik *accuracy*, *precision*, dan *recall*. Berdasarkan hasil yang diperoleh, dapat disimpulkan bahwa algoritma XGBoost mampu memberikan performa yang baik dalam melakukan klasifikasi multikelas, dengan tingkat akurasi sebesar 88,39%. Model menunjukkan kinerja yang optimal pada kelas risiko rendah dan tinggi, yang ditunjukkan oleh nilai *precision* dan *recall* yang tinggi, sementara performa pada kelas risiko sedang masih relatif lebih rendah. Hal ini mengindikasikan bahwa model masih menghadapi tantangan dalam membedakan kelas dengan karakteristik yang saling beririsan serta adanya ketidakseimbangan distribusi data.

Secara keseluruhan, penelitian ini membuktikan bahwa pendekatan berbasis *machine learning*, khususnya *XGBoost*, efektif digunakan untuk mendukung analisis risiko kecelakaan kereta api berbasis data historis. Model yang dihasilkan dapat menjadi dasar dalam pengambilan keputusan untuk mitigasi risiko, terutama dalam mengidentifikasi tingkat keparahan kecelakaan berdasarkan jumlah korban. Dengan demikian, penelitian ini memberikan kontribusi dalam pengembangan sistem klasifikasi risiko yang dapat dimanfaatkan untuk meningkatkan keselamatan transportasi kereta api di Indonesia.

Adapun untuk penelitian selanjutnya, disarankan untuk meningkatkan jumlah dan kualitas dataset agar model dapat mempelajari pola yang lebih kompleks dan meningkatkan kemampuan generalisasi. Selain itu, penerapan teknik penanganan data tidak seimbang, seperti *oversampling* atau *undersampling*, dapat dipertimbangkan untuk meningkatkan performa pada kelas minoritas. Penelitian berikutnya juga dapat mengintegrasikan metode *Explainable Artificial Intelligence (XAI)* untuk meningkatkan interpretabilitas model, sehingga faktor-faktor yang memengaruhi tingkat risiko kecelakaan dapat diidentifikasi secara lebih jelas. Di samping itu, perbandingan dengan algoritma *machine learning* lainnya serta optimasi *hyperparameter* yang lebih mendalam diharapkan dapat menghasilkan model dengan performa yang lebih optimal.

Daftar Pustaka

- Anisah, S., & Irwansyah, I. (2025). Analisis Data Mining untuk Klasifikasi Kafe Populer di Jakarta Menggunakan Decision Tree dan Visualisasi dengan Tableau. *Jutisi : Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi*, 14(2), 890. <https://doi.org/10.35889/jutisi.v14i2.2660>
- Ari Rama Novryadi, Irwansyah, & Moh Shidqon. (2026). Analisis Sentimen Masyarakat Terhadap Kinerja Presiden Indonesia Joko Widodo Periode Kedua Menggunakan Metode Naïve Bayes dan SVM. *DIGINTEL-AI : DIGital INnovation and inTELLigence – AI*, 1(1), 11–24. <https://doi.org/10.66217/digintel-ai.v1i1.2>
- Bagaskara, W., Pusparini, N. N., & Irwansyah, I. (2024). Klasifikasi Penjadwalan Kerja Perawatan Air Conditioner (Ac) Menggunakan Algoritma Decision Tree (C4.5) Pada Pt Xyz. *Infotech: Journal of Technology Information*, 10(1), 11–20. <https://doi.org/10.37365/jti.v10i1.240>
- Bridgelall, R., & Tolliver, D. D. (2021). Railroad accident analysis using extreme gradient boosting. *Accident Analysis & Prevention*, 156, 106126. <https://doi.org/10.1016/j.aap.2021.106126>
- Budiyantara, A., Irwansyah, I., Prengki, E., Pratama, P. A., & Wiliani, N. (2020). Komparasi Algoritma Decision Tree, Naive Bayes Dan K-Nearest Neighbor Untuk Memprediksi Mahasiswa Lulus Tepat

- Waktu. *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, 5(2), 265–270. <https://doi.org/10.33480/jitk.v5i2.1214>
- Ihza Kurniawan, D., Irwansyah, & Taufik, A. (2026). Analisis Sentimen Terhadap Komentar Video IShowSpeed Tour Indonesia Pada YouTube Menggunakan Metode SVM. *DIGINTEL-AI: DIGital INnovation and inTELLigence – AI*, 1(1), 50–62. <https://doi.org/10.66217/digintel-ai.v1i1.5>
- Irwansyah, I., Dittyata, R., Rizal, R., & Wiyono, W. (2024). Optimalisasi Klasifikasi Uji Emisi Sepeda Motor Menggunakan Algoritma Naïve Bayes. *Infotech: Journal of Technology Information*, 10(2), 337–242. <https://doi.org/10.37365/jti.v10i2.327>
- Liu, J., Wang, Y., Deng, C., Hou, F., Jin, Z., Qiao, L., & Wang, G. (2025). A new procedure for assessing and predicting the severity of accidents: A case study on freight-train derailments. *Journal of Loss Prevention in the Process Industries*, 94, 105511. <https://doi.org/10.1016/j.jlp.2024.105511>
- Liu, X., Turla, T., & Zhang, Z. (2018). Accident-Cause-Specific Risk Analysis of Rail Transport of Hazardous Materials. *Transportation Research Record: Journal of the Transportation Research Board*, 2672(10), 176–187. <https://doi.org/10.1177/0361198118794532>
- Nafisah Nurul Hakim. (2020). Implementasi Machine Learning pada Sistem Prediksi Kejadian dan Lokasi Patah Rel Kereta Api di Indonesia. *Jurnal Sistem Cerdas*, 3(1), 25–35. <https://doi.org/10.37396/jsc.v3i1.58>
- Permatasari, M. P., Kristanto, D., Ervianty, R. M., Salam, M. D., Dwikesumasari, P. R., & Zulfa, V. R. (2024). *Persepsi Kepuasan Layanan Transportasi Kereta Api Pt. Kereta Api Indonesia (Pt.Kai): Survey Pada Mahasiswa Yang Berkuliah Di Surabaya*. 4(10).
- Salsa Billa Permana Putri, Irwansyah, & Tri M, T. (2026). Implementasi Algoritma K-NN Pada Sosial Media X Untuk Analisis Sentimen Pengalaman Warganet Tinggal Di Luar Negeri. *DIGINTEL-AI: DIGital INnovation and inTELLigence – AI*, 1(1), 37–49. <https://doi.org/10.66217/digintel-ai.v1i1.4>
- Senkondo, E., Chimba, D., Madalo, M., Yeboah, A., & Blue, S. (2025). Comparative Analysis of Machine Learning and Statistical Models for Railroad–Highway Grade Crossing Safety. *Vehicles*, 7(4), 163. <https://doi.org/10.3390/vehicles7040163>
- Sinai, N. K., Dzulkifli, D., Pertiwi, S., Chentia, R., Halimatussakdiyah, H., Furqan, M., Alfaiz, A. B., Nafyla, A., Arrafif, C. R., Zawira, S., & Windasari, N. (2024). Kematian Akibat Kecelakaan Kereta Api: Laporan Kasus. *Jurnal Ilmu Kesehatan Indonesia*, 5(3), 268–272. <https://doi.org/10.25077/jikesi.v5i3.1246>
- Wang, D., He, Q., Peng, J., & Li, G. (2025). Enhancing Intelligent Transportation Safety with Explainable AI: A Framework for Uncovering Crash Severity Factors at Highway–Rail Grade Crossings. *World Electric Vehicle Journal*, 16(11), 637. <https://doi.org/10.3390/wevj16110637>
- Zhou, X., Lu, P., Zheng, Z., Tolliver, D., & Keramati, A. (2020). Accident Prediction Accuracy Assessment for Highway-Rail Grade Crossings Using Random Forest Algorithm Compared with Decision Tree. *Reliability Engineering & System Safety*, 200, 106931. <https://doi.org/10.1016/j.ress.2020.106931>