

Model Prediksi Tingkat Keparahan Kecelakaan Lalu Lintas di DKI Jakarta dengan Regresi Logistik Ordinal

Prima Nuriyan Saputra^{1*}, Budi Hartanto Susilo², Darmawan Pontan³

¹⁻³Program Studi Transportasi, Teknik Sipil dan Perencanaan, Universitas Trisakti, Indonesia

Corresponding Author's e-mail: primamilano88@gmail.com



e-ISSN: 2964-2981

ARMADA : Jurnal Penelitian Multidisiplin

<https://ejournal.45mataram.ac.id/index.php/armada>

Vol. 04, No. 07 Juli, 2026

Page: 2407-2422

DOI:

<https://doi.org/10.55681/armada.v4i7.2827>

Article History:

Received: Mei 21, 2026

Revised: Juni 13, 2026

Accepted: Juli 18, 2026

Abstract : Road traffic accidents are a transport safety problem in Jakarta because of mobility and complex risks. This study aims to develop an Ordinal Logistic Regression model to analyse accident severity trends using IRSMS data from the Indonesian National Police Traffic Corps for 2022–2024. The dependent variable covers material damage, minor injuries, serious injuries, and fatalities, while the independent variables include time, weather, location, accident type, vehicle type, age, and victims' sex. The analysis involved descriptive statistics, assumption testing, Proportional Odds Model estimation, model fit evaluation, and risk simulation. Results indicate that the models satisfied the statistical assumptions. The 2023 model was the best, with a Model Fitting Information significance value of 0.006 and a Nagelkerke R-Square of 0.360, whereas the 2022 and 2024 models reached only 0.142 and 0.144. Most independent variables were not significant at the 5% level, but age emerged as the most dominant factor based on the odds ratio. Limited valid data and missing values reduced explanatory power. The model is therefore more suitable for risk exploration than operational prediction.

Keywords: Traffic Accidents, Accident Severity, Ordinal Logistic Regression, Road Safety, Metropolitan

Abstrak : Kecelakaan lalu lintas menjadi masalah utama keselamatan transportasi di DKI Jakarta karena tingginya mobilitas dan kompleksitas risikonya. Penelitian ini bertujuan mengembangkan model Regresi Logistik Ordinal untuk menganalisis kecenderungan tingkat keparahan kecelakaan berdasarkan data IRSMS Korlantas Polri periode 2022–2024. Variabel dependen mencakup kerugian material, luka ringan, luka berat, dan meninggal dunia, sedangkan variabel independen meliputi waktu kejadian, cuaca, lokasi, jenis kecelakaan, jenis kendaraan, usia, dan jenis kelamin korban. Analisis dilakukan melalui statistik deskriptif, pengujian asumsi, estimasi Proportional Odds Model, evaluasi kecocokan model, dan simulasi risiko. Hasil menunjukkan bahwa model secara umum memenuhi asumsi statistik. Model tahun 2023 menjadi model terbaik dengan nilai signifikansi Model Fitting Information sebesar 0,006 dan Nagelkerke R-Square sebesar 0,360, sedangkan model 2022 dan 2024 masing-masing hanya mencapai 0,142 dan 0,144. Sebagian besar variabel independen belum berpengaruh signifikan pada taraf 5%, tetapi usia korban menunjukkan kecenderungan sebagai faktor paling dominan berdasarkan nilai odds ratio. Keterbatasan data valid dan tingginya missing value mengurangi daya jelas model. Model lebih tepat digunakan untuk eksplorasi risiko dan

identifikasi kecenderungan keparahan dibandingkan sebagai alat prediksi operasional.

Kata Kunci: Kecelakaan Lalu Lintas, Tingkat Keparahannya, Kecelakaan, Regresi Logistik Ordinal, Keselamatan Jalan, Metropolitan

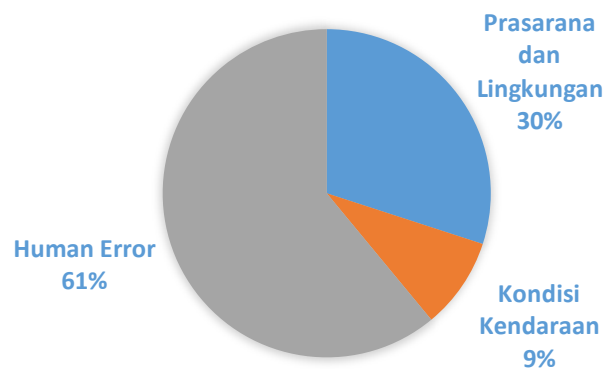
PENDAHULUAN

Lebih dari 1,35 juta orang meninggal dunia akibat kecelakaan lalu lintas setiap tahun, dan puluhan juta lainnya mengalami cedera serius yang berdampak jangka panjang pada kehidupan mereka dan keluarga mereka, menurut (World Health Organization, 2023). Sebagai negara dengan tingkat urbanisasi yang tinggi, Indonesia juga menghadapi masalah serupa. Sebagai informasi yang diperoleh dari *Indonesia Road Safety Profile 2025*, sekitar 31.000 orang meninggal akibat kecelakaan lalu lintas pada tahun 2021, atau sekitar 2% dari total kematian nasional (Summary, 2007). Data ini menunjukkan bahwa tujuannya untuk mengurangi angka kematian di jalan masih jauh dari target yang ditetapkan dalam *Decade of Action for Road Safety*. Kecenderungan kecelakaan lalu lintas sangat berbeda di seluruh negara. Jumlah kecelakaan menurun menjadi 100.028 pada tahun 2020 dari 116.411 kasus pada tahun 2019 (Kusumastutie *et al.*, 2024). Namun, kembali meningkat menjadi 103.645 kasus pada tahun 2021. Menurut penelitian yang dilakukan oleh (Basri *et al.*, 2021) mayoritas kecelakaan terjadi pada jam sibuk, dan dipengaruhi oleh pengendara sepeda motor dan perilaku pengguna jalan.

Kondisi lalu lintas yang padat, berbagai jenis pengguna jalan, dan jaringan jalan perkotaan yang kompleks meningkatkan risiko kecelakaan di wilayah DKI Jakarta. Sebuah studi mengenai fasilitas pejalan kaki menemukan 2.646 kecelakaan pejalan kaki di wilayah DKI Jakarta dari tahun 2020 hingga 2022, dengan tren meningkat dari 837 kasus pada tahun 2020 menjadi 931 kasus pada tahun 2022 (Siregar *et al.*, 2024). Ini menunjukkan peningkatan sekitar 50% dalam angka kecelakaan dari 2018 hingga 2019, menunjukkan betapa pentingnya perbaikan sistem keselamatan jalan raya (Ricardianto *et al.*, 2021). Sayangnya, publikasi resmi lebih banyak menyoroti jumlah kasus tanpa menguraikan faktor-faktor yang mempengaruhi tingkat keparahan kecelakaan.

Dari sudut pandang akademis, kebanyakan penelitian kecelakaan lalu lintas di Indonesia menggunakan pendekatan deskriptif atau regresi logistik binari/multinomial. Untuk menilai tingkat keparahan, beberapa penelitian menggunakan regresi logistik ordinal, seperti penelitian di Yogyakarta dan Padang. Namun, mereka tidak memperhitungkan variabel spasial atau heterogenitas antar lokasi kejadian, dan mereka juga terbatas dalam hal cakupan wilayah dan periode data mereka. Namun, penelitian global menunjukkan bahwa karakteristik pengemudi, jenis kendaraan, cuaca, pencahayaan, dan kondisi permukaan jalan memengaruhi tingkat fatalitas (W. Darwish & Darwish, 2023). Menurut Penelitian (Shiran *et al.*, 2021), metode regresi logistik ordinal memiliki tingkat interpretabilitas yang lebih tinggi dibandingkan dengan model pembelajaran mesin, yang membuatnya tetap relevan. Namun, prediksinya mungkin kurang akurat.

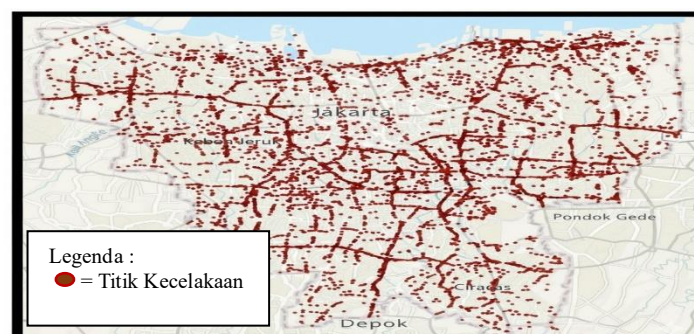
Selain itu, data resmi Korlantas Polri menunjukkan bahwa penyebab kecelakaan di Indonesia didominasi oleh faktor manusia (61%), prasarana dan lingkungan (30%), serta kondisi kendaraan (9%) (Gambar 1). Temuan ini konsisten dengan (Soehodho, 2009), (Santosa, 2017), dan (Puspasari *et al.*, 2023) yang menegaskan dominasi human error dalam kecelakaan nasional. Literatur internasional seperti (H. H. Darwish, 2021) serta (O'Donnell & Connor, 1996) juga mendukung bahwa faktor perilaku, lingkungan, dan kondisi teknis kendaraan memengaruhi tingkat fatalitas. Integrasi ini memperkuat urgensi penelitian menggunakan Regresi Logistik Ordinal pada data Korlantas Polri.



Gambar 1. Persentase Penyebab Kecelakaan
(Sumber: Korlantas polri)

Berdasarkan telaah literatur, terdapat tiga kesenjangan penelitian yang masih ditemukan. Pertama, sebagian besar penelitian di Indonesia masih menggunakan pendekatan deskriptif atau regresi logistik biner sehingga belum sepenuhnya mempertimbangkan karakter ordinal dari tingkat keparahan kecelakaan. Kedua, penelitian yang memanfaatkan data mikro kecelakaan pada wilayah metropolitan dengan periode observasi multiyear masih relatif terbatas. Ketiga, penelitian terdahulu umumnya berhenti pada identifikasi faktor risiko tanpa mengembangkan interpretasi risiko yang dapat digunakan sebagai dasar pertimbangan kebijakan keselamatan transportasi. Penggunaan data mikro kecelakaan Korlantas Polri selama tiga tahun berturut-turut (2022–2024) serta pengembangan model prediksi yang dapat digunakan sebagai dasar prioritas kebijakan keselamatan lalu lintas perkotaan menggunakan regresi logistik ordinal pada wilayah metropolitan di Indonesia dan kurangnya fokus pada penerapan hasil penelitian untuk kebijakan publik. Oleh karena itu, penelitian ini menawarkan inovasi berupa pengembangan model prediksi tingkat keparahan kecelakaan dengan regresi logistik ordinal yang memanfaatkan data kepolisian terkini (Tahun 2022 – 2024) dengan mengintegrasikan faktor waktu kejadian, kondisi cuaca, lokasi kejadian, jenis kecelakaan, jenis kendaraan, usia korban, dan jenis kelamin korban dalam model Regresi Logistik Ordinal.

Meskipun berbagai penelitian mengenai kecelakaan lalu lintas telah dilakukan di Indonesia, sebagian besar masih berfokus pada analisis deskriptif atau menggunakan cakupan wilayah yang terbatas (gambar 2). Penelitian yang memanfaatkan data mikro kecelakaan pada wilayah metropolitan dengan pendekatan Regresi Logistik Ordinal masih relatif terbatas. Oleh karena itu penelitian ini dilakukan untuk mengisi kesenjangan tersebut. Kebaruan penelitian ini terletak pada penggunaan data mikro kecelakaan Korlantas Polri selama tiga tahun berturut-turut (2022–2024), penerapan Regresi Logistik Ordinal pada wilayah metropolitan Indonesia, serta pengembangan interpretasi risiko keparahan kecelakaan berdasarkan model terbaik yang diperoleh.



Gambar 2. Data Kecelakaan Lalu Lintas Provinsi DKI Jakarta Periode Januari 2019 –Desember 2021 (Irfan et al., 2021).

METODE PENELITIAN

Penelitian ini dilaksanakan di wilayah DKI Jakarta yang meliputi Jakarta Pusat, Jakarta Barat, Jakarta Timur, Jakarta Selatan, dan Jakarta Utara. Pemilihan lokasi penelitian didasarkan pada tingginya kepadatan lalu lintas di DKI Jakarta yang mencapai 18.213 kendaraan per km² berdasarkan data BPS tahun 2024, menjadikannya wilayah dengan beban lalu lintas tertinggi di Indonesia. Selain itu, Jakarta merupakan pusat aktivitas ekonomi, pemerintahan, dan transportasi nasional dengan mobilitas kendaraan yang sangat tinggi, yaitu mencapai 44.991.781 kendaraan yang keluar masuk wilayah Jakarta pada tahun 2023. Kondisi tersebut menyebabkan tingginya potensi kecelakaan lalu lintas. Penelitian ini juga didukung oleh ketersediaan data kecelakaan yang terdokumentasi dengan baik melalui *IRSMS* Korlantas Polri, sehingga relevan untuk pengembangan model prediksi keparahan kecelakaan. Data yang digunakan berupa data sekunder kecelakaan lalu lintas tahun 2022–2024 dengan unit analisis berupa setiap kasus kecelakaan yang terjadi di wilayah Jakarta. Variabel penelitian terdiri atas tingkat keparahan kecelakaan sebagai variabel terikat, serta waktu kejadian, kondisi cuaca, lokasi kejadian, jenis kecelakaan, jenis kendaraan, usia korban, dan jenis kelamin sebagai variabel bebas. Penelitian ini menggunakan pendekatan kuantitatif dengan desain observasional retrospektif, memanfaatkan data historis tanpa intervensi peneliti, serta dianalisis secara statistik menggunakan regresi ordinal. Teknik sampling yang digunakan adalah total sampling terhadap seluruh kasus kecelakaan yang tercatat selama periode penelitian, dengan instrumen berupa formulir coding data yang digunakan untuk merekap variabel dari database kecelakaan Korlantas Polri yang diperoleh dalam format Microsoft Excel.

HASIL DAN PEMBAHASAN

Analisis Inferensial

Analisis Model Tahun 2022 – 2024

Rekapitulasi Hasil Uji Multikolinearitas Data Tahun 2022-2024

Tabel 1. Rekapitulasi Hasil Uji Multikolinearitas Data Tahun 2022-2024

Tahun	Variabel yang Dianalisis	Rentang Tolerance	Rentang VIF	Kesimpulan
2022	Kondisi Cuaca, Lokasi, Jenis Kecelakaan, Jenis Kendaraan, Usia, Jenis Kelamin	0,95 – 0,97	1,02 – 1,04	Tidak terjadi multikolinearitas
2023	Jenis Kendaraan, Usia, Jenis Kelamin	0,91 – 0,95	1,04 – 1,09	Tidak terjadi multikolinearitas
2024	Lokasi, Jenis Kecelakaan, Usia, Jenis Kelamin	0,64 – 0,78	1,27 – 1,54	Tidak terjadi multikolinearitas

Berdasarkan hasil rekapitulasi uji multikolinearitas tahun 2022 sampai dengan tahun 2024 diketahui bahwa seluruh variabel independen memiliki nilai tolerance lebih besar dari 0,10 dan nilai Variance Inflation Factor (VIF) lebih kecil dari 10. Nilai VIF tertinggi ditemukan pada tahun 2024 yaitu sebesar 1,54 pada variabel jenis kecelakaan. Namun demikian, nilai tersebut masih jauh di bawah batas maksimum 10 sehingga belum menunjukkan adanya hubungan korelasi yang kuat antar variabel independen. Sementara itu, nilai VIF terendah ditemukan pada tahun 2022 yaitu sebesar 1,02 pada variabel jenis kecelakaan dan jenis kendaraan. Nilai tersebut menunjukkan bahwa hubungan antar variabel bebas dalam model penelitian sangat rendah.

Rekapitulasi Hasil Regresi Logistik Ordinal Tahun 2022–2024

Tabel 2. Rekapitulasi Hasil Regresi Logistik Ordinal Tahun 2022 ~ 2024

Tahun	Model Fit Sig.	Goodness of Fit	Nagelkerke	Parallel Lines	Variabel Signifikan	Kesimpulan Model		
2022	0,83	Fit	0,14	0,68	Tidak ada	Model layak kemampuan rendah	namun prediksi	

2023	0,006	Fit	0,36	Tidak konvergen	Tidak ada	Model cukup baik namun data terbatas
2024	0,95	Fit	0,14	0,32	Tidak ada	Model layak namun kemampuan prediksi rendah

Hasil penelitian menunjukkan bahwa tidak ditemukannya variabel yang signifikan secara statistik bukan berarti faktor-faktor yang diuji tidak memiliki hubungan dengan tingkat keparahan kecelakaan. Temuan tersebut lebih menunjukkan bahwa hubungan antar variabel belum dapat dijelaskan secara kuat oleh data yang tersedia. Kondisi ini diduga dipengaruhi oleh keterbatasan jumlah observasi valid, distribusi kategori yang tidak seimbang, serta tingginya proporsi data yang tidak lengkap pada beberapa variabel penelitian.

Berdasarkan hasil rekapitulasi regresi logistik ordinal tahun 2022 sampai dengan tahun 2024 diketahui bahwa seluruh model penelitian memiliki tingkat *goodness of fit* yang baik karena nilai signifikansi *Pearson* dan *Deviance* berada di atas 0,05. Nilai *Nagelkerke R-Square* tertinggi ditemukan pada tahun 2023 yaitu sebesar 0,36 atau 36,0%. Hal tersebut menunjukkan bahwa model penelitian tahun 2023 memiliki kemampuan penjelasan yang lebih baik dibandingkan tahun lainnya. Sementara itu, model penelitian tahun 2022 dan 2024 menunjukkan kemampuan penjelasan yang relatif rendah dengan nilai *Nagelkerke* masing-masing sebesar 0,14 dan 0,14.

Pada pengujian test of parallel lines, model tahun 2022 dan 2024 memenuhi asumsi proportional odds karena memiliki nilai signifikansi lebih besar dari 0,05. Namun pada tahun 2023, model mengalami masalah konvergensi sehingga pengujian parallel lines tidak dapat dilakukan secara sempurna. Selain itu, seluruh model penelitian menunjukkan bahwa belum terdapat variabel independen yang memiliki pengaruh signifikan terhadap tingkat keparahan kecelakaan lalu lintas secara statistik. Kondisi tersebut dipengaruhi oleh keterbatasan jumlah data valid dan tingginya missing value dalam data penelitian. Secara umum, model regresi logistik ordinal pada penelitian ini masih dapat digunakan sebagai pendekatan eksploratif untuk melihat kecenderungan hubungan antar variabel terhadap tingkat keparahan kecelakaan lalu lintas di wilayah DKI Jakarta. Dengan demikian, hasil penelitian ini lebih menggambarkan kecenderungan hubungan antar variabel terhadap tingkat keparahan kecelakaan dibandingkan sebagai model prediksi yang kuat.

Pembentukan Model Regresi Logistik Ordinal

Berdasarkan hasil evaluasi model Regresi Logistik Ordinal pada data kecelakaan lalu lintas tahun 2022, 2023, dan 2024, diperoleh perbedaan kemampuan model dalam menjelaskan variasi tingkat keparahan kecelakaan. Hasil perbandingan menunjukkan bahwa model tahun 2023 memiliki performa yang lebih baik dibandingkan model tahun 2022 dan 2024, ditunjukkan oleh nilai *Nagelkerke R-Square* sebesar 0,36 atau 36,0%. Nilai tersebut merupakan yang tertinggi dibandingkan model tahun 2022 sebesar 0,14 dan model tahun 2024 sebesar 0,14. Selain memiliki nilai *Nagelkerke R-Square* tertinggi, model tahun 2023 juga menghasilkan nilai signifikansi Model Fitting Information sebesar 0,006 yang lebih kecil dari tingkat signifikansi 0,05. Hasil tersebut menunjukkan bahwa model yang dibentuk mampu memberikan peningkatan kecocokan yang signifikan dibandingkan model yang hanya memuat intersep. Dengan demikian, model tahun 2023 dipandang sebagai model yang paling representatif untuk menjelaskan tingkat keparahan kecelakaan lalu lintas pada penelitian ini.

Berdasarkan hasil estimasi parameter, variabel yang membentuk model tahun 2023 terdiri atas jenis kendaraan (X5), usia korban (X6), dan jenis kelamin korban (X7). Ketiga variabel tersebut merupakan faktor yang memberikan kontribusi dalam menjelaskan perubahan tingkat keparahan kecelakaan lalu lintas pada periode penelitian. Model Regresi Logistik Ordinal yang terbentuk menggunakan fungsi logit kumulatif (*cumulative logit model*) dengan persamaan (1) sebagai berikut:

$$\text{Logit } [P(Y \leq j)] = \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \dots \dots \dots (1)$$

dimana:

Y = tingkat keparahan kecelakaan;
 α_j = nilai threshold pada kategori ke-j;
 β = koefisien regresi;
 X = variabel independen.

Berdasarkan hasil estimasi parameter tahun 2023, diperoleh persamaan (2) sebagai berikut:

$$\text{Logit } [P(Y \leq j)] = \alpha_j - 0,587(X51) - 3,298(X52) - 21,256(X53) + 0,210(X61) - 0,647(X62) + 0,152(X71) \dots\dots (2)$$

dengan nilai threshold:

$$\alpha_1 = -1,51$$

$$\alpha_2 = 0,52$$

$$\alpha_3 = 2,53$$

Persamaan tersebut menunjukkan bahwa perubahan kategori jenis kendaraan, usia korban, dan jenis kelamin korban akan memengaruhi peluang suatu kejadian kecelakaan untuk berada pada kategori tingkat keparahan tertentu. Tanda koefisien positif menunjukkan kecenderungan peningkatan peluang pada kategori tingkat keparahan yang lebih tinggi, sedangkan tanda koefisien negatif menunjukkan kecenderungan penurunan peluang terhadap kategori tingkat keparahan yang lebih tinggi. Model yang terbentuk selanjutnya digunakan sebagai dasar dalam penyusunan Simulasi Risiko tingkat keparahan kecelakaan lalu lintas setelah tahun 2024. Penggunaan model tahun 2023 sebagai dasar Simulasi Risiko dilakukan karena model tersebut memiliki kemampuan prediktif yang lebih baik dibandingkan model tahun 2022 dan 2024 berdasarkan hasil evaluasi statistik yang telah dilakukan sebelumnya.

Pemilihan Model Terbaik dan Simulasi Risiko Keparahannya Kecelakaan Pasca Tahun 2024

Berdasarkan hasil analisis Regresi Logistik Ordinal pada data kecelakaan lalu lintas tahun 2022–2024, model tahun 2023 dipilih sebagai model terbaik karena memiliki nilai *Nagelkerke R-Square* tertinggi sebesar 0,36 dibandingkan model tahun 2022 dan 2024. Selain itu, nilai signifikansi Model Fitting Information sebesar 0,006 menunjukkan bahwa model yang dibentuk mampu menjelaskan tingkat keparahan kecelakaan lebih baik dibandingkan model tanpa variabel prediktor. Hasil tersebut menunjukkan bahwa model tahun 2023 memiliki kemampuan penjelasan yang relatif lebih baik dalam menggambarkan hubungan antara karakteristik korban, jenis kendaraan, dan tingkat keparahan kecelakaan lalu lintas. Persamaan model Regresi Logistik Ordinal tahun 2023 yang diperoleh persamaan (2) sebagai berikut:

$$\text{Logit } [P(Y \leq j)] = \alpha_j - 0,587(X51) - 3,298(X52) - 21,256(X53) + 0,210(X61) - 0,647(X62) + 0,152(X71) \dots\dots (2)$$

dengan nilai threshold:

$$\alpha_1 = -1,516$$

$$\alpha_2 = 0,526$$

$$\alpha_3 = 2,537$$

Meskipun hasil pengujian parameter menunjukkan bahwa belum terdapat variabel yang signifikan secara statistik pada tingkat kepercayaan 95%, model yang terbentuk tetap memberikan informasi mengenai arah dan pola hubungan antarvariabel dalam penelitian. Nilai *Nagelkerke R-Square* sebesar 36,00% menunjukkan bahwa model mampu menjelaskan sebagian variasi tingkat keparahan kecelakaan, sedangkan sisanya dipengaruhi oleh faktor lain yang belum tersedia dalam basis data penelitian, seperti kecepatan kendaraan, kondisi geometrik jalan, kondisi pencahayaan, perilaku pengemudi, penggunaan alat keselamatan, serta faktor lingkungan lainnya.

Untuk menginterpretasikan kondisi pasca tahun 2024, penelitian ini tidak melakukan prediksi jumlah kecelakaan, melainkan melakukan simulasi indeks risiko keparahan berdasarkan model terbaik yang diperoleh. Pendekatan ini dipilih karena model Regresi Logistik Ordinal yang terbentuk bertujuan untuk menjelaskan tingkat keparahan kecelakaan, bukan memprediksi jumlah kejadian kecelakaan pada tahun tertentu. Data tahun 2024 digunakan sebagai kondisi dasar simulasi karena merupakan data terbaru dalam periode penelitian. Berdasarkan hasil analisis deskriptif, kecelakaan lalu lintas pada tahun 2024 masih didominasi oleh pengguna sepeda motor sebesar 75,79%, kelompok usia dewasa sebesar 79,86%, dan korban berjenis kelamin laki-laki sebesar 67,66%. Karakteristik tersebut kemudian disubstitusikan ke dalam koefisien model tahun 2023 untuk memperoleh nilai indeks risiko keparahan dasar.

Perhitungan indeks risiko dilakukan sebagai berikut:

a. Kontribusi Variabel Jenis Kendaraan

$$\begin{aligned} &= (-0,587 \times 0,7579) + (-3,298 \times 0,1623) + (-21,256 \times 0,0710) \\ &= -0,445 - 0,535 - 1,509 \\ &= -2,489 \end{aligned}$$

b. Kontribusi Variabel Usia Korban

$$\begin{aligned} &= (0,210 \times 0,1184) + (-0,647 \times 0,7986) \\ &= 0,025 - 0,517 \\ &= -0,492 \end{aligned}$$

c. Kontribusi Variabel Jenis Kelamin

$$\begin{aligned} &= (0,152 \times 0,6766) \\ &= 0,103 \end{aligned}$$

Sehingga nilai indeks risiko keparahan tahun dasar 2024 diperoleh sebagai berikut:

$$IRK_{2024} = -2,489 - 0,492 + 0,103$$

$$IRK_{2024} = -2,878$$

Nilai indeks risiko sebesar -2,878 digunakan sebagai representasi kondisi dasar tingkat keparahan kecelakaan berdasarkan karakteristik dominan yang terjadi pada tahun 2024. Nilai tersebut bukan merupakan probabilitas maupun jumlah kecelakaan yang diprediksi, melainkan indikator yang menggambarkan posisi relatif risiko keparahan berdasarkan struktur model yang terbentuk. Berdasarkan hasil simulasi tersebut, dapat diinterpretasikan bahwa pola tingkat keparahan kecelakaan pasca tahun 2024 masih berpotensi dipengaruhi oleh dominasi pengguna sepeda motor, kelompok usia dewasa, dan korban berjenis kelamin laki-laki. Apabila komposisi karakteristik tersebut tidak mengalami perubahan yang signifikan pada masa mendatang, maka kecenderungan tingkat keparahan kecelakaan diperkirakan masih mengikuti pola yang relatif sama dengan periode penelitian.

Hasil ini menunjukkan bahwa upaya peningkatan keselamatan lalu lintas perlu difokuskan pada kelompok pengguna jalan yang memiliki tingkat keterpaparan risiko tertinggi. Penguatan edukasi keselamatan berkendara bagi pengguna sepeda motor, peningkatan kepatuhan terhadap peraturan lalu lintas pada kelompok usia produktif, serta peningkatan pengawasan terhadap perilaku berkendara berisiko merupakan langkah yang dapat dilakukan untuk menekan potensi terjadinya kecelakaan dengan tingkat keparahan yang lebih tinggi pada masa mendatang. Dengan demikian, model Regresi Logistik Ordinal yang diperoleh tidak hanya memberikan gambaran mengenai faktor-faktor yang membentuk tingkat keparahan kecelakaan pada periode penelitian, tetapi juga dapat digunakan sebagai dasar untuk memahami kecenderungan risiko keparahan kecelakaan setelah tahun 2024 berdasarkan karakteristik dominan yang terdapat dalam data penelitian.

Pembahasan

Evaluasi Kelayakan dan Kemampuan Penjelasan Model

Hasil uji multikolinearitas menunjukkan bahwa seluruh variabel independen pada model tahun 2022, 2023, dan 2024 mempunyai nilai *tolerance* lebih besar dari 0,10 dan nilai *variance inflation factor* (VIF) jauh di bawah 10. Nilai VIF berada pada rentang 1,02–1,04 pada tahun 2022,

1,04–1,09 pada tahun 2023, dan 1,27–1,54 pada tahun 2024. Hasil tersebut menunjukkan bahwa tidak terdapat korelasi linear yang berlebihan di antara variabel kondisi cuaca, lokasi, jenis kecelakaan, jenis kendaraan, usia, dan jenis kelamin. Dengan demikian, ketidakstabilan atau tidak signifikannya koefisien regresi dalam penelitian ini tidak terutama disebabkan oleh multikolinearitas.

Tidak ditemukannya multikolinearitas merupakan kondisi yang mendukung penggunaan variabel-variabel independen secara bersamaan dalam model. Namun, kondisi tersebut tidak secara otomatis menjamin bahwa model mempunyai kemampuan penjelasan atau prediksi yang tinggi. Kemampuan model regresi keparahan kecelakaan juga dipengaruhi oleh ukuran sampel efektif, distribusi kategori variabel, ketidakseimbangan kategori keparahan, kelengkapan data, ketepatan pengelompokan variabel, dan keberadaan faktor lain yang belum dimasukkan ke dalam model. Literatur menunjukkan bahwa model keparahan kecelakaan sering menghadapi persoalan ketidakseimbangan kelas karena kejadian fatal dan cedera berat umumnya lebih sedikit dibandingkan kejadian dengan tingkat keparahan rendah (Kim et al., 2021; Li et al., 2023; Zhu et al., 2021).

Evaluasi regresi logistik ordinal memperlihatkan adanya perbedaan kemampuan penjelasan antara model tahun 2022, 2023, dan 2024. Model tahun 2023 menghasilkan nilai Nagelkerke R -square sebesar 0,36, sedangkan model tahun 2022 dan 2024 masing-masing menghasilkan nilai sebesar 0,14. Nilai tersebut menunjukkan bahwa model tahun 2023 memiliki daya jelas relatif lebih tinggi dibandingkan dua model lainnya. Namun, nilai Nagelkerke R -square tidak dapat diperlakukan sama dengan koefisien determinasi dalam regresi linear. Nilai 0,36 tidak berarti bahwa model mempunyai ketepatan prediksi sebesar 36%, tetapi mengindikasikan adanya peningkatan kecocokan relatif dibandingkan model yang hanya memuat intersep.

Nilai signifikansi *Model Fitting Information* sebesar 0,006 pada tahun 2023 menunjukkan bahwa model yang memasukkan jenis kendaraan, usia, dan jenis kelamin memberikan peningkatan kecocokan yang bermakna dibandingkan model tanpa prediktor. Sebaliknya, nilai signifikansi sebesar 0,83 pada tahun 2022 dan 0,95 pada tahun 2024 menunjukkan bahwa penambahan prediktor pada kedua tahun tersebut belum menghasilkan peningkatan kecocokan yang signifikan dibandingkan model intersep. Dengan demikian, nilai *goodness of fit* yang berada di atas 0,05 pada model tahunan harus dibedakan dari hasil *Model Fitting Information*. Nilai *goodness of fit* yang tidak signifikan menunjukkan bahwa tidak ditemukan bukti kuat mengenai ketidakcocokan model terhadap data, tetapi tidak membuktikan bahwa model mempunyai kemampuan prediksi yang tinggi.

Model tahun 2023 juga menghadapi ketidakberhasilan konvergensi pada pengujian *parallel lines*. Kondisi tersebut perlu diperhatikan karena dapat mengindikasikan adanya kategori dengan jumlah observasi yang sangat sedikit, distribusi data yang tidak seimbang, pemisahan kategori yang mendekati sempurna, atau estimasi parameter yang belum stabil. Koefisien sebesar $-21,256$ pada salah satu kategori jenis kendaraan turut menunjukkan kemungkinan adanya kategori dengan informasi yang terbatas. Oleh karena itu, model tahun 2023 lebih tepat dinyatakan sebagai model dengan kemampuan penjelasan relatif tertinggi di antara model yang diuji, bukan sebagai model prediksi yang sepenuhnya stabil dan telah tervalidasi.

Hasil penelitian menunjukkan bahwa belum terdapat variabel independen yang berpengaruh signifikan secara individual terhadap tingkat keparahan kecelakaan pada tingkat kepercayaan 95%. Temuan tersebut berbeda dari penelitian Bekelcho et al. (2023), yang menemukan bahwa karakteristik kendaraan, jenis korban, kondisi pencahayaan, jenis tabrakan, dan karakteristik pengemudi berhubungan dengan tingkat keparahan kecelakaan. Lee et al. (2023) juga menemukan bahwa jenis kendaraan, waktu kejadian, jenis jalan, kondisi permukaan jalan, lingkungan terbangun, usia, dan jenis kelamin pengemudi dapat memengaruhi kemungkinan terjadinya kecelakaan yang mengakibatkan cedera atau kematian.

Perbedaan hasil tersebut tidak menunjukkan bahwa temuan penelitian ini bertentangan secara mutlak dengan penelitian terdahulu. Pengaruh suatu variabel terhadap keparahan kecelakaan dapat berbeda menurut karakteristik wilayah, sistem pencatatan kecelakaan, struktur lalu lintas, komposisi pengguna jalan, ukuran sampel, definisi kategori, dan metode analisis. Wang

et al. (2021) menunjukkan bahwa hubungan antara faktor penyebab dan keparahan kecelakaan dapat mengalami ketidakstabilan temporal. Variabel yang berpengaruh pada satu periode belum tentu mempunyai pengaruh yang sama pada periode berikutnya karena terjadi perubahan perilaku perjalanan, karakteristik kendaraan, lingkungan jalan, pola lalu lintas, dan kebijakan keselamatan.

Se et al. (2021) juga menemukan bahwa faktor manusia, kendaraan, jalan, lingkungan, waktu, dan karakteristik spasial tidak selalu menunjukkan pengaruh yang stabil dari tahun ke tahun. Hasil tersebut mendukung pemisahan model tahun 2022, 2023, dan 2024 dalam penelitian ini. Perbedaan nilai Nagelkerke dan komposisi prediktor antartahun dapat dipahami sebagai indikasi bahwa mekanisme keparahan kecelakaan di DKI Jakarta bersifat dinamis, bukan sebagai ketidakkonsistenan penelitian.

Temuan mengenai perbedaan model antartahun juga sejalan dengan Islam dan Mannering (2021), yang menunjukkan adanya perbedaan pengaruh faktor kecelakaan menurut jenis kelamin dan periode pengamatan. Jafari et al. (2025) menemukan bahwa sebagian faktor keparahan kecelakaan sepeda motor bersifat stabil, sedangkan faktor lainnya hanya berpengaruh pada periode tertentu. Oleh sebab itu, pengaruh jenis kendaraan, usia, jenis kelamin, lokasi, cuaca, dan jenis kecelakaan tidak seharusnya diasumsikan konstan sepanjang waktu.

Tidak signifikannya variabel independen dalam penelitian ini juga tidak dapat langsung dimaknai bahwa variabel-variabel tersebut tidak mempunyai hubungan substantif dengan tingkat keparahan. Hasil tidak signifikan menunjukkan bahwa bukti statistik yang tersedia belum cukup kuat untuk menghasilkan estimasi individual yang presisi. Kondisi tersebut dapat terjadi karena ukuran sampel efektif berkurang akibat data tidak lengkap, sedikitnya observasi pada kategori tertentu, distribusi tingkat keparahan yang tidak seimbang, atau terdapat heterogenitas yang belum direpresentasikan dalam model.

Jenis Kendaraan dan Dominasi Pengguna Sepeda Motor

Analisis deskriptif menunjukkan bahwa pengguna sepeda motor mendominasi kecelakaan lalu lintas tahun 2024 dengan proporsi sebesar 75,79%. Besarnya proporsi tersebut menunjukkan bahwa pengguna sepeda motor merupakan kelompok yang paling banyak muncul dalam data kecelakaan. Namun, dominasi tersebut tidak secara langsung membuktikan bahwa sepeda motor menyebabkan tingkat keparahan yang lebih tinggi. Besarnya jumlah korban juga dipengaruhi oleh banyaknya sepeda motor yang beroperasi, intensitas perjalanan, jarak tempuh, pola penggunaan kendaraan, tingkat keterpaparan lalu lintas, dan sistem pelaporan kecelakaan.

Li et al. (2021) menemukan bahwa faktor-faktor yang memengaruhi keparahan pengendara sepeda motor berbeda antara kecelakaan kendaraan tunggal, kecelakaan yang melibatkan dua kendaraan, dan kecelakaan yang melibatkan lebih dari dua kendaraan. Karakteristik pengendara, kondisi jalan, situasi sebelum tabrakan, dan konfigurasi tabrakan dapat menghasilkan tingkat keparahan yang berbeda. Temuan tersebut menunjukkan bahwa pengelompokan jenis kendaraan saja belum tentu cukup untuk menggambarkan mekanisme keparahan apabila tidak disertai informasi mengenai posisi korban, kendaraan lawan, arah tumbukan, penggunaan helm, kecepatan, dan jenis tabrakan.

Zahid et al. (2024) menunjukkan bahwa faktor-faktor yang berhubungan dengan keparahan kecelakaan sepeda motor dapat berbeda menurut kelompok usia pengendara. Hasil tersebut menegaskan bahwa jenis kendaraan dan usia tidak selalu bekerja secara terpisah, tetapi dapat berinteraksi dengan karakteristik jalan, waktu kejadian, perilaku pengendara, dan kondisi lingkungan. Tidak signifikannya jenis kendaraan dalam penelitian ini dapat disebabkan oleh kategori kendaraan yang masih luas dan belum sepenuhnya menggambarkan heterogenitas pengguna serta konfigurasi kecelakaan.

Se et al. (2024) menemukan adanya perbedaan faktor keparahan kecelakaan sepeda motor antara lajur utama, lajur samping, dan tipe lajur jalan lainnya. Temuan tersebut menunjukkan bahwa pengaruh penggunaan sepeda motor tidak dapat dilepaskan dari lingkungan jalan tempat kecelakaan terjadi. Jenis kendaraan yang sama dapat menghasilkan tingkat keparahan yang berbeda ketika beroperasi pada karakteristik jalan, kecepatan, kepadatan, dan pola konflik lalu lintas yang berbeda.

Dengan demikian, dominasi pengguna sepeda motor dalam penelitian ini sebaiknya dimaknai sebagai dasar untuk menentukan kelompok prioritas dalam program keselamatan, bukan sebagai bukti kausal bahwa sepeda motor secara mandiri meningkatkan keparahan. Untuk mengukur risiko yang lebih akurat, diperlukan ukuran keterpaparan, seperti jumlah sepeda motor, volume perjalanan, jarak tempuh, atau jumlah perjalanan menurut jenis kendaraan.

Usia dan Jenis Kelamin Korban

Kelompok usia dewasa mendominasi data kecelakaan tahun 2024 dengan proporsi sebesar 79,86%. Proporsi tersebut menunjukkan bahwa sebagian besar korban berasal dari kelompok usia yang mempunyai mobilitas tinggi. Akan tetapi, dominasi kelompok usia dewasa tidak secara langsung menunjukkan bahwa kelompok tersebut mempunyai probabilitas keparahan paling tinggi. Jumlah korban yang besar dapat berkaitan dengan tingginya aktivitas perjalanan untuk bekerja, pendidikan, kegiatan ekonomi, dan kepentingan rumah tangga.

Lee et al. (2023) menunjukkan bahwa hubungan usia dengan keparahan kecelakaan tidak selalu bersifat linear. Pengaruh usia dapat berubah pada rentang usia tertentu dan berinteraksi dengan kondisi jalan, jenis kendaraan, lingkungan terbangun, dan karakteristik lokasi. Pengelompokan usia yang terlalu luas dapat menyamarkan variasi antara pengendara dewasa muda, dewasa menengah, dan lanjut usia. Oleh karena itu, tidak signifikannya usia dalam penelitian ini mungkin berkaitan dengan penggunaan kategori usia yang belum cukup rinci untuk menangkap pola nonlinear.

Nickkar et al. (2024) menunjukkan bahwa variasi keparahan menurut usia dan jenis kelamin dapat mengandung heteroskedastisitas, yaitu perbedaan variasi yang tidak seragam pada setiap kelompok. Model ordinal standar yang mengasumsikan variasi seragam berpotensi belum sepenuhnya menangkap perbedaan tersebut. Temuan tersebut dapat menjelaskan mengapa usia dan jenis kelamin dalam penelitian ini belum menunjukkan pengaruh individual yang signifikan meskipun keduanya memiliki pola distribusi yang berbeda.

Korban laki-laki mendominasi data tahun 2024 dengan proporsi sebesar 67,66%. Dominasi tersebut dapat berkaitan dengan intensitas berkendara, jenis pekerjaan, durasi perjalanan, kepemilikan kendaraan, serta keterpaparan lalu lintas yang lebih besar. Dengan demikian, besarnya jumlah korban laki-laki tidak secara otomatis berarti bahwa jenis kelamin laki-laki meningkatkan tingkat keparahan setelah faktor lainnya dikendalikan.

Islam dan Mannering (2021) menemukan bahwa perbedaan keparahan antara pengemudi laki-laki dan perempuan bergantung pada kondisi kecelakaan dan dapat berubah sepanjang periode pengamatan. Pengaruh jenis kelamin dapat berinteraksi dengan kecepatan, kondisi cuaca, penggunaan alat keselamatan, jenis kendaraan, waktu berkendara, konsumsi alkohol, dan perilaku berisiko. Oleh sebab itu, tidak signifikannya jenis kelamin dalam penelitian ini dapat mengindikasikan bahwa variasi perilaku dan keterpaparan belum sepenuhnya terwakili oleh variabel yang tersedia.

Kondisi Cuaca, Lokasi, dan Jenis Kecelakaan

Kondisi cuaca, lokasi, dan jenis kecelakaan belum menunjukkan pengaruh yang signifikan dalam model tahunan yang memasukkan variabel-variabel tersebut. Meskipun demikian, ketiga variabel tetap mempunyai dasar teoritis yang kuat. Kondisi cuaca dapat memengaruhi jarak pandang, daya cengkeram permukaan jalan, jarak pengereman, kestabilan kendaraan, dan keputusan pengemudi. Akan tetapi, pengaruh cuaca juga ditentukan oleh intensitas hujan, waktu kejadian, kondisi pencahayaan, kualitas drainase, kecepatan kendaraan, dan kondisi permukaan jalan.

Kategori cuaca yang terlalu umum dapat menggabungkan kondisi dengan tingkat bahaya yang berbeda. Misalnya, hujan ringan, hujan lebat, genangan air, dan kondisi sesaat setelah hujan mungkin ditempatkan dalam kategori yang sama meskipun mempunyai konsekuensi keselamatan yang berbeda. Jafari et al. (2025) menunjukkan bahwa pengaruh cuaca dan kondisi lingkungan terhadap keparahan kecelakaan sepeda motor dapat berubah secara temporal dan mengandung heterogenitas yang tidak teramati.

Lokasi kecelakaan juga merupakan variabel yang bersifat multidimensional. Satu kategori lokasi dapat mencakup perbedaan bentuk geometrik jalan, jumlah lajur, keberadaan

persimpangan, fungsi kawasan, penerangan, kepadatan aktivitas, fasilitas pejalan kaki, dan kecepatan operasional. Lee et al. (2023) menunjukkan bahwa karakteristik lingkungan terbangun dan kondisi jalan berkontribusi terhadap perbedaan keparahan kecelakaan. Oleh karena itu, lokasi yang diklasifikasikan secara umum mungkin belum cukup sensitif untuk menggambarkan variasi risiko di dalam setiap kategori.

Jenis kecelakaan juga dapat mencakup perbedaan arah tumbukan, lawan tabrakan, energi benturan, posisi korban, dan jumlah kendaraan yang terlibat. Li et al. (2021) memperlihatkan adanya perbedaan faktor keparahan antara kecelakaan sepeda motor tunggal dan kecelakaan yang melibatkan beberapa kendaraan. Pengelompokan jenis kecelakaan yang terlalu luas dapat mengurangi kemampuan model dalam membedakan mekanisme keparahan pada setiap konfigurasi tabrakan.

Ketidaksignifikanan Variabel dan Daya Penjelasan Model

Nilai Nagelkerke R -square sebesar 0,14 pada model tahun 2022 dan 2024 menunjukkan bahwa variabel yang tersedia baru memberikan daya jelas yang relatif terbatas. Kondisi tersebut dapat dipahami karena keparahan kecelakaan merupakan hasil interaksi yang kompleks antara manusia, kendaraan, jalan, lingkungan, energi tumbukan, penggunaan alat keselamatan, dan penanganan medis setelah kecelakaan. Ketika sebagian faktor tersebut tidak tersedia, variasi keparahan yang mampu dijelaskan oleh model menjadi lebih rendah.

Model tahun 2023 mempunyai nilai Nagelkerke yang lebih tinggi, tetapi belum menghasilkan prediktor individual yang signifikan. Kondisi ini dapat terjadi ketika penambahan prediktor memperbaiki kecocokan model secara keseluruhan, tetapi simpangan baku koefisien masih besar sehingga estimasi masing-masing parameter belum cukup presisi. Kategori dengan jumlah observasi sangat sedikit juga dapat menghasilkan koefisien yang ekstrem dan interval kepercayaan yang lebar.

Ketidakseimbangan kategori tingkat keparahan merupakan salah satu faktor yang perlu dipertimbangkan. Data kecelakaan biasanya mempunyai jumlah kejadian ringan yang lebih besar dibandingkan kecelakaan berat atau fatal. Kim et al. (2021) menunjukkan bahwa sifat ordinal dan ketidakseimbangan kelas perlu ditangani secara bersamaan dalam pemodelan keparahan. Zhu et al. (2021) juga menegaskan bahwa metode klasifikasi yang mengabaikan urutan kategori dapat menghasilkan kesalahan klasifikasi yang lebih besar, terutama pada kategori yang berjauhan tingkat keparahannya.

Li et al. (2023) menemukan bahwa penyeimbangan distribusi data dapat meningkatkan kemampuan model dalam mengenali kategori keparahan yang jarang ditemukan. Jika ketidakseimbangan kelas tidak diperhitungkan, model dapat mempunyai kecocokan keseluruhan yang cukup baik, tetapi lebih mampu mengenali kategori dominan daripada kecelakaan berat atau fatal. Oleh karena itu, evaluasi model tidak cukup hanya menggunakan kecocokan keseluruhan, tetapi juga perlu memperhatikan sensitivitas atau *recall* pada setiap kategori keparahan.

Missing value juga berpotensi mengurangi ukuran sampel efektif dan mengubah komposisi data yang dianalisis. Apabila observasi yang tidak lengkap dikeluarkan, hasil model hanya merepresentasikan kejadian dengan data lengkap. Gao et al. (2025) menunjukkan bahwa pemilihan metode imputasi pada data kecelakaan yang tidak lengkap dapat memengaruhi kemampuan model dalam memprediksi tingkat keparahan. Penanganan data hilang perlu mempertimbangkan pola kehilangan data, jenis variabel, dan kemungkinan hubungan antara ketidaklengkapan data dengan karakteristik kecelakaan.

Berdasarkan kondisi tersebut, tidak signifikkannya variabel dalam penelitian ini lebih tepat ditafsirkan sebagai belum cukup kuatnya bukti statistik pada data yang tersedia. Temuan tersebut tidak membuktikan bahwa kondisi cuaca, lokasi, jenis kecelakaan, jenis kendaraan, usia, dan jenis kelamin tidak relevan terhadap keparahan kecelakaan.

Pemaknaan Simulasi Indeks Risiko Keparahahan

Model tahun 2023 digunakan sebagai dasar simulasi karena mempunyai nilai Nagelkerke R -square dan *Model Fitting Information* yang relatif lebih baik dibandingkan model tahun 2022 dan 2024. Substitusi komposisi jenis kendaraan, usia, dan jenis kelamin tahun 2024 ke dalam koefisien model menghasilkan indeks risiko keparahan sebesar $-2,878$.

Nilai $-2,878$ bukan merupakan probabilitas terjadinya kecelakaan, jumlah kecelakaan yang diprediksi, atau probabilitas individual mengalami kategori keparahan tertentu. Nilai tersebut merupakan skor komposit relatif yang dibentuk berdasarkan koefisien model dan proporsi kategori dalam data tahun 2024. Oleh karena itu, istilah yang lebih tepat digunakan adalah indeks atau skor skenario eksploratif, bukan nilai prediksi absolut.

Dalam regresi logistik ordinal, probabilitas setiap kategori keparahan pada dasarnya dihitung melalui fungsi logit kumulatif untuk setiap kombinasi karakteristik observasi dan setiap nilai *threshold*. Penggunaan proporsi agregat sebagai pengganti karakteristik individual tidak secara langsung menghasilkan probabilitas populasi. Meskipun demikian, indeks tersebut tetap mempunyai kegunaan sebagai kondisi dasar internal yang dapat dibandingkan dengan skenario perubahan komposisi pengguna jalan.

Penggunaan model tahun 2023 untuk menggambarkan periode setelah tahun 2024 juga perlu memperhatikan kemungkinan ketidakstabilan temporal. Wang et al. (2021), Se et al. (2021), dan Jafari et al. (2025) menunjukkan bahwa koefisien model keparahan dapat berubah antarperiode. Oleh karena itu, indeks $-2,878$ dapat digunakan sebagai *baseline* eksploratif, tetapi perlu dikalibrasi dan diuji kembali ketika data kecelakaan setelah tahun 2024 telah tersedia.

Penelitian ini memperkuat pemahaman bahwa tingkat keparahan kecelakaan merupakan hasil interaksi banyak faktor dan tidak selalu dapat dijelaskan oleh pengaruh langsung satu variabel. Jenis kendaraan, usia, jenis kelamin, kondisi cuaca, lokasi, dan jenis kecelakaan dapat berinteraksi dengan karakteristik lain yang tidak tercatat. Hasil penelitian juga menunjukkan pentingnya mempertimbangkan heterogenitas dan ketidakstabilan temporal dalam analisis keparahan kecelakaan. Perbedaan hasil model tahun 2022, 2023, dan 2024 memberikan indikasi bahwa hubungan antarvariabel dapat berubah sepanjang waktu. Temuan tersebut mendukung pengembangan model yang lebih fleksibel, seperti *partial proportional odds model*, *generalized ordered logit*, *random-parameter logit*, *latent-class model*, atau model yang mengakomodasi heterogenitas dalam nilai rata-rata dan varians.

Pemisahan model berdasarkan tahun merupakan langkah yang relevan karena memungkinkan identifikasi perubahan pola hubungan antarfaktor. Penelitian selanjutnya dapat melengkapi pendekatan tersebut dengan pengujian transferabilitas parameter, validasi temporal, dan validasi menggunakan data di luar sampel. Santos et al. (2022) menegaskan bahwa model statistik dan pembelajaran mesin mempunyai keunggulan yang berbeda. Model statistik lebih kuat dalam memberikan interpretasi parameter, sedangkan beberapa metode pembelajaran mesin dapat memberikan performa klasifikasi yang lebih tinggi ketika terdapat hubungan nonlinear dan interaksi yang kompleks. Evaluasi model selanjutnya perlu memasukkan ukuran kinerja yang sensitif terhadap ketidakseimbangan data, seperti *precision*, *recall*, *F1-score*, *balanced accuracy*, dan area di bawah kurva *precision-recall*. Penggunaan akurasi keseluruhan saja dapat menghasilkan kesimpulan yang terlalu optimistis apabila kategori ringan sangat mendominasi data. Penanganan missing value dapat dikembangkan melalui analisis pola kehilangan data dan metode imputasi yang sesuai. Penggunaan imputasi sebaiknya disertai analisis sensitivitas untuk membandingkan hasil model sebelum dan setelah imputasi. Langkah tersebut dapat membantu menentukan apakah kesimpulan penelitian stabil terhadap metode penanganan data yang berbeda.

Hasil penelitian menunjukkan perlunya peningkatan kelengkapan dan integrasi basis data kecelakaan. Data kepolisian dapat dihubungkan dengan data rumah sakit, kondisi jalan, cuaca, volume lalu lintas, serta karakteristik kendaraan. Variabel tambahan yang penting antara lain kecepatan kendaraan, penggunaan helm atau sabuk keselamatan, konsumsi alkohol atau obat, kepemilikan surat izin mengemudi, kondisi pencahayaan, bentuk geometrik jalan, kondisi permukaan, posisi korban, dan waktu respons medis. Sistem pencatatan juga perlu menggunakan definisi kategori yang konsisten antarpetugas dan antarperiode. Formulir digital dengan validasi otomatis, variabel wajib, pilihan kategori terstandar, dan pemeriksaan kelengkapan dapat mengurangi missing value. Peningkatan kualitas pencatatan akan memperkuat kemampuan model sekaligus mendukung evaluasi program keselamatan secara lebih objektif.

Dominasi pengguna sepeda motor, kelompok usia dewasa, dan korban laki-laki dapat digunakan sebagai dasar awal untuk menentukan kelompok sasaran program keselamatan.

Namun, prioritas intervensi sebaiknya tidak hanya didasarkan pada jumlah atau persentase korban. Penentuan prioritas juga perlu mempertimbangkan jumlah pengguna, volume perjalanan, jarak tempuh, lokasi kejadian, jenis tabrakan, dan tingkat keparahan.

Program keselamatan bagi pengendara sepeda motor dapat diarahkan pada peningkatan penggunaan helm berstandar, pengendalian kecepatan, peningkatan visibilitas, pemeriksaan kelayakan kendaraan, dan pengawasan pelanggaran berisiko. Intervensi untuk kelompok usia produktif dapat disesuaikan dengan pola perjalanan menuju tempat kerja, perjalanan pada malam hari, dan kawasan dengan mobilitas tinggi. Penanganan berbasis lokasi juga perlu dikembangkan melalui identifikasi titik atau koridor dengan konsentrasi kecelakaan berat. Intervensi dapat mencakup peningkatan penerangan, pengendalian kecepatan, penataan akses, perbaikan marka dan rambu, perlindungan pengguna jalan rentan, dan pengurangan konflik pada persimpangan. Kebijakan yang dihasilkan perlu mengombinasikan hasil model statistik dengan audit keselamatan jalan dan observasi lapangan.

Penelitian ini menggunakan data kecelakaan yang tercatat dalam sistem administrasi sehingga hasilnya merepresentasikan karakteristik kejadian yang terdokumentasi selama tahun 2022–2024. Penggunaan data aktual memberikan dasar empiris yang kuat untuk menggambarkan pola kecelakaan di DKI Jakarta. Namun, sejumlah faktor yang secara teoretis berkaitan dengan keparahan, seperti kecepatan, penggunaan alat keselamatan, perilaku pengemudi, bentuk geometrik jalan, kondisi pencahayaan, dan penanganan medis, belum tersedia secara konsisten. Penambahan variabel tersebut pada penelitian selanjutnya berpotensi memperluas kemampuan penjelasan model. Jumlah observasi valid dan distribusi kategori berbeda pada setiap tahun. Kondisi tersebut merupakan karakteristik yang umum ditemukan pada data kecelakaan administratif, terutama karena kecelakaan berat dan fatal mempunyai frekuensi lebih rendah. Penelitian ini telah merespons kondisi tersebut dengan mengevaluasi model secara terpisah untuk setiap tahun dan menempatkan hasilnya sebagai analisis kecenderungan, bukan sebagai kesimpulan kausal yang bersifat absolut.

Model tahun 2023 mempunyai daya jelas relatif tertinggi, tetapi pengujian *parallel lines* belum mencapai konvergensi. Oleh karena itu, interpretasi model dibatasi pada arah hubungan dan pembentukan skor eksploratif. Pembatasan tersebut merupakan bentuk kehati-hatian analitis agar kesimpulan tetap proporsional terhadap informasi yang tersedia. Validasi menggunakan data setelah tahun 2024 dapat digunakan untuk menilai kestabilan koefisien dan konsistensi hasil model. Penelitian ini belum menggunakan ukuran keterpaparan, seperti jumlah kendaraan, volume perjalanan, atau jarak tempuh menurut kelompok pengguna jalan. Akibatnya, proporsi korban menggambarkan distribusi kejadian yang tercatat dan belum dapat diartikan sebagai risiko per perjalanan. Meskipun demikian, hasil penelitian tetap memberikan dasar penting untuk mengidentifikasi kelompok dominan dan menentukan kebutuhan pengembangan data berbasis keterpaparan.

Simulasi indeks risiko menggunakan komposisi agregat tahun 2024 sehingga tidak dapat diperlakukan sebagai probabilitas individual. Pendekatan tersebut digunakan untuk membentuk kondisi dasar yang dapat dibandingkan dengan skenario perubahan komposisi. Penelitian lanjutan dapat memperkuat simulasi melalui perhitungan probabilitas pada tingkat individu, *bootstrap*, analisis sensitivitas, dan validasi temporal. Secara keseluruhan, keterbatasan tersebut tidak mengurangi kontribusi utama penelitian. Penelitian ini memberikan gambaran empiris mengenai dinamika model keparahan kecelakaan selama tiga tahun, menunjukkan perbedaan kemampuan penjelasan antarmodel, mengidentifikasi persoalan kelengkapan serta ketidakseimbangan data, dan menyediakan dasar eksploratif bagi pengembangan kebijakan keselamatan lalu lintas berbasis data.

KESIMPULAN DAN SARAN

Penelitian ini bertujuan menganalisis faktor-faktor yang berkaitan dengan tingkat keparahan kecelakaan lalu lintas di DKI Jakarta pada periode 2022–2024, membentuk model regresi logistik ordinal, serta menyusun simulasi indeks risiko keparahan berdasarkan model dengan kemampuan penjelasan relatif terbaik. Hasil penelitian menunjukkan bahwa seluruh variabel independen tidak

mengalami masalah multikolinearitas, tetapi belum ditemukan variabel yang berpengaruh signifikan secara individual terhadap tingkat keparahan kecelakaan pada tingkat kepercayaan 95%. Model tahun 2023 memiliki daya jelas relatif lebih tinggi dibandingkan model tahun 2022 dan 2024, ditunjukkan oleh nilai Nagelkerke *R-square* sebesar 0,36 dan nilai *Model Fitting Information* sebesar 0,006. Meskipun demikian, tidak tercapainya konvergensi pada pengujian *parallel lines* menunjukkan bahwa model tersebut perlu ditafsirkan secara hati-hati dan lebih tepat digunakan sebagai model eksploratif daripada sebagai model prediksi yang kuat. Simulasi yang dilakukan dengan menggunakan komposisi karakteristik tahun 2024 menghasilkan indeks risiko keparahan sebesar -2,878. Nilai tersebut tidak merepresentasikan probabilitas atau jumlah kecelakaan, tetapi berfungsi sebagai skor dasar relatif yang menggambarkan kecenderungan risiko berdasarkan dominasi pengguna sepeda motor, kelompok usia dewasa, dan korban laki-laki dalam data penelitian. Secara keseluruhan, penelitian ini menunjukkan bahwa tingkat keparahan kecelakaan merupakan hasil interaksi yang kompleks antara faktor manusia, kendaraan, jalan, dan lingkungan, sehingga tidak dapat dijelaskan secara memadai hanya melalui beberapa variabel administratif yang tersedia.

Penelitian selanjutnya disarankan menggunakan basis data yang lebih lengkap dan terintegrasi dengan memasukkan variabel kecepatan kendaraan, penggunaan helm atau sabuk keselamatan, kondisi geometrik jalan, pencahayaan, kondisi permukaan, perilaku pengemudi, volume lalu lintas, tingkat keterpaparan, serta waktu penanganan medis. Pengembangan model juga dapat dilakukan melalui *partial proportional odds model*, *generalized ordered logit*, *random-parameter model*, atau pendekatan pembelajaran mesin yang mampu mengakomodasi ketidakseimbangan kategori dan heterogenitas temporal. Bagi praktisi keselamatan lalu lintas, hasil penelitian dapat digunakan sebagai dasar awal untuk memprioritaskan program edukasi, pengawasan, dan intervensi keselamatan bagi pengguna sepeda motor serta kelompok usia produktif, dengan tetap mempertimbangkan lokasi, jenis tabrakan, dan tingkat keterpaparan. Pada tingkat kebijakan, diperlukan penguatan sistem pencatatan kecelakaan yang terstandar, integrasi data kepolisian dengan rumah sakit dan instansi transportasi, serta penerapan intervensi berbasis lokasi melalui pengendalian kecepatan, peningkatan penerangan, perbaikan marka dan rambu, serta penanganan titik konflik. Validasi menggunakan data setelah tahun 2024 juga diperlukan agar kestabilan koefisien, konsistensi pola, dan kegunaan indeks risiko dapat dievaluasi secara lebih kuat.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada seluruh pihak yang telah memberikan dukungan, bantuan, dan kontribusi selama proses pelaksanaan penelitian hingga penyusunan artikel ini, khususnya kepada instansi penyedia data, serta pihak-pihak terkait yang telah membantu kelancaran penelitian.

DAFTAR PUSTAKA

- Basri, B., Budiharto, B., & Iswanto, B. T. (2021). *The Characteristics of Traffic Accident in Indonesia*. 4–8. <https://doi.org/10.4108/eai.18-11-2020.2311695>
- Bekelcho, T., Olani, A. B., Woldemeskel, A., Alemayehu, M., & Guta, G. (2023). Identification of determinant factors for crash severity levels occurred in Addis Ababa City, Ethiopia, from 2017 to 2020: Using ordinal logistic regression model approach. *BMC Public Health*, 23, Article 1884. <https://doi.org/10.1186/s12889-023-16785-3>
- Darwish, H. H. (2021). *Ordinal regression analysis of traffic collision accidents in Jordan 2021 : factors and severity assessment* Ordinal regression analysis of traffic collision accidents in Jordan 2021 : factors and severity assessment. April 2023. <https://doi.org/10.22306/al.v10i2.390>
- Darwish, W., & Darwish, H. H. (2023). Ordinal regression analysis of traffic collision accidents in Jordan 2021: factors and severity assessment. *Acta Logistica*, 10(2), 241–250. <https://doi.org/10.22306/al.v10i2.390>

- Gao, X., Jiang, X., Zhuang, D., Haworth, J., Wang, S., Ilyankou, I., & Chen, H. (2025). Reliable imputation of incomplete crash data for predicting driver injury severity. *Accident Analysis & Prevention*, 216, Article 108020. <https://doi.org/10.1016/j.aap.2025.108020>
- Irfan, R., Sartavie, A., Arfan, A., Cahyo, D., Anwar, S., Studi, P., Informatika, T., Komputer, F. I., Unggul, U. E., Jeruk, K., Barat, J., Autocorrelation, I. S., Density, K., & Lintas, L. (2021). *Implementasi Kernel Density Pada Analisa Daerah Provinsi Dki Jakarta*. 159–168.
- Islam, M., & Mannering, F. (2021). The role of gender and temporal instability in driver-injury severities in crashes caused by speeds too fast for conditions. *Accident Analysis & Prevention*, 153, Article 106039. <https://doi.org/10.1016/j.aap.2021.106039>
- Jafari, M., Starewich, M. J., Hossain, A., Barua, S., Alnawmasi, N., Ye, X., & Das, S. (2025). Assessing motorcyclist injury severity on curved road segments with temporal dynamics and unobserved heterogeneity. *Scientific Reports*, 15, Article 13110. <https://doi.org/10.1038/s41598-025-97972-7>
- Kusumastutie, N. S., Patria, B., Kusrohmaniah, S., & Hastjarjo, T. D. (2024). A review of accident data for traffic safety studies in Indonesia. *IOP Conference Series: Earth and Environmental Science*, 1294(1). <https://doi.org/10.1088/1755-1315/1294/1/012012>
- Kim, S., Lym, Y., & Kim, K.-J. (2021). Developing crash severity model handling class imbalance and implementing ordered nature: Focusing on elderly drivers. *International Journal of Environmental Research and Public Health*, 18(4), Article 1966. <https://doi.org/10.3390/ijerph18041966>
- Lee, D., Guldman, J.-M., & von Rabenau, B. (2023). Impact of driver's age and gender, built environment, and road conditions on crash severity: A logit modeling approach. *International Journal of Environmental Research and Public Health*, 20(3), Article 2338. <https://doi.org/10.3390/ijerph20032338>
- Li, J., Fang, S., Guo, J., Fu, T., & Qiu, M. (2021). A motorcyclist-injury severity analysis: A comparison of single-, two-, and multi-vehicle crashes using latent class ordered probit model. *Accident Analysis & Prevention*, 151, Article 105953. <https://doi.org/10.1016/j.aap.2020.105953>
- Li, Y., Yang, Z., Xing, L., Yuan, C., Liu, F., Wu, D., & Yang, H. (2023). Crash injury severity prediction considering data imbalance: A Wasserstein generative adversarial network with gradient penalty approach. *Accident Analysis & Prevention*, 192, Article 107271. <https://doi.org/10.1016/j.aap.2023.107271>
- Nickkar, A., Pourfalatoun, S., Miller, E. E., & Lee, Y.-J. (2024). Applying the heteroskedastic ordered probit model on injury severity for improved age and gender estimation. *Traffic Injury Prevention*, 25(2), 202–209. <https://doi.org/10.1080/15389588.2023.2286429>
- O'Donnell, C. J., & Connor, D. H. (1996). Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. *Accident Analysis and Prevention*, 28(6), 739–753. [https://doi.org/10.1016/S0001-4575\(96\)00050-4](https://doi.org/10.1016/S0001-4575(96)00050-4)
- Puspasari, M. A., Madani, S. T., Iqbal, B. M., Muslim, E., Sanjaya, B. P., Pribadyo, C. Y. P., Junistya, K. N., Ghanny, A., Syaifullah, D. H., & Arista, S. A. (2023). Effect of Distraction and Driving Behaviour to Traffic Accidents in Jakarta Using Partial Least Squares Structural Equation Modeling (PLS-SEM). *International Journal of Technology*, 14(7), 1548–1559. <https://doi.org/10.14716/ijtech.v14i7.6676>
- Ricardianto, P., Hidayat, S., Manik, P., Widiyanto, P., & Susanto, P. C. (2021). *International Journal of Research in Commerce and Management Studies GUIDELINES FOR MEASURING THE SUCCESS OF TRAFFIC SAFETY ACTION PLAN*. 3(01), 37–50.
- Santosa, S. P. (2017). *Anatomy of injuries*. 17, 302.
- Shiran, G., Imaninasab, R., & Khayamim, R. (2021). Crash severity analysis of highways based on multinomial logistic regression model, decision tree techniques and artificial neural network: A modeling comparison. *Sustainability (Switzerland)*, 13(10). <https://doi.org/10.3390/su13105670>
- Siregar, M. L., Mauramdha, G., & Adwitiya, R. I. (2024). Benefit-cost analysis for signalised pedestrian crossing implementation in Jakarta, Indonesia. *International Journal of*

- Multidisciplinary Research and Growth Evaluation*, 5(4), 91–96. <https://doi.org/10.54660/ijmrge.2024.5.4.91-96>
- Soehodho, S. (2009). Road accidents in Indonesia. *IATSS Research*, 33(2), 122–124. [https://doi.org/10.1016/S0386-1112\(14\)60250-8](https://doi.org/10.1016/S0386-1112(14)60250-8)
- Summary, C. (2007). *Country Summary*. 2021, 1–7.
- Santos, K., Dias, J. P., & Amado, C. (2022). A literature review of machine learning algorithms for crash injury severity prediction. *Journal of Safety Research*, 80, 254–269. <https://doi.org/10.1016/j.jsr.2021.12.007>
- Se, C., Champahom, T., Jomnonkwao, S., Karoonsoontawong, A., & Ratanavaraha, V. (2021). Temporal stability of factors influencing driver-injury severities in single-vehicle crashes: A correlated random parameters with heterogeneity in means and variances approach. *Analytic Methods in Accident Research*, 32, Article 100179. <https://doi.org/10.1016/j.amar.2021.100179>
- Se, C., Champahom, T., Wisutwattanasak, P., Jomnonkwao, S., Chanpariyavatevong, K., & Ratanavaraha, V. (2024). Modeling of motorcyclist injury severities: A comparison between crashes on main-, frontage-, and standard-lane of roadway. *IATSS Research*, 48(3), 288–298. <https://doi.org/10.1016/j.iatssr.2024.06.005>
- World Health Organization. (2023). *Global status report on road safety 2023*.
- Wang, K., Shirani-Bidabadi, N., Shaon, M. R. R., Zhao, S., & Jackson, E. (2021). Correlated mixed logit modeling with heterogeneity in means for crash severity and surrogate measure with temporal instability. *Accident Analysis & Prevention*, 160, Article 106332. <https://doi.org/10.1016/j.aap.2021.106332>
- Zahid, M., Habib, M. F., Ijaz, M., Ameer, I., Ullah, I., Ahmed, T., & He, Z. (2024). Factors affecting injury severity in motorcycle crashes: Different age groups analysis using CatBoost and SHAP techniques. *Traffic Injury Prevention*, 25(3), 472–481. <https://doi.org/10.1080/15389588.2023.2297168>
- Zhu, S., Wang, K., & Li, C. (2021). Crash injury severity prediction using an ordinal classification machine learning approach. *International Journal of Environmental Research and Public Health*, 18(21), Article 11564. <https://doi.org/10.3390/ijerph182111564>