

Prediksi Titik Kritis Jaringan KRL Jabodetabek Menggunakan Node2Vec dan LightGBM

¹Siti Zulhi Nirma Saidah, ²Yuyun Umaidah, ³Betha Nurina Sari

^{1,2,3}Program Studi Informatika, Universitas Singaperbangsa Karawang, Karawang

E-mail: ¹2210631170103@student.unsika.ac.id, ²yuyun.umaidah@staff.unsika.ac.id, ³betha.nurina@staff.unsika.ac.id

ABSTRAK

KRL Commuter Line Jabodetabek melayani jutaan penumpang setiap hari dengan jaringan yang mencakup 94 stasiun aktif dan 421 relasi layanan langsung. Pada jaringan dengan keterkaitan tinggi seperti ini, terdapat stasiun-stasiun yang bersifat kritis, yaitu stasiun yang apabila mengalami gangguan akan menyebabkan dampak terbesar terhadap konektivitas dan efisiensi waktu tempuh jaringan secara keseluruhan. Penelitian ini membangun model prediktif berbasis Node2Vec dan LightGBM dalam kerangka Knowledge Discovery in Databases (KDD) untuk memperkirakan tingkat kekritisian setiap stasiun secara numerik. Data jadwal operasional diperoleh dari API Comuline (30.989 rekaman), diproses menjadi directed weighted graph (94 simpul, 421 sisi), kemudian digunakan untuk menghitung impact score melalui simulasi single-node removal sebagai label regresi. Sebanyak 38 fitur digunakan: 6 metrik topologi dan 32 dimensi embedding Node2Vec. Evaluasi 5-Fold Cross-Validation menghasilkan MAE 0,3152%, RMSE 0,8903%, dan korelasi Spearman 0,7802 ($p < 0,001$) dengan Bootstrap CI 95% [0,6622; 0,8700]. Model mengidentifikasi 9 dari 10 stasiun paling kritis (Top-10 recall 90%). Betweenness centrality teridentifikasi sebagai fitur paling dominan berdasarkan analisis Gain Importance dan SHAP. Dari 94 stasiun, 8 dikategorikan Sangat Kritis, 21 Kritis, 20 Cukup Kritis, dan 45 Relatif Aman. Manggarai dan Tanah Abang merupakan stasiun paling kritis dengan impact score 9,60% dan 8,89%.

Kata kunci: Graf Operasional; Impact Score; LightGBM; Node2Vec; Titik Kritis Jaringan

ABSTRACT

The Jabodetabek Commuter Rail (KRL) serves millions of passengers daily across a network of 94 active stations and 421 direct service relations. In highly interconnected networks, certain stations are critical, defined as those whose disruption causes the greatest impact on network connectivity and average travel time efficiency. This study develops a predictive model based on Node2Vec and LightGBM under the Knowledge Discovery in Databases (KDD) framework to numerically estimate station criticality levels. Operational schedule data were collected via the Comuline API (30,989 records), processed into a directed weighted graph (94 nodes, 421 edges), and used to compute impact scores through single-node removal simulation as regression labels. A total of 38 features were used: 6 topology metrics and 32 Node2Vec embedding dimensions. 5-Fold Cross-Validation yielded MAE 0.3152%, RMSE 0.8903%, and Spearman correlation 0.7802 ($p < 0.001$) with 95% Bootstrap CI [0.6622; 0.8700]. The model identified 9 of 10 most critical stations (Top-10 recall 90%). Betweenness centrality was the most dominant predictor based on Gain Importance and SHAP analysis. Of 94 stations, 8 are Very Critical, 21 Critical, 20 Moderately Critical, and 45 Relatively Safe. Manggarai and Tanah Abang are the most critical with impact scores of 9.60% and 8.89%.

Keywords: Critical Node Detection; Graph Embedding; Impact Score; LightGBM; Node2Vec

1. PENDAHULUAN

KRL Commuter Line Jabodetabek merupakan moda transportasi massal utama di kawasan Jakarta, Bogor, Depok, Tangerang, dan Bekasi. Berdasarkan data BPS, jumlah penumpang kumulatif periode Januari 2024 - September 2025 mencapai 583.250.000 penumpang dengan rata-rata bulanan 27,8 juta (BPS, 2025). Pada jaringan dengan tingkat penggunaan setinggi ini, karakteristik interconnected multi-lintasan KRL menyebabkan gangguan pada satu simpul berpotensi memicu efek berantai (*cascading failures*) yang berdampak luas (Zhang et al., 2022; Hu et al., 2024).

Di Indonesia, kajian akademik yang secara spesifik menganalisis kekritisitas simpul pada jaringan KRL Jabodetabek masih sangat terbatas, dan pendekatan operasional saat ini cenderung reaktif. Perkembangan metode analitik berbasis graf membuka peluang analisis yang lebih komprehensif. Chen dan Su (2024) menunjukkan bahwa karakteristik struktural jaringan rel dapat dimodelkan untuk analisis robustness, sementara Huang dan Ma (2024) mengaplikasikan optimisasi berbasis simetri jaringan untuk perencanaan rel urban.

Node2Vec memungkinkan pembangkitan *embedding* yang menangkap pola struktural laten antar simpul (Grover & Leskovec, 2016; Hiraoka et al., 2024), sedangkan LightGBM terbukti efektif untuk data tabular berdimensi kecil-menengah (Ke et al., 2017). Dalam penelitian ini, istilah prediksi mengacu pada kerangka *surrogate modeling* (Müller-Hannemann et al., 2021; Natterer et al., 2025), yaitu model machine learning yang dilatih menggunakan hasil simulasi sebagai *ground truth* sehingga dapat memperkirakan kekritisitas stasiun pada kondisi jaringan yang berubah tanpa simulasi ulang.

Titik kritis didefinisikan sebagai stasiun yang apabila mengalami gangguan akan menyebabkan dampak terbesar terhadap konektivitas dan efisiensi waktu tempuh jaringan, diukur melalui *impact score* hasil simulasi *single-node removal*. Penelitian ini bertujuan: (1) membangun model prediktif berbasis Node2Vec dan LightGBM; (2) mengevaluasi performa model; dan (3) mengidentifikasi fitur jaringan paling berpengaruh terhadap kekritisitas stasiun.

2. LANDASAN TEORI

2.1 Analisis Kekritisitas Simpul Jaringan

Analisis kekritisitas simpul (*critical node detection*) bertujuan mengidentifikasi simpul yang penghapusannya memberikan dampak terbesar terhadap kinerja jaringan. Lu et al. (2024) memodelkan kegagalan berantai pada jaringan rel urban berbasis heterogenitas spasiotemporal penumpang. Ippolito dan Cats (2024) menganalisis robustness jaringan rel multi-modal di Eropa menggunakan pendekatan *multi-layer*. Pal dan Ganguly (2024) mengkaji determinan topologi resiliensi jaringan rel menghadapi *multi-hazard disruption*.

2.2 Node2Vec

Node2Vec adalah metode pembelajaran representasi grafik yang menghasilkan *embedding* berdimensi tetap melalui *random walk* terbiased dengan dua hiperparameter p (return parameter) dan q (in-out parameter) yang mengontrol eksplorasi BFS versus DFS (Grover & Leskovec, 2016). Hiraoka et al. (2024) memperluas Node2Vec dengan *persistent homology* untuk menangkap fitur topologi yang lebih kaya. Kalogeropoulos et al. (2024) menunjukkan keunggulan *embedding* berbasis graf untuk tugas peringkat dan klasifikasi.

2.3 LightGBM

LightGBM adalah implementasi *gradient boosting* berbasis histogram yang efisien untuk dataset berdimensi tinggi (Ke et al., 2017). Algoritma ini dipilih karena kemampuannya menangani dataset kecil-menengah, dukungan native terhadap analisis Gain feature importance, dan kompatibilitas dengan SHAP untuk interpretabilitas model.

2.4 Knowledge Discovery in Database (KDD)

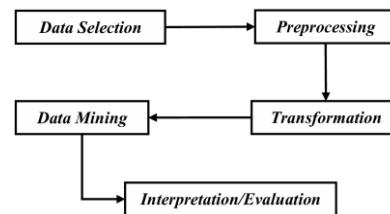
KDD adalah kerangka proses analitik yang terdiri dari Data Selection, Preprocessing, Transformation, Data Mining, dan Interpretation/Evaluation (Fayyad et al., 1996). Kerangka ini digunakan sebagai metodologi utama penelitian untuk menjamin sistematisitas alur dari data mentah hingga pengetahuan yang dapat diinterpretasi.

2.5 Surrogate Modeling

Surrogate modeling adalah pendekatan di mana model *machine learning* dilatih untuk mengaproksimasi output simulasi komputasional mahal (Müller-Hannemann et al., 2021). Natterer et al. (2025) menerapkan *surrogate machine learning* untuk analisis kebijakan transportasi berbasis model agen, menunjukkan efisiensi evaluasi skenario tanpa simulasi penuh.

3. METODOLOGI

Penelitian ini menggunakan kerangka KDD dengan lima tahap sebagaimana ditunjukkan pada Gambar 1. Implementasi menggunakan Python 3.10 pada Google Colaboratory dengan pustaka NetworkX, Node2vec, LightGBM, dan SHAP.



Gambar 1. Diagram alur metodologi penelitian berbasis KDD

3.1 Data Selection

Data jadwal operasional KRL Jabodetabek diperoleh dari API Comuline (api.comuline.com) periode 2025. Dari 113 stasiun nasional, diterapkan dua kriteria inklusi: (1) Daerah Operasi 1 Jakarta, dan (2) tipe layanan KRL Commuterline. Hasilnya: 94 stasiun aktif dengan 30.989 baris data jadwal mentah.

3.2 Preprocessing

Preprocessing mencakup enam langkah: (1) strategi penyelamatan data lintas periode 2024-2025 untuk stasiun belum diperbarui; (2) penghapusan 23 anomali waktu mundur; (3) pemeriksaan missing values; (4) penghapusan duplikasi; (5) penghapusan jadwal bocor melewati stasiun luar Daop 1; dan (6) seleksi akhir stasiun aktif terhubung. Hasil: 18.297 baris data bersih (tingkat retensi 59,0%).

3.3 Transformation

Data jadwal dibangun menjadi *directed weighted graph* $G = (V, E)$ dengan 94 simpul dan 421 sisi, dengan bobot sisi berupa durasi rata-rata perjalanan. Dari graf diekstraksi enam metrik topologi: *in-degree*, *out-degree*, *in-degree weighted*, *out-degree weighted*, *betweenness centrality*, dan *closeness centrality*. Selanjutnya dijalankan simulasi *single-node removal* untuk seluruh 94 stasiun, menghasilkan impact score per stasiun menggunakan Persamaan (1):

$$IS(v) = \alpha \cdot \Delta C(v) + \beta \cdot \Delta T(v) \quad (1)$$

Dengan $\alpha = 0,2$ (bobot komponen konektivitas) dan $\beta = 0,8$ (bobot komponen waktu tempuh), ditentukan melalui analisis sensitivitas. Node2Vec dijalankan dengan konfigurasi optimal (dimensions=32, walk_length=30, num_walks=200, p=1, q=0,5, window=10) menghasilkan 32 dimensi embedding per stasiun.

3.4 Data Mining

Model LightGBM Regressor dilatih menggunakan 38 fitur (6 topologi + 32 embedding) dengan impact score sebagai target regresi. Hyperparameter tuning menggunakan Optuna (100 trial, 5-Fold CV inner loop) menghasilkan konfigurasi optimal: n_estimators=500, learning_rate=0,05, num_leaves=15, max_depth=4, min_child_samples=5, feature_fraction=0,8, bagging_fraction=0,8, bagging_freq=5, reg_alpha=0,1, reg_lambda=1,0.

3.5 Interpretation/Evaluation

Evaluasi menggunakan 5-Fold Cross-Validation dengan metrik MAE, RMSE, dan korelasi Spearman. Bootstrap CI 95% (1.000 iterasi) dihitung untuk Spearman. Interpretasi model menggunakan Gain feature importance dan SHAP (SHapley Additive exPlanations). Ablation study membandingkan tiga konfigurasi fitur: Full Model (38 fitur), Topologi Saja (6 fitur), dan Node2Vec Saja (32 fitur). Analisis articulation point dilakukan sebagai validasi topologi independen.

4. HASIL DAN PEMBAHASAN

4.1 Seleksi Data dan Preprocessing

Dari 113 stasiun nasional, 94 memenuhi kriteria inklusi. Satu stasiun (Cikampek, CKP) dikecualikan meskipun berada di Daop 1 karena bertipe LOCAL (Kereta Api Lokal/KAJJ, bukan KRL Commuterline). Setelah preprocessing enam tahap, diperoleh 18.297 baris data

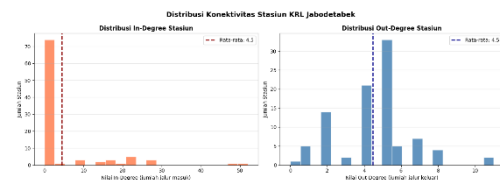
bersih dari 30.989 baris mentah. Tabel 1 menyajikan ringkasan seleksi data.

Tabel 1. Ringkasan Hasil Seleksi dan Preprocessing Data

Tahap	Keterangan	Jumlah
Data Mentah	Total rekaman dari API Comuline	30.989 baris
Setelah Filter Daop 1 + KRL	Stasiun aktif KRL Jabodetabek	94 stasiun
Anomali dihapus	Waktu mundur (tiba lebih awal dari berangkat)	23 rekaman
Jadwal bocor dihapus	Melewati stasiun luar Daop 1	Terfilter
Data Bersih	Hasil akhir preprocessing (retensi 59,0%)	18.297 baris

4.2 Graf Operasional dan Distribusi Fitur

Graf operasional memiliki 94 simpul dan 421 sisi dengan rata-rata frekuensi 43,5 kereta per rute. Distribusi in-degree menunjukkan nilai maksimum 52 pada stasiun terminal utama, sementara out-degree maksimum hanya 11. Asimetri ini mencerminkan perbedaan fungsional antara stasiun terminal (tujuan akhir banyak layanan) dan stasiun distribusi. Distribusi keenam metrik sentralitas bersifat right-skewed sebagaimana ditunjukkan pada Gambar 2, mengindikasikan dominasi segelintir stasiun dalam konektivitas jaringan.



Gambar 2. Distribusi in-degree dan out-degree 94 stasiun KRL Jabodetabek

Simulasi single-node removal menghasilkan distribusi impact score

sangat *right-skewed*: mean 0,87%, median 0,40%, nilai minimum 0% (18 stasiun terminus dengan nilai identik), dan maksimum 9,60% (Manggarai). Tabel 2 menyajikan statistik deskriptif lengkap.

Tabel 2. Statistik deskriptif impact score

Statistik	Nilai	Keterangan
Mean	0,87%	Rata-rata <i>impact score</i>
Median	0,40%	Nilai tengah distribusi
Minimum	0,21%	18 stasiun terminus dengan <i>impact identik</i>
Maksimum	9,6%	Manggarai – hub transit utama
Standar Deviasi	1,48%	Distribusi sangat <i>right-skewed</i>

4.3 Evaluasi Performa Model

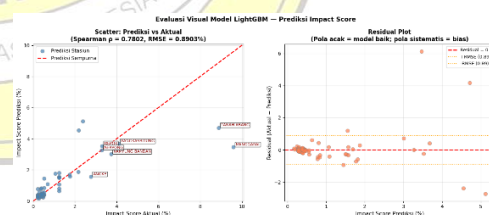
Evaluasi *5-Fold Cross-Validation* menghasilkan MAE 0,3152% ($\pm 0,1022\%$), RMSE 0,8903% ($\pm 0,4027\%$), dan korelasi Spearman 0,7802 ($p < 0,001$) dengan Bootstrap CI 95% [0,6622; 0,8700]. Tabel 3 menyajikan hasil per fold.

Tabel 3. Hasil Evaluasi Model - 5-Fold Cross-Validation

Statistik	Nilai	Interpretasi
MAE (Mean Absolute Error)	0,3152%	Rata-rata kesalahan prediksi $\pm 0,32$ poin persentase per stasiun
RMSE (Root Mean Squared Error)	0,8903%	RMSE > MAE (0,8903 vs 0,3152) ada beberapa prediksi dengan error besar (wajar untuk outlier Manggarai)

Statistik	Nilai	Interpretasi
Spearman Rank Correlation (ρ)	0,7802 ($p < 0,001$)	Kesesuaian urutan prioritas 78,0% - signifikan secara statistik; ini metrik primer penelitian
Fold pelatihan	~75 stasiun / fold	Setiap fold melatih dengan 75 stasiun, menguji 18 19 stasiun
Jumlah fold	5	K=5 dipilih sebagai keseimbangan bias-variansi untuk $n=94$

Scatter plot prediksi vs aktual pada Gambar 3 menunjukkan korelasi positif kuat. Top-10 *recall* mencapai 90%: 9 dari 10 stasiun paling kritis berhasil diidentifikasi. Satu stasiun yang terlewat (Angke, *rank* aktual 7, *impact score* 2,76%) berada di perbatasan kategori Sangat Kritis. Validasi Kendall Tau Sangat Kritis. Validasi Kendall Tau menghasilkan 0,6101 ($p < 0,001$). Lebar *Bootstrap CI* 0,2078 merupakan konsekuensi matematis *outlier* Manggarai yang nilainya $11\times$ di atas rata-rata, setiap *bootstrap iteration* yang tidak menyertakannya menghasilkan *Spearman* lebih rendah.



Gambar 3. Scatter plot prediksi vs actual impact score model LightGBM

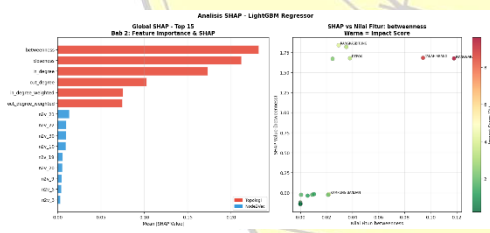
4.4 Analisis Fitur Penting

Berdasarkan konvergensi *Gain Importance* dan SHAP, *betweenness centrality* adalah prediktor paling dominan (*Gain* 3.377,5; Mean |SHAP| 0,2323), diikuti *closeness centrality* (*Gain* 2.531,6; SHAP 0,2122). Fitur topologi

mendominasi 95,9% total Gain dan 89,6% total SHAP. Dominasi ini merupakan konsekuensi logis: definisi *impact score* yang mengukur dampak penghapusan simpul terhadap jalur terpendek secara matematis berkorelasi tinggi dengan *betweenness centrality* yang mengukur proporsi jalur terpendek melewati suatu simpul. Tabel 4 menyajikan top-10 fitur.

Tabel 4. Fitur Paling Berpengaruh – Gain Importance dan SHAP

Rank Gain	Rank SHAP	Fitur	Total Gain	Mean /SHAP/
1	1	Betweenness	3.377,5	0,2323
2	2	Closeness	2.531,6	0,2122
3	4	Out_degree	814,2	0,1027
4	3	In_degree	770,9	0,1733
5	5	In_degree_weighted	550,9	0,0753
6	6	Out_degree_weighted	451,7	0,0746
7-38	7-38	Node2Vec (32 dim. Total)	359,7 total	10,4% total



Gambar 4. SHAP summary plot – kontribusi fitur terhadap prediksi model

Untuk memvalidasi kontribusi gabungan Node2Vec dan LightGBM, dilakukan *ablation study* dengan membandingkan tiga konfigurasi fitur: *Full Model* (38 fitur), *Topologi Saja* (6 fitur), dan *Node2Vec Saja* (32 fitur). Model *Topologi Saja* menghasilkan *Spearman* lebih tinggi (0,8281) dibanding *Full Model* (0,7802), karena penambahan 32 dimensi Node2Vec meningkatkan

rasio fitur/sampel dari 0,064 menjadi 0,404 sehingga menambah *noise* pada dataset berukuran kecil (n=94). Meskipun demikian, Node2Vec tetap digunakan dalam 37% *total split* LightGBM, menunjukkan kontribusinya dalam menangkap pola konektivitas laten yang tidak dapat direpresentasikan oleh metrik topologi skalar. Pada jaringan dengan jumlah simpul lebih besar, kontribusi *embedding* ini berpotensi lebih signifikan.

4.5 Peringkat Kekritisan Stasiun

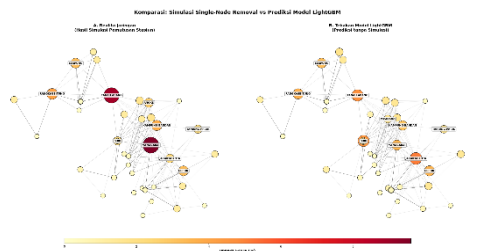
Dari 94 stasiun, 8 dikategorikan Sangat Kritis, 21 Kritis, 20 Cukup Kritis, dan 45 Relatif Aman. Tabel 5 menyajikan 10 stasiun paling kritis berdasarkan prediksi model.

Tabel 5. Top 10 stasiun paling kritis – output prediksi LightGBM

Rank Pred.	Stasiun	Impact Aktual (%)	Impact Prediksi (%)	Kategori
1	JAKARTA KOTA	2,3819	5,1160	Sangat Kritis
2	TANAH ABANG	8,8864	4,7107	Sangat Kritis
3	DURI	2,1685	4,5496	Sangat Kritis
4	RANGKASBITUNG	4,0927	3,6775	Sangat Kritis
5	DEPOK	3,2897	3,5237	Sangat Kritis
6	MANGGARAI	9,5957	3,4658	Sangat Kritis
7	SERPONG	3,2806	3,2680	Sangat Kritis
8	KAMPUNG BANDAN	3,7311	3,0080	Sangat Kritis
9	TANJUNG PRIOK	2,1657	1,8730	Kritis
10	CIKARANG	1,2436	1,8157	Kritis

Manggarai (*impact score* 9,6%) adalah hub transit murni dengan *betweenness centrality* tertinggi (0,1179). Tanah Abang (8,89%) merupakan satu-satunya stasiun yang sekaligus berfungsi

sebagai *articulation point* dan hub transit menjadikannya simpul dengan kerentanan paling komprehensif. Dua articulation point yang teridentifikasi (Tanah Abang dan Rangkasbitung) menunjukkan jenis kekritisan berbeda dari hub: pemutusan konektivitas total ketika terganggu, bukan sekadar degradasi efisiensi. Gambar 5 menampilkan komparasi visual peta kekritisan *ground truth* vs prediksi.



Gambar 5. Komparasi peta kekritisan: *ground truth* simulasi (kiri) vs model prediksi (kanan)

Model *surrogate* yang dihasilkan dapat dijalankan ulang secara periodik setiap terjadi perubahan jadwal operasional KRL, cukup mengekstrak ulang fitur dari graf yang diperbarui tanpa simulasi penuh, memungkinkan evaluasi cepat distribusi kekritisan stasiun pasca-perubahan rute.

5. KESIMPULAN

Penelitian ini berhasil membangun model prediktif berbasis Node2Vec dan LightGBM dalam kerangka KDD untuk memperkirakan tingkat kekritisan stasiun KRL Jabodetabek. Kesimpulan utama:

- (1) Model prediktif berhasil dibangun dari 30.989 baris data mentah melalui *pipeline* KDD komprehensif, menghasilkan 38 fitur gabungan. Evaluasi 5-Fold CV menghasilkan MAE 0,3152%, RMSE 0,8903%, Spearman 0,7802 ($p < 0,001$), dan Top-10 recall 90%.
- (2) *Betweenness centrality* teridentifikasi sebagai fitur paling dominan (95,9%

Gain, 89,6% SHAP), membuktikan secara empiris bahwa kekritisan stasiun sangat dipengaruhi oleh perannya sebagai perantara jalur terpendek dalam jaringan.

- (3) Manggarai dan Tanah Abang adalah stasiun paling kritis (9,60% dan 8,89%). Model *surrogate* yang dihasilkan dapat digunakan untuk evaluasi berkala *robustness* jaringan setiap terjadi perubahan jadwal operasional tanpa simulasi ulang.

6. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ibu Yuyun Umaidah, M.Kom. dan Ibu Betha Nurina Sari, M.Kom. selaku dosen pembimbing atas bimbingan dan arahnya, serta kepada seluruh civitas akademika Program Studi Informatika Universitas Singaperbangsa Karawang yang telah memberikan dukungan selama penelitian ini berlangsung.

DAFTAR PUSTAKA

- Badan Pusat Statistik (BPS). (2025). *PerkembanganTransportasi IndonesiaSeptember 2025*.
- Boldini, D., Grisoni, F., Kuhn, D., Friedrich, L., & Sieber, S. A. (2023). Practical guidelines for the use of gradient boosting for molecular property prediction. *Journal of Cheminformatics*, 15(1). <https://doi.org/10.1186/s13321-023-00743-7>
- BPS. (2024). Statistik Komuter Jabodetabek - Hasil Survei Komuter Jabodetabek 2023. *Statistik Komuter Jabodetabek - Hasil Survei Komuter Jabodetabek 2023*, 3.
- BPTJ. (2025). *Laporan Kinerja Tahun 2024 BPTJ | Badan Pengelola Transportasi Jabodetabek*.
- Chen, B., & Su, M. (2024). *Complex characteristics and robustness analysis of urban rail transit*

- network. 115.
<https://doi.org/10.1117/12.3024089>
- Corrias, R., Gjoreski, M., & Langheinrich, M. (2023). Exploring Transformer and Graph Convolutional Networks for Human Mobility Modeling. *Sensors*, 23(10), 4803. <https://doi.org/10.3390/s23104803>
- Damayanti, D., Megahnia, R., Dinda, P., Jannah, A. R., Makhfudloh, I. I., Agatha, A. B., & Wulandari, Y. N. (n.d.). *Application of Graph Theory in Mapping KRL Commuter Line Services in Jabodetabek*.
- Dong, H., Pan, X., & Chen, X. (2024). Granularity Optimization of Travel Trajectory Based on Node2vec: A Case Study on Urban Travel Time Prediction. *ISPRS International Journal of Geo-Information*, 13(12). <https://doi.org/10.3390/ijgi13120435>
- Giamphy, E., Guillaume, J.-L., Doucet, A., & Sanchis, K. (2023). A survey on bipartite graphs embedding. *Social Network Analysis and Mining*, 13(1), 54. <https://doi.org/10.1007/s13278-023-01058-z>
- Hiraoka, Y., Imoto, Y., Lacombe, T., Meehan, K., & Yachimura, T. (2024). Topological Node2vec: Enhanced Graph Embedding via Persistent Homology. In *Journal of Machine Learning Research* (Vol. 25). <http://jmlr.org/papers/v25/23-1185.html>.
- Hu, J., Yang, M., & Zhen, Y. (2024). A Review of Resilience Assessment and Recovery Strategies of Urban Rail Transit Networks. In *Sustainability (Switzerland)* (Vol. 16, Number 15). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/su16156390>
- Hu, X., Huang, J., & Shi, F. (2022). A robustness assessment with passenger flow data of high-speed rail network in China. *Chaos, Solitons & Fractals*, 165, 112792. <https://doi.org/10.1016/j.chaos.2022.112792>
- Huang, Z., & Ma, C. (2024). Symmetry-Based Urban Rail Transit Network Planning Using Two-Stage Robust Optimization. *Symmetry*, 16(9). <https://doi.org/10.3390/sym16091149>
- Ippolito, N., & Cats, O. (2024a). Multi-modal and multi-layer robustness analysis of the European rail and air networks. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-76264-6>
- Ippolito, N., & Cats, O. (2024b). Multi-modal and multi-layer robustness analysis of the European rail and air networks. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-76264-6>
- Kalogeropoulos, N. R., Ioannou, D., Stathopoulos, D., & Makris, C. (2024). On Embedding Implementations in Text Ranking and Classification Employing Graphs. *Electronics (Switzerland)*, 13(10). <https://doi.org/10.3390/electronics13101897>
- Kalogeropoulos, N.-R., Ioannou, D., Stathopoulos, D., & Makris, C. (2024). On Embedding Implementations in Text Ranking and Classification Employing Graphs. *Electronics*, 13(10), 1897. <https://doi.org/10.3390/electronics13101897>
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (n.d.). *LightGBM: A Highly Efficient Gradient Boosting Decision Tree*. Retrieved <https://github.com/Microsoft/LightGBM>.
- Lu, Q.-C., Li, J., Xu, P.-C., Zhang, L., & Cui, X. (2024a). Modeling cascading failures of urban rail transit network based on passenger spatiotemporal heterogeneity. *Reliability Engineering & System*

- Safety*, 242, 109726.
<https://doi.org/10.1016/j.ress.2023.109726>
- Lu, Q.-C., Li, J., Xu, P.-C., Zhang, L., & Cui, X. (2024b). Modeling cascading failures of urban rail transit network based on passenger spatiotemporal heterogeneity. *Reliability Engineering & System Safety*, 242, 109726.
<https://doi.org/10.1016/j.ress.2023.109726>
- Müller-Hannemann, M., Rückert, R., Schiewe, A., & Schöbel, A. (2021). *Estimating the Robustness of Public Transport Systems Using Machine Learning*.
<http://arxiv.org/abs/2106.08967>
- Natterer, E. S., Rao, S. R., Tejada Lapuerta, A., Engelhardt, R., Hörl, S., & Bogenberger, K. (2025). Machine learning surrogates for agent-based models in transportation policy analysis. *Transportation Research Part C: Emerging Technologies*, 180.
<https://doi.org/10.1016/j.trc.2025.105360>
- Nur, E., & Iman, E. (2024). Comparative Analysis of Classification Algorithms for Citizens Welfare Status using PCA as Feature Selection. *International Journal of Computer Applications*, 186(5), 30–37.
<https://doi.org/10.5120/ijca2024923386>
- Pal, A. K., & Ganguly, A. R. (2024). *Topological Determinants of Resilience in Urban Rail Networks Facing Multi-Hazard Disruptions*.
<http://arxiv.org/abs/2407.06359>
- Pohan, N. A., Perwira, A., Tarigan, M., Perdana, Z., & Nasution, Z. P. (2024). Connectivity Analysis of Non-Container Ports at PT Pelindo Multi Terminal Using SNA and GIS. *Eduvest-Journal of Universal Studies*, 4(9), 8103–8117.
<http://eduvest.greenvest.co.id>
- Sánchez, A., Vidal-Silva, C., Mancilla, G., Tupac-Yupanqui, M., & Rubio, J. M. (2023). Sustainable e-Learning by Data Mining—Successful Results in a Chilean University. *Sustainability (Switzerland)*, 15(2).
<https://doi.org/10.3390/su15020895>
- Sun, G. (2023). Robustness Analysis of an Urban Public Traffic Network Based on a Multi-Subnet Composite Complex Network Model. *Entropy*, 25(10).
<https://doi.org/10.3390/e25101377>
- Transportasi Umum Massal Ramah Lingkungan di Kawasan Jakarta, I., & Bekasi Kementerian Koordinator Bidang Kemaritiman dan Investasi, dan. (2024). *Kajian Sosial Ekonomi*.
- Wang, N., Xu, Y., & Wang, S. (2022). Interpretable boosting tree ensemble method for multisource building fire loss prediction. *Reliability Engineering & System Safety*, 225, 108587.
<https://doi.org/10.1016/j.ress.2022.108587>
- Wei, Y., Yang, X., Xiao, X., Ma, Z., Zhu, T., Dou, F., Wu, J., Chen, A., & Gao, Z. (2024a). Understanding the Resilience of Urban Rail Transit: Concepts, Reviews, and Trends. In *Engineering* (Vol. 41, pp. 7–18). Elsevier Ltd.
<https://doi.org/10.1016/j.eng.2024.01.022>
- Wei, Y., Yang, X., Xiao, X., Ma, Z., Zhu, T., Dou, F., Wu, J., Chen, A., & Gao, Z. (2024b). Understanding the Resilience of Urban Rail Transit: Concepts, Reviews, and Trends. In *Engineering* (Vol. 41, pp. 7–18). Elsevier Ltd.
<https://doi.org/10.1016/j.eng.2024.01.022>
- Yang, L., Chatelain, C., & Adam, S. (2024). Dynamic Graph Representation Learning With Neural Networks: A Survey. *IEEE Access*, 12, 43460–43484.

<https://doi.org/10.1109/ACCESS.2024.3378111>

