

Graph Convolutional Network (GCN) untuk Klasifikasi Konektivitas Bandara

Rendy Saputro, Dian Puspita Hapsari

Teknik Informatika, Fakultas Teknik Elektro dan Teknologi Informasi,

Institut Teknologi Adhi Tama Surabaya

Email: rendyys418@gmail.com, dian.puspita@itats.ac.id

Diajukan: 12-08-2026

Direvisi: 09-09-2026

Disetujui: 14-09-2026

Diterbitkan: 30-09-2026

Abstract. The global flight network naturally forms a complex transportation system with a spatial graph structure. Understanding airport connectivity hierarchies proves crucial for route analysis and infrastructure optimisation. However, conventional tabular machine learning approaches face unique challenges when modelling graph-structured relational data. This study proposes a graph-based deep learning architecture, specifically the Graph Convolutional Network (GCN), to classify global airport connectivity levels. The dataset comprises 2024 global flight operation records, representing 7,698 nodes (airports) and 36,589 unique edges (flight routes) formed from 66,316 raw route records (one row per airline per route) after the data-cleaning stage. Since the dataset lacks explicit class annotations, the study applies the unsupervised Fuzzy C-Means (FCM) algorithm, using degree centrality metrics to generate pseudo-ground truth. This method classifies the airports into three categories: Low, Medium, and High. The GCN architecture incorporates two graph convolution layers with ReLU activation, a Dropout regularisation technique set at a 0.5 ratio, and a Softmax classification output function. A transductive node masking scheme splits the data into an 70% training, 15% validation, and 15% testing ratio, preserving the integrity of the network topology. Testing reveals that the GCN model, optimised with the Adam algorithm over 900 epochs, predicts connectivity classes with an overall accuracy of 95% and a macro F1-score of 79%. Overall, this study demonstrates that the GCN algorithm effectively and robustly captures and aggregates spatial attribute features alongside connection topology, enabling accurate node classification for large-scale, imbalanced air transportation systems.

Keywords: Airport Connectivity, Flight Network, Fuzzy C-Means, Graph Convolutional Network, Node Classification

Abstrak. Jaringan penerbangan global merupakan sistem transportasi kompleks yang secara alamiah membentuk struktur graf keruangan (spasial). Pemahaman terhadap hierarki konektivitas bandara sangat krusial bagi analisis rute dan optimalisasi infrastruktur. Namun, pemodelan data relasional berstruktur graf memiliki tantangan tersendiri jika diselesaikan menggunakan pendekatan machine learning konvensional berskala tabular. Penelitian ini mengusulkan penerapan arsitektur Deep Learning berbasis graf, yaitu Graph Convolutional Network (GCN), untuk mengklasifikasikan tingkat konektivitas bandara di seluruh dunia. Dataset yang digunakan merupakan rekaman operasional penerbangan global tahun 2024 yang merepresentasikan 7.698 simpul (bandara) dan 36.589 sisi (edge) unik antarbandara, yang dibentuk dari 66.316 baris data rute mentah (satu baris per maskapai per rute) setelah melalui tahap pembersihan data.. Karena dataset tidak memiliki anotasi kelas yang eksplisit, algoritma Fuzzy C-Means (FCM) diterapkan secara unsupervised berdasarkan metrik degree centrality untuk menghasilkan pseudo-ground truth. Melalui metode ini, bandara diklasifikasikan ke dalam tiga kategori: Rendah, Sedang, dan Tinggi. Arsitektur GCN dirancang menggunakan dua lapisan konvolusi graf dengan fungsi aktivasi ReLU, teknik regularisasi Dropout dengan rasio 0.5, serta fungsi keluaran klasifikasi Softmax. Pembagian data dilakukan menggunakan skema transductive node masking dengan rasio 70% data latih, 15% data validasi, dan 15% data uji guna menjaga keutuhan topologi jaringan. Hasil pengujian menunjukkan bahwa model GCN yang dioptimasi menggunakan algoritma Adam pada 900 Epoch mampu memprediksi kelas konektivitas dengan tingkat akurasi keseluruhan sebesar 95% dan Macro F1-Score sebesar 79%. Secara keseluruhan, penelitian ini membuktikan bahwa algoritma GCN efektif dan tangguh dalam menangkap dan mengagregasi fitur atribut spasial secara simultan dengan topologi

koneksi untuk pemodelan node classification pada sistem transportasi udara skala besar yang tidak seimbang (imbalanced data).

Kata Kunci: Fuzzy C-Means, Graph Convolutional Network, Jaringan Penerbangan, Klasifikasi Node, Konektivitas Bandara

1. Pendahuluan

Konektivitas bandara memainkan peran krusial dalam menentukan posisi dan tingkat kepentingan sebuah lokasi di dalam jaringan penerbangan global (Zanin & Lillo, 2013). Analisis konektivitas tidak hanya sebatas menghitung *volume* kedatangan dan keberangkatan penerbangan, melainkan tentang bagaimana suatu bandara terhubung secara struktural dengan bandara lainnya di seluruh dunia (Guimerà et al., 2005). Struktur relasional ini secara matematis dapat direpresentasikan sebagai sebuah graf, di mana lokasi bandara bertindak sebagai simpul (*node*) dan rute penerbangan bertindak sebagai sisi (*edge*). Representasi graf terbukti mampu menyederhanakan kompleksitas data tabular menjadi bentuk visual dan topologis yang lebih intuitif untuk mengkaji hubungan antarbandara (Sun & Wandelt, 2021).

Perkembangan teknologi Pembelajaran Mesin (*Machine Learning*) saat ini telah memberikan solusi komputasional yang signifikan dalam memproses data jaringan yang kompleks. Jika sebelumnya analisis konektivitas membutuhkan kalkulasi analitik tradisional yang lama, pendekatan kecerdasan buatan memungkinkan sistem untuk mempelajari dan mengekstraksi pola secara komprehensif. Namun, tantangan mendasar yang selama ini terjadi adalah metode evaluasi konektivitas konvensional yang cenderung statis dan terisolasi, di mana total penerbangan atau kesibukan bandara diukur secara independen (Jiang & Luo, 2022). Kelemahan utama dari metode konvensional adalah pengabaian terhadap pengaruh struktural dari bandara tetangga yang terhubung dengannya. Akibatnya, hasil analisis kerap kali bias dan gagal menggambarkan posisi strategis bandara dalam topologi jaringan global, khususnya bagi bandara kecil atau perintis yang sering kali dianulir kepentingannya akibat frekuensi penerbangannya yang minim.

Algoritma Graph Neural Network (GNN) dan variannya, Graph Convolutional Network (GCN), hadir sebagai solusi inovatif untuk mengatasi kelemahan analitik struktural tersebut (Kipf & Welling, 2017; Wu, Pan, Chen, Long, Zhang, & Yu, 2019). Meskipun riset mutakhir telah mengadopsi teori jaringan kompleks untuk menganalisis ketahanan rute penerbangan global, implementasi GNN pada sektor transportasi udara saat ini masih didominasi oleh pemodelan prediktif berbasis deret waktu (*time-series forecasting*), seperti prediksi *volume* lalu lintas atau keterlambatan penerbangan (Jiang & Luo, 2022; Huang, Yang, & Yan, 2025). Hal ini memperlihatkan adanya celah penelitian (*research gap*) yang nyata, di mana pemanfaatan arsitektur GCN untuk klasifikasi tingkat konektivitas bandara (*node classification*) masih sangat jarang dieksplorasi. Padahal, algoritma ini memiliki kapabilitas khusus untuk memproses pola relasional (matriks ketetanggaan) dan atribut spasial simpul secara simultan. Studi banding pada domain transportasi lain turut menekankan bahwa klaim keunggulan GCN idealnya disertai baseline pembandingan metode konvensional seperti Random Forest, karena performanya kerap kompetitif dan bahkan setara pada sejumlah kasus (Ting, Li, Sanner, & Abdulhai, 2023).

Penelitian ini bertujuan untuk mengimplementasikan model GCN guna mengklasifikasikan bandara ke dalam kategori tingkat konektivitas tinggi, menengah, dan rendah berdasarkan struktur relasional jaringan penerbangan global, serta mengevaluasi performanya menggunakan metrik pengujian yang komprehensif seperti Accuracy, Precision, Recall, dan F1-Score. Untuk mengatasi ketiadaan label kelas acuan (*ground truth*) pada dataset publik Kaggle (Rafiee, 2024), penelitian ini mengintegrasikan algoritma Fuzzy C-Means (FCM) berdasarkan nilai Degree Centrality sebagai metode pelabelan otomatis. Penggunaan prinsip Max-Membership pada FCM memungkinkan penetapan label tegas yang objektif sebelum dilatih menggunakan model GCN.

2. Tinjauan Pustaka

Kecerdasan Buatan (Artificial Intelligence) telah berkembang pesat dengan melahirkan paradigma Deep Learning yang mampu mengekstraksi hierarki fitur kompleks secara otomatis dari data mentah (Goodfellow, Bengio, & Courville, 2016). Dalam domain data non-Euclidean yang memiliki struktur tidak beraturan seperti graf jaringan transportasi, arsitektur konvensional seperti Convolutional Neural Network (CNN) maupun Recurrent Neural Network (RNN) tidak dapat diterapkan secara langsung karena ketidakaturan jumlah tetangga pada setiap simpul (Zhou, Cui, Hu, Zhang, Yang, Liu, Wang, Li, & Sun, 2020). GNN mengatasi keterbatasan ini melalui mekanisme perambatan pesan (neural message passing), di mana representasi fitur suatu simpul diperbarui secara iteratif dengan mengagregasi informasi dari simpul-simpul tetangga yang terhubung langsung dengannya (Wu, Pan, Chen, Long, Zhang, & Yu, 2019).

Graph Convolutional Network (GCN) yang diperkenalkan oleh (Kipf & Welling, 2017) menyederhanakan konvolusi graf spektral menggunakan aproksimasi polinomial Chebyshev orde pertama. Secara matematis, aturan propagasi per lapisan pada model GCN dirumuskan sebagai berikut:

$$H^{l+1} = \sigma(\hat{D}^{(-\frac{1}{2})} \hat{A} \hat{D}^{(-\frac{1}{2})} H^l W^l) \quad (1)$$

dimana $\hat{A} = A + I_N$ merupakan matriks ketetanggaan yang ditambah dengan matriks identitas (*self-loops*) agar fitur lokal simpul tetap dipertahankan, $\hat{D}$ adalah matriks derajat diagonal di mana $\hat{D}_{ii} = \sum_j \hat{A}_{ij}$; $H^{(l)}$ adalah matriks representasi fitur pada lapisan ke- l , $W^{(l)}$ adalah matriks bobot parameter yang dipelajari, dan σ adalah fungsi aktivasi non-linier seperti Rectified Linear Unit (ReLU) (Nair & Hinton, 2010). Operasi normalisasi simetris $\hat{D}^{-1/2} \hat{A} \hat{D}^{-1/2}$ yang dikenal sebagai renormalization trick sangat krusial untuk mencegah ketidakstabilan numerik seperti ledakan gradien (exploding gradients) pada simpul hub berderajat tinggi dan hilangnya sinyal (vanishing gradients) pada simpul kecil.

Dalam analisis jaringan transportasi udara, ukuran sentralitas dasar yang digunakan untuk mengukur tingkat keterhubungan suatu bandara adalah *Degree Centrality* (Zanin & Lillo, 2013). Karena rentang nilai derajat konektivitas sangat lebar, algoritma pengelompokan berbasis logika fuzzy yaitu Fuzzy C-Means (FCM) diterapkan untuk mengelompokkan simpul berdasarkan atribut topologisnya (Bezdek, 1981; Kadhim Abdulzahra & Al-Qurabat, 2022). Fungsi objektif FCM yang diminimalkan adalah jumlah kuadrat galat terbobot:

$$J_m(U, V) = \sum_{i=1}^N \sum_{j=1}^C (u_{ij})^m \|x_i - v_j\|^2 \quad (2)$$

dimana u_{ij} adalah derajat keanggotaan data x_i pada kluster v_j , dan $m > 1$ adalah parameter pembobot fuzzy. Setelah konvergensi tercapai, penetapan label kelas tegas (hard label) dilakukan menggunakan prinsip Max-Membership yang memilih indeks kluster dengan nilai derajat keanggotaan tertinggi. Beberapa penelitian terdahulu membuktikan efektivitas GCN pada tugas klasifikasi simpul dan optimasi jaringan, seperti (Chen, Huang, Xu, Deng, Huang, He, & Li, 2022) yang meningkatkan ketangguhan GCN melalui kernel fitur tetangga, serta (Sejan, Rahman, Aziz, Baik, You, & Song, 2023) yang meningkatkan akurasi klasifikasi simpul pada graf berskala besar.

3. Metode Penelitian

Penelitian ini dilaksanakan melalui alur kerja eksperimental yang sistematis guna memastikan model GCN dapat mempelajari struktur topologi jaringan penerbangan secara optimal. Tahapan metodologi meliputi: (1) Pengumpulan dan prapemrosesan data. (a) Seleksi fitur dan pembersihan nilai kosong. (b) Normalisasi koordinat spasial. (2) Konstruksi graf berbobot dan pelabelan otomatis. (a) Perhitungan matriks ketetanggaan berbobot. (b) Klasterisasi FCM. (3) Perancangan dan pelatihan model GCN. (a) Strategi pembagian data. (b) Konfigurasi lapisan dan optimasi.



Gambar 1. Bagan Alur Penelitian

Dataset yang digunakan bersumber dari repositori publik Kaggle berjudul "Airports, Airlines, Planes, and Routes Update 2024" (Rafiee, 2024). Dua berkas utama yang diekstraksi adalah airports.csv (7.698 data bandara) dan routes.csv (67.663 data rute). Proses pembersihan data meliputi penghapusan baris dengan nilai kosong, eliminasi duplikasi rute, serta validasi relasi silang (cross-relational check) untuk menghapus rute yatim (orphan routes) yang terhubung ke bandara yang tidak terdaftar. Setelah prapemrosesan, diperoleh 7.698 bandara valid dan 66.316 baris rute mentah. Atribut koordinat geografis yaitu Latitude dan Longitude dinormalisasi ke dalam rentang [0, 1] menggunakan Min-Max Scaler agar tidak terjadi dominasi skala nilai selama pembelajaran gradien.

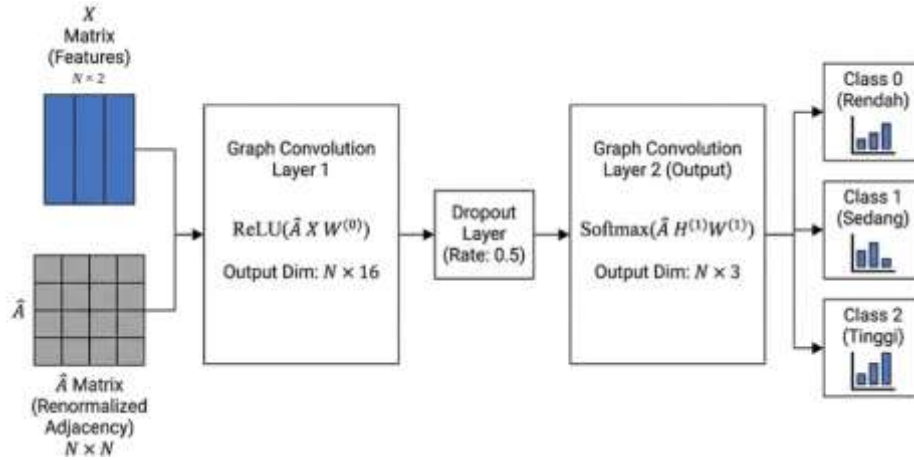
Konstruksi graf $G = (V, E)$ memetakan 7.698 bandara sebagai himpunan simpul V dengan matriks fitur X berdimensi 7698×2 (koordinat lintang dan bujur ternormalisasi). Himpunan sisi E dibentuk dari pasangan bandara asal dan tujuan pada data rute. Berbeda dengan graf biner, penelitian ini membangun matriks ketetanggaan berbobot A di mana elemen A_{ij} diisi dengan jumlah maskapai unik yang melayani rute langsung dari bandara i ke bandara j , menghasilkan total 36.589 sisi berbobot unik. Matriks A kemudian dinormalisasi secara spektral simetris menjadi $\hat{A}$ sesuai Persamaan (1).

Pembatasan fitur masukan hanya pada dua atribut spasial (Latitude dan Longitude) merupakan keputusan desain yang disengaja, bukan akibat keterbatasan dataset, sebab atribut lain seperti Altitude dan Type sebenarnya tersedia pada dataset Kaggle namun sengaja tidak diikutsertakan. Tujuannya agar performa klasifikasi yang dicapai murni mencerminkan kapabilitas GCN dalam mempelajari topologi jaringan (struktur ketetanggaan) melalui mekanisme message passing, bukan kontribusi fitur atribut tambahan. Desain ini berfungsi sebagai bentuk pengujian terkontrol terhadap kapabilitas struktural murni arsitektur GCN.

Pelabelan otomatis dilakukan dengan menghitung nilai Degree Centrality setiap bandara pada graf, yang kemudian dinormalisasi [0, 1] dan diinputkan ke algoritma FCM dengan parameter jumlah klaster $C = 3$ (Rendah, Sedang, Tinggi), eksponen fuzzy $m = 2.0$, dan ambang toleransi galat $E = 0,005$. Hasil keanggotaan fuzzy didefuzzifikasi dengan fungsi argmax (Max-Membership) untuk menghasilkan vektor target Y berindeks diskrit $\{0, 1, 2\}$.

Arsitektur model GCN dirancang dengan kedalaman dua lapisan konvolusi graf (2-layer GCN) untuk menghindari fenomena over-smoothing pada jaringan dunia-kecil (small-world network). Lapisan pertama mentransformasi fitur input berdimensi $N \times 2$ menjadi ruang representasi tersembunyi berdimensi $N \times 16$ menggunakan fungsi aktivasi ReLU. Regularisasi Dropout dengan probabilitas $p = 0.5$ diterapkan setelah lapisan pertama untuk mencegah overfitting. Lapisan kedua memetakan fitur tersembunyi menjadi keluaran logit berdimensi $N \times 3$ yang kemudian diubah

menjadi distribusi probabilitas menggunakan fungsi Softmax. Spesifikasi lengkap parameter model dirangkum pada Tabel 1.



Gambar 2. Arsitektur Model Graph Convolutional Network (GCN)

Tabel 1. Spesifikasi Arsitektur dan Konfigurasi Model GCN

Komponen / Parameter	Spesifikasi / Nilai	Keterangan Fungsional
Dimensi Fitur Masukan (Input)	$N \times 2$	Koordinat Latitude dan Longitude ternormalisasi
Lapisan Tersembunyi (Hidden Units)	16 Unit	Dimensi fitur laten pada lapisan GCN pertama
Dimensi Keluaran (Output Layer)	$N \times 3$	Probabilitas kelas: Rendah (0), Sedang (1), Tinggi (2)
Fungsi Aktivasi Tersembunyi	ReLU	Non-linieritas dan mitigasi vanishing gradient
Fungsi Aktivasi Keluaran	Softmax	Distribusi probabilitas klasifikasi multikelas
Rasio Regularisasi (Dropout)	0.5 (50%)	Mencegah overfitting pada lapisan tersembunyi
Algoritma Optimasi (Optimizer)	Adam	Laju pembelajaran adaptif (learning rate = 0.01)
Weight Decay (L2 Regularization)	5×10^{-4}	Penalti regularisasi bobot pada Adam
Fungsi Kerugian (Loss Function)	Cross-Entropy	Dilengkapi class weighting untuk imbalanced data

Karena pemodelan graf membutuhkan konektivitas antar-simpul yang utuh, pembagian dataset diterapkan menggunakan teknik transductive node masking. Seluruh 7.698 simpul tetap berada dalam matriks ketetanggaan $\hat{A}$, namun evaluasi kerugian hanya diaktifkan pada penanda biner tertentu. Dataset dibagi dengan rasio 70% data latih (5.388 simpul), 15% data validasi (1.155 simpul), dan 15% data uji (1.155 simpul). Pelatihan dijalankan menggunakan metode Full-Batch Gradient Descent dengan optimalisasi Adam (Kingma & Ba, 2017). Untuk menangani ketimpangan distribusi kelas, fungsi kerugian Categorical Cross-Entropy diberikan bobot kelas (class weighting) berbanding terbalik dengan frekuensi sampel. Kinerja model diukur menggunakan metrik Confusion Matrix, Accuracy, Precision, Recall, dan Macro F1-Score.

Selain keluaran klasifikasi diskrit, penelitian ini juga menghitung skor konektivitas kontinu untuk setiap simpul berdasarkan nilai ekspektasi distribusi probabilitas Softmax, guna memungkinkan pemeringkatan yang lebih presisi antarbandara dalam kelas prediksi yang sama:

$$Skor_i = \sum_{k=0}^{C-1} k \cdot (P(k))_i \quad (3)$$

di mana $Skor_i$ adalah skor konektivitas kontinu untuk simpul i , k adalah indeks kelas (0 = Rendah, 1 = Sedang, 2 = Tinggi), C adalah jumlah total kelas, dan $(P(k))_i$ adalah probabilitas simpul i terhadap kelas k hasil fungsi Softmax.

4. Hasil dan Pembahasan

Eksplorasi awal terhadap 7.698 simpul bandara dan 36.589 sisi berbobot memperlihatkan bahwa jaringan penerbangan global memiliki karakteristik kuat dari jaringan skala-bebas (scale-free network) dengan distribusi derajat berekor tebal (heavy-tailed distribution). Sebagian besar bandara hanya melayani sedikit rute penerbangan, sementara sebagian kecil bandara mendominasi konektivitas sebagai pusat hub antarbentang benua. Sepuluh bandara dengan derajat sentralitas tertinggi didominasi oleh fasilitas internasional raksasa seperti Frankfurt am Main Airport (FRA, 477 rute), Charles de Gaulle Airport (CDG, 470 rute), Amsterdam Airport Schiphol (AMS, 463 rute), dan Hartsfield-Jackson Atlanta International Airport (ATL, 433 rute).

Proses klusterisasi otomatis menggunakan algoritma FCM berhasil mencapai konvergensi pada iterasi ke-42 dengan nilai Fuzzy Partition Coefficient (FPC) sebesar 0,9691. Nilai FPC yang mendekati 1,0 ini membuktikan bahwa batas antarkelompok topologi telah dipisahkan secara tajam dan konsisten. Berdasarkan penetapan prinsip Max-Membership, distribusi keanggotaan kelas acuan semu (pseudo-ground truth) disajikan pada Tabel 2.

Tabel 2. Distribusi Populasi Bandara Berdasarkan Kelas Konektivitas FCM

Indeks Kelas	Kategori / Interpretasi Topologi	Total Bandara	Persentase (%)	Pusat Centroid
0	Konektivitas Rendah (Bandara Perintis / Lokal)	7.300	94,83%	0,0058
1	Konektivitas Sedang (Bandara Regional)	321	4,17%	0,1788
2	Konektivitas Tinggi (Global Hubs)	77	1,00%	0,5720
Total	Total Keseluruhan Populasi Simpul	7.698	100,00%	-

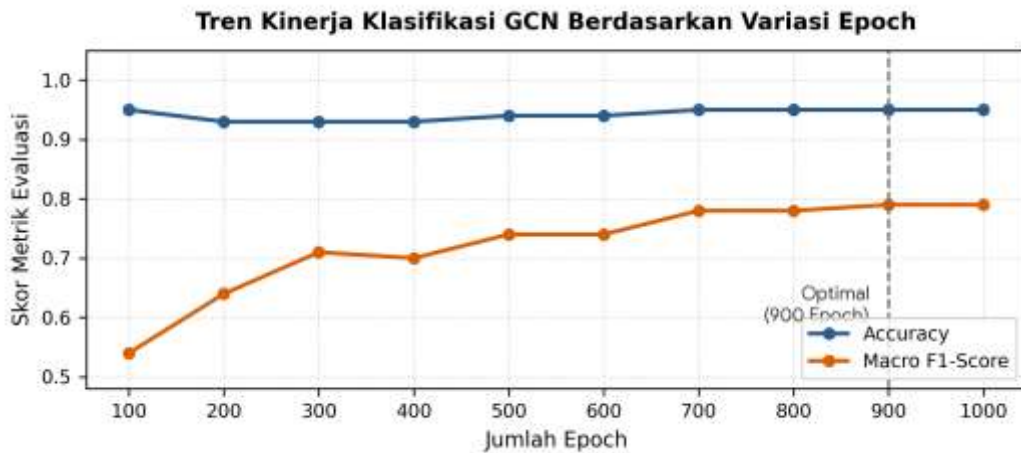
Tabel 2 mengonfirmasi bahwa dataset jaringan penerbangan mengalami ketimpangan kelas yang sangat ekstrem (highly imbalanced dataset). Bandara berstatus Global Hubs (Kelas 2) hanya mencakup 1,00% dari total populasi dunia (77 bandara), sementara bandara perintis dan lokal (Kelas 0) mendominasi sebanyak 94,83% (7.300 bandara). Fenomena ini menjadi alasan utama mengapa pemantauan model GCN diprioritaskan pada metrik Macro F1-Score dan Recall pada kelas minoritas agar model tidak terjebak pada bias akurasi mayoritas.

Guna menentukan konfigurasi iterasi pembelajaran yang paling optimal, eksperimen pengujian dilakukan terhadap 10 skenario variasi epoch mulai dari 100 hingga 1.000 Epoch dengan kelipatan 100. Evaluasi performa pada data uji (1.155 simpul asing) diuji menggunakan urutan prioritas: Macro F1-Score sebagai metrik utama, nilai AUC kelas Sedang sebagai penentu pemecah keseimbangan (tie-breaker), dan Accuracy keseluruhan sebagai metrik pelengkap. Hasil komparasi lengkap disajikan pada Tabel 3 dan Gambar 3.

Tabel 3. Perbandingan Kinerja Klasifikasi GCN Berdasarkan Variasi Epoch

Skenario Epoch	Accuracy	Macro F1-Score	AUC Sedang	Indikasi dan Karakteristik Model
100 Epoch	0,95	0,54	0,68	Underfitting; gagal mengenali kelas Sedang (F1 = 0,00)
200 Epoch	0,93	0,64	0,89	Konvergensi awal; mulai memetakan kelas minoritas
300 Epoch	0,93	0,71	0,93	Peningkatan performa klasifikasi kelas regional
500 Epoch	0,94	0,74	0,95	Performa meningkat secara konsisten dan stabil
700 Epoch	0,95	0,78	0,96	Lonjakan akurasi signifikan pada kelas minoritas

900 Epoch	0,95	0,79	0,97	Model Optimal (Sweet Spot); F1 kelas Tinggi 0,81
1.000 Epoch	0,95	0,79	0,97	Titik jenuh (plateau); tidak ada peningkatan lanjut



Gambar 3. Tren pergerakan kinerja klasifikasi model GCN berdasarkan variasi epoch pelatihan

Berdasarkan Gambar 3 dan Tabel 3, performa klasifikasi model GCN memperlihatkan peningkatan yang konsisten dari epoch ke epoch. Pada 100 Epoch, model masih mengalami underfitting di mana kelas Sedang gagal terdeteksi sama sekali ($F1 = 0,00$). Konvergensi yang stabil mulai tercapai pada 500 hingga 700 Epoch, dan mencapai titik puncaknya pada skenario 900 Epoch dengan Macro F1-Score sebesar 0,79 dan AUC kelas Sedang sebesar 0,97. Skenario 900 Epoch dipilih sebagai model final karena memberikan F1-Score tertinggi pada kelas Tinggi (0,81 berbanding 0,79 pada 1.000 Epoch). Kesetaraan performa antara 900 dan 1.000 Epoch menandakan model telah mencapai titik jenuh (plateau), sehingga 900 Epoch ditetapkan sebagai konfigurasi komputasi yang paling optimal dan efisien.

Evaluasi mendalam terhadap prediksi model terbaik (900 Epoch) pada 1.155 sampel uji dilakukan dengan menganalisis rincian performa per kelas dan visualisasi matriks kebingungan (Confusion Matrix), sebagaimana disajikan pada Tabel 4 dan Tabel 5.

Tabel 4. Laporan Kinerja Klasifikasi GCN pada Data Uji (Test Set, N = 1.155)

Kelas Konektivitas	Precision	Recall	F1-Score	Jumlah Sampel (Support)
Konektivitas Rendah (0)	1,00	0,96	0,98	1.091
Konektivitas Sedang (1)	0,46	0,80	0,59	51
Konektivitas Tinggi (2)	0,68	1,00	0,81	13
Rata-rata Makro (Macro Average)	0,71	0,92	0,79	1.155
Rata-rata Tertimbang (Weighted)	0,97	0,95	0,96	1.155
Akurasi Keseluruhan (Accuracy)	0,95 (95,00%)	-	-	1.155

Tabel 5. Matriks Kebingungan (Confusion Matrix) Model GCN pada Data Uji

Ground Truth (Aktual FCM)	Rendah (0)	Sedang (1)	Tinggi (2)	Total Aktual
Rendah (0)	1.043	48	0	1.091
Sedang (1)	4	41	6	51
Tinggi (2)	0	0	13	13
Total Prediksi Model	1.047	89	19	1.155

Analisis terhadap Tabel 4 dan Tabel 5 memperlihatkan bahwa model GCN memiliki ketepatan prediksi yang sangat tinggi. Pada kelas Rendah, model mengenali 1.043 dari 1.091 bandara dengan tepat (Recall 0,96 dan Precision 1,00), dan tidak ada satupun bandara lokal yang keliru diklasifikasikan sebagai bandara hub global (0 salah tebak ke kelas Tinggi). Pada kelas Sedang (bandara regional), model berhasil mengidentifikasi 41 dari 51 bandara dengan Recall 0,80. Sedikit pergeseran klasifikasi (4 sampel salah ke Rendah dan 6 ke Tinggi) mencerminkan karakteristik topologi bandara regional yang memiliki rentang rute transisi dan tumpang tindih dengan hub nasional. Precision rendah pada kelas Sedang (0,46) turut dipengaruhi oleh kesalahan pada arah sebaliknya: sebanyak 48 bandara kelas Rendah pada Tabel 5 justru diprediksi sebagai Sedang, jauh melebihi 41 prediksi benar kelas Sedang itu sendiri. Sebagian bandara Rendah dengan derajat konektivitas mendekati batas atas distribusinya berpotensi terdorong melewati ambang keputusan model menuju kelas Sedang, sehingga precision kelas ini tertekan meskipun recall-nya tetap tinggi (0,80).

Temuan paling signifikan adalah presisi sempurna pada kelas Tinggi (Global Hubs), di mana model GCN berhasil mengklasifikasikan seluruh 13 sampel bandara kelas Tinggi dengan tepat 100% (Recall = 1,00; 0 sampel keliru ke kelas lain). Hal ini membuktikan bahwa mekanisme perambatan pesan pada GCN mampu mengekstraksi informasi struktur topologi graf yang membedakan Global Hubs di seluruh benua. Pada analisis kurva Receiver Operating Characteristic (ROC) dengan skema One-vs-Rest, model memperoleh nilai Area Under Curve (AUC) sebesar 0,99 untuk kelas Rendah, 0,97 untuk kelas Sedang, dan 1,00 sempurna untuk kelas Tinggi.

Selain pemetaan diskrit tiga kelas, model GCN menghasilkan pemeringkatan (ranking) konektivitas yang kontinu berdasarkan skor ekspektasi keluaran Softmax sesuai definisi Persamaan (3) berentang [0, 2]. Sepuluh bandara dengan skor konektivitas tertinggi disajikan pada Tabel 6.

Tabel 6. Peringkat Sepuluh Bandara dengan Skor Konektivitas GCN Tertinggi

Peringkat	Nama Bandara	Kode IATA	Negara	Jumlah Rute	Skor GCN
1	Hartsfield-Jackson Atlanta International Airport	ATL	United States	433	1,9990
2	Chicago O'Hare International Airport	ORD	United States	409	1,9940
3	Dallas/Fort Worth International Airport	DFW	United States	372	1,9908
4	London Heathrow Airport	LHR	United Kingdom	340	1,9884
5	Los Angeles International Airport	LAX	United States	295	1,9827
6	Domodedovo International Airport	DME	Russia	375	1,9802
7	Charles de Gaulle International Airport	CDG	France	470	1,9791
8	Denver International Airport	DEN	United States	335	1,9787
9	Frankfurt am Main Airport	FRA	Germany	477	1,9778
10	Beijing Capital International Airport	PEK	China	408	1,9700

Tabel 7. Perbandingan GCN dengan Baseline Random Forest pada Data Uji yang Identik

Model	Fitur Masukan	Accuracy	Macro F1-Score
GCN (hasil penelitian ini)	Latitude, Longitude, dan topologi graf	0,95	0,79
Random Forest	Latitude dan Longitude (identik dengan GCN)	0,94	0,32
Random Forest	Degree Centrality (basis pembentukan label FCM)	1,00	1,00

Berdasarkan Tabel 6, seluruh sepuluh bandara teratas merupakan hub penerbangan dunia yang memiliki rute penerbangan masif. Sebaliknya, bandara dengan skor konektivitas terendah didominasi oleh landasan perintis di wilayah terisolasi Arktik (seperti Mys Shmidt Airport Rusia dengan skor 0,0063 dan Cape Lisburne LRRS Airport Alaska dengan skor 0,0072). Kemampuan model memberikan skor yang sangat konsisten antara ujung hierarki tertinggi dan terendah membuktikan bahwa GCN berhasil memetakan topologi jaringan penerbangan global. Seluruh alur papemrosesan, klasterisasi FCM, pelatihan GCN, serta pemetaan graf interaktif telah diintegrasikan

ke dalam antarmuka dasbor berbasis web interaktif. Untuk memperkuat klaim keunggulan GCN atas pendekatan konvensional, penelitian ini menambahkan pembanding empiris menggunakan algoritma Random Forest yang diuji pada data uji (test set) identik dengan yang digunakan GCN (1.155 sampel), sebagaimana disajikan pada Tabel 7.

Pada fitur masukan yang identik dengan GCN (Latitude dan Longitude), Random Forest memperoleh Accuracy tinggi (0,94) namun Macro F1-Score sangat rendah (0,32) karena model gagal memprediksi satupun sampel pada kelas Sedang dan Tinggi (recall = 0), sehingga nilai Accuracy yang tinggi semata-mata mencerminkan dominasi kelas Rendah pada dataset yang tidak seimbang. Perbandingan ini menegaskan bahwa keunggulan GCN dalam menangani kelas minoritas (Macro F1-Score 0,79) berasal dari kemampuannya mengagregasi informasi struktural melalui message passing, bukan sekadar dari fitur spasial yang digunakan. Sebaliknya, ketika Random Forest diberi fitur Degree Centrality, yaitu atribut yang menjadi basis pembentukan label FCM, model mencapai skor sempurna (Accuracy dan Macro F1-Score = 1,00). Temuan ini secara eksplisit mengonfirmasi kekhawatiran keterbatasan metodologis yang dibahas berikut ini.

Perlu dicermati bahwa label kelas konektivitas pada penelitian ini dihasilkan melalui klusterisasi FCM berdasarkan nilai Degree Centrality, yang dihitung langsung dari matriks ketetanggaan A , yaitu matriks yang sama yang digunakan GCN sebagai basis operasi message passing. Hubungan struktural yang erat ini menyebabkan tugas klasifikasi cenderung mendekati pemetaan deterministik dari topologi graf ke label, sehingga tingginya nilai akurasi dan AUC yang diperoleh tidak sepenuhnya mengejutkan dan perlu dimaknai secara proporsional. Interpretasi bahwa GCN mampu menangkap struktur topologi secara robust pada penelitian ini lebih tepat dipahami sebagai kemampuan model merekonstruksi kembali fungsi struktural yang menjadi basis pelabelan itu sendiri, ketimbang generalisasi terhadap definisi konektivitas yang independen dari struktur graf. Karakteristik ini merupakan keterbatasan metodologis yang melekat pada pendekatan pseudo-labeling berbasis fitur graf, dan perlu menjadi perhatian ketika interpretasi hasil diperluas ke konteks praktis yang menggunakan definisi konektivitas dari sumber eksternal, misalnya volume penumpang aktual atau klasifikasi hub versi otoritas penerbangan.

5. Simpulan

Penelitian ini telah membuktikan bahwa klasifikasi tingkat konektivitas bandara global (Tinggi, Sedang, Rendah) dapat dilaksanakan secara efektif menggunakan model Graph Convolutional Network (GCN). Arsitektur GCN mampu mengatasi kelemahan metode analitik statis dengan cara mengagregasi fitur atribut spasial (koordinat geografis) secara simultan dengan struktur relasional rute penerbangan berbobot maskapai melalui matriks ketetanggaan. Algoritma Fuzzy C-Means (FCM) terbukti efektif menghasilkan label acuan semu (pseudo-ground truth) yang objektif pada dataset yang tidak berlabel.

Hasil pengujian pada 1.155 data uji menunjukkan bahwa model GCN yang dilatih selama 900 Epoch mencapai kinerja optimal dengan tingkat Akurasi keseluruhan 95% dan Macro F1-Score 79%. Penalti bobot kelas (class weighting) mampu mengatasi tantangan ketimpangan kelas yang ekstrem (imbalanced data), dibuktikan dengan perolehan Recall sempurna (1,00) dan nilai AUC 1,00 pada bandara kelas Tinggi (Global Hubs), serta AUC 0,97 pada kelas Sedang. Untuk penelitian selanjutnya, disarankan untuk mengintegrasikan fitur eksternal non-struktural seperti Produk Domestik Bruto (PDB) regional dan kebijakan penerbangan, serta mengimplementasikan arsitektur Temporal Graph Neural Network (T-GNN) guna memodelkan dinamika perubahan konektivitas penerbangan dari waktu ke waktu. Selain class weighting, teknik augmentasi data berbasis graf seperti GraphSMOTE (Zhao, Zhang, & Wang, 2021) juga dapat dieksplorasi untuk mengatasi ketimpangan kelas secara lebih mendalam.

Referensi

- Bandara, D., and Riccardi, K. 2024. Graph Node Classification to Predict Autism Risk in Genes. *Genes*, 15(4): 447-458.
- Bezdek, J.C. 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press.

- Chen, H., Huang, Z., Xu, Y., Deng, Z., Huang, F., He, P., and Li, Z. 2022. Neighbor Enhanced Graph Convolutional Networks for Node Classification and Recommendation. arXiv preprint, arXiv:2203.16097.
- Dunn, J.C. 1973. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *Journal of Cybernetics*, 3(3): 32-57.
- Goodfellow, I., Bengio, Y., and Courville, A. 2016. *Deep Learning*. Cambridge: MIT Press.
- Guimerà, R., Mossa, S., Turttschi, A., and Amaral, L.A.N. 2005. The World-Wide Air Transportation Network: Anomalous Centrality, Community Structure, and Cities' Global Roles. *Proceedings of the National Academy of Sciences*, 102(22): 7794-7799.
- Hammond, D.K., Vandergheynst, P., and Gribonval, R. 2011. Wavelets on Graphs via Spectral Graph Theory. *Applied and Computational Harmonic Analysis*, 30(2): 129-150.
- Huang, Y., Yang, H., and Yan, Z. 2025. A Dynamic Multi-Graph Convolutional Spatial-Temporal Network for Airport Arrival Flow Prediction. *Aerospace*, 12(5): 395-410.
- Jiang, W., and Luo, J. 2022. Graph Neural Network for Traffic Forecasting: A Survey. *Expert Systems with Applications*, 207: 117921.
- Kadhim Abdulzahra, A.M., and Al-Qurabat, A.K.M. 2022. A Clustering Approach Based on Fuzzy C-Means in Wireless Sensor Networks for IoT Applications. *Karbala International Journal of Modern Science*, 8(4): 579-595.
- Kingma, D.P., and Ba, J. 2017. Adam: A Method for Stochastic Optimization. arXiv preprint, arXiv:1412.6980.
- Kipf, T.N., and Welling, M. 2017. Semi-Supervised Classification with Graph Convolutional Networks. *Proceedings of the International Conference on Learning Representations (ICLR)*, hal. 1-14.
- Nair, V., and Hinton, G.E. 2010. Rectified Linear Units Improve Restricted Boltzmann Machines. *Proceedings of the 27th International Conference on Machine Learning (ICML)*, hal. 807-814.
- Rafiee, A. 2024. Airports, Airlines, Planes, and Routes Update 2024 Dataset. Kaggle Repository.
- Russell, S., and Norvig, P. 2020. *Artificial Intelligence: A Modern Approach Fourth Edition*. Hoboken: Pearson Education.
- Seddik, M.E.A., Wu, C., Lutzeyer, J.F., and Vazirgiannis, M. 2022. Node Feature Kernels Increase Graph Convolutional Network Robustness. arXiv preprint, arXiv:2109.01785.
- Sejan, M.A.S., Rahman, M.H., Aziz, M.A., Baik, J.I., You, Y.H., and Song, H.K. 2023. Graph Convolutional Network Design for Node Classification Accuracy Improvement. *Mathematics*, 11(17): 3680-3695.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15(1): 1929-1958.
- Sun, X., and Wandelt, S. 2021. Robustness of Air Transportation as Complex Networks: Systematic Review of 15 Years of Research and Outlook into the Future. *Sustainability*, 13(11): 6446-6468.
- Ting, T.J., Li, X., Sanner, S., and Abdulhai, B. 2023. Revisiting Random Forests in a Comparative Evaluation of Graph Convolutional Neural Network Variants for Traffic Prediction. arXiv preprint, arXiv:2305.19292.
- Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Yu, P.S. 2019. A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1): 4-24.
- Zang, H., Zhu, J., and Gao, Q. 2022. Deep Learning Architecture for Flight Flow Spatiotemporal Prediction in Airport Network. *Electronics*, 11(23): 4058-4071.
- Zanin, M., and Lillo, F. 2013. Modelling the Air Transport with Complex Networks: A Short Review. *The European Physical Journal Special Topics*, 215(1): 5-21.
- Zhao, T., Zhang, X., and Wang, S. 2021. GraphSMOTE: Imbalanced Node Classification on Graphs with Graph Neural Networks. *Proceedings of the 14th ACM International Conference on Web Search and Data Mining (WSDM)*, 833-841.
- Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., and Sun, M. 2020. Graph Neural Networks: A Review of Methods and Applications. *AI Open*, 1: 57-81.