

PERKEMBANGAN PROJEK RANCANGAN SISTEM REKOMENDASI FILM DENGAN MODEL DEEP LEARNING BERBASIS GNN

Atta Zulfahrizan, Adamsyach Prana Walidin, Mhd Ilyasyah Drilanang,

Muhammad Raffi Akbar Tjg, Kana Saputra

Prodi Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Medan

Jl. William Iskandar Ps. V Medan, Sumatera Utara 20221

latifhamzah106@gmail.com.

ABSTRAK

Di tengah pesatnya perkembangan dunia digital, sistem rekomendasi film menjadi bagian penting dalam meningkatkan kenyamanan dan kepuasan pengguna saat mengakses platform hiburan. Namun, metode tradisional seperti *collaborative filtering* dan *content-based filtering* masih menghadapi kendala dalam memahami pola interaksi yang kompleks antara pengguna dan film. Penelitian ini hadir untuk menjawab tantangan tersebut dengan mengembangkan sistem rekomendasi berbasis *deep learning* menggunakan pendekatan *Graph Neural Network* (GNN). Tujuan utama dari penelitian ini adalah membangun model yang mampu merepresentasikan hubungan antara pengguna dan film secara lebih mendalam dan kontekstual. Dalam pengembangannya, digunakan dataset MovieLens untuk membentuk graf yang mencerminkan relasi rating antara pengguna dan film. Proses pelatihan model dilakukan dengan arsitektur GraphSAGE dua lapis yang mampu menyebarkan informasi antar node secara efektif dalam graf. Hasil evaluasi menggunakan *Mean Squared Error* (MSE) menunjukkan performa yang baik dengan nilai loss terbaik sebesar 0.2780. Selain itu, hasil uji menggunakan *cosine similarity* memperlihatkan bahwa model ini mampu merekomendasikan film yang relevan dengan preferensi pengguna. Temuan ini menunjukkan bahwa pendekatan GNN dapat menjadi solusi yang menjanjikan untuk meningkatkan kualitas rekomendasi serta pengalaman pengguna dalam menjelajahi konten digital.

Kata kunci : Sistem Rekomendasi, Deep Learning, Graph Neural Network, GraphSAGE, MovieLens, Cosine Similarity.

1. PENDAHULUAN

Dalam era digital saat ini, sistem rekomendasi telah menjadi bagian penting dalam berbagai platform hiburan, termasuk layanan streaming film seperti Netflix, Disney+, dan Hulu. Sistem ini bertujuan untuk memberikan rekomendasi yang relevan kepada pengguna berdasarkan preferensi dan pola perilaku mereka. Metode tradisional seperti *collaborative filtering* dan *content-based filtering* telah banyak digunakan, tetapi memiliki keterbatasan dalam menangani hubungan kompleks antara pengguna dan film, terutama dalam skenario dengan data yang tidak lengkap atau sangat besar. Oleh karena itu, pendekatan berbasis *deep learning*, khususnya dengan *Graph Neural Networks* (GNN), menjadi semakin populer karena kemampuannya dalam menangkap hubungan kompleks dalam bentuk graf.

Graph Neural Networks (GNN) merupakan salah satu metode *deep learning* yang dirancang untuk bekerja dengan data berbentuk graf, di mana entitas seperti pengguna dan film direpresentasikan sebagai node, sedangkan hubungan atau interaksi di antara mereka direpresentasikan sebagai edge. Dengan menggunakan GNN, sistem dapat memahami hubungan tidak langsung antara pengguna dan film, misalnya melalui preferensi pengguna lain dengan pola serupa. Selain itu, GNN memungkinkan pemodelan konteks yang lebih kaya, seperti genre, aktor, dan ulasan, yang dapat memperkuat prediksi rekomendasi. *Graph Neural Network* (GNN) memiliki keunggulan utama dibandingkan metode lain (CNN &

RNN) karena kemampuannya untuk secara efektif memodelkan data yang memiliki struktur graf, seperti hubungan antar entitas dalam jaringan sosial, sistem rekomendasi, atau koneksi dalam jaringan sensor. Berbeda dengan metode konvensional yang mengandalkan data berbentuk vektor atau grid, GNN mampu menangkap dependensi kompleks antar node melalui proses agregasi informasi dari tetangga-tetangganya, sehingga menghasilkan representasi yang lebih kontekstual dan akurat. Keunggulan ini membuat GNN sangat unggul dalam tugas-tugas seperti prediksi link, klasifikasi node, dan deteksi komunitas. Dibandingkan metode konvensional, GNN lebih efektif dalam menangani *cold-start problem*, yaitu kondisi ketika terdapat pengguna baru atau film baru dengan sedikit data historis.

Penerapan GNN dalam sistem rekomendasi film juga memberikan keuntungan dalam menangani skala besar dan dinamika data yang terus berubah. Dengan kemampuan propagasi informasi melalui graf, model dapat melakukan pembelajaran representasi yang lebih kaya dan menangkap hubungan latar belakang yang lebih kompleks. Hal ini dapat meningkatkan akurasi rekomendasi, sehingga pengalaman menonton pengguna menjadi lebih personal dan memuaskan. Dengan perkembangan teknologi *deep learning* yang pesat, pemanfaatan GNN dalam rekomendasi film berpotensi menjadi solusi unggulan dalam menyajikan rekomendasi yang lebih cerdas, relevan, dan adaptif terhadap preferensi pengguna. Sistem pembelajaran dan manajemen proyek berbasis teknologi sangat

diperlukan untuk siswa dan karyawan di era digital yang semakin berkembang.

2. TINJAUAN PUSTAKA

2.1. Deep Learning

Deep learning adalah sebuah pendekatan untuk penyelesaian masalah pada sistem pembelajaran komputer yang menggunakan konsep hierarki; itu memanfaatkan banyak lapisan pengolahan informasi nonlinier untuk melakukan ekstraksi fitur, pengenalan pola, dan klasifikasi. Jaringan ini terdiri dari neuron atau unit, yaitu simpul-simpul berlapis yang saling terhubung. Setelah mendapatkan informasi dari lapisan di atasnya, setiap neuron menerapkan fungsi aktivasi dan mengeluarkan keluaran, yang merupakan masukan bagi lapisan di atasnya. Pembaruan dari neural network adalah bagian dari kecerdasan buatan yang digunakan untuk deteksi citra digital, pengenalan suara, dan tugas lainnya. [1], [4], [9].

2.2. Graph Neural Network (GNN)

GNN adalah metode berbasis pembelajaran mendalam yang berjalan pada domain grafik. Karena kinerjanya yang luar biasa, GNN baru-baru ini menjadi metode analisis grafis yang banyak digunakan. Struktur data yang memodelkan sekumpulan objek (simpul) dan hubungannya (tepi) disebut grafik. Analisis grafik berkonsentrasi pada tugas-tugas seperti pengelompokan, prediksi tautan, dan klasifikasi simpul. Ini adalah jenis pembelajaran mesin yang unik dengan struktur data non-Euclidean. Metode berbasis grafik telah sangat populer dalam beberapa tahun terakhir dan terbukti menjadi cara yang efisien untuk penalaran relasi. CRF dan jaringan random walk diusulkan berdasarkan model grafik untuk segmentasi gambar yang efektif. Menurut cara seseorang mendefinisikan operator konvolusi grafik, GNN dapat dibagi menjadi dua jenis. Graph Neural Network juga disebut Graph Convolution Network (GCN) oleh banyak orang. Pelabelan peran semantik, klasifikasi relasi, dan terjemahan mesin adalah beberapa tugas pengolahan bahasa natural, di mana GCN digunakan untuk mengkodekan struktur sintaksis kalimat. Dalam metode ini, setiap simpul pusat dan simpul tetangga digabungkan dengan mendefinisikan fungsi agregasi. Metode spasial bertujuan untuk menggabungkan setiap simpul pusat dan simpul tetangga dengan menentukan fungsi agregasi dari domain simpul. Graph attention network (GAT) dan GraphSAGE adalah contoh metode yang biasa digunakan dalam kategori ini. [6], [8], [9].

2.3. Graf

Graf adalah struktur yang digunakan untuk menunjukkan objek diskrit dan hubungan mereka satu sama lain. Graf biasanya digambarkan sebagai pasangan dua nilai atau tuple: $G = (V, E)$ di mana V adalah himpunan tidak kosong (vertices) dan E adalah himpunan sisi yang menghubungkan dua simpul. Graf terbagi menjadi beberapa jenis berdasarkan apakah

mereka memiliki gelang atau sisi ganda. Yang pertama adalah graf sederhana, yang tidak memiliki sisi ganda, dan yang kedua adalah graf sisi ganda, yang merupakan kebalikannya, dengan sisi ganda. Selain itu, ada jenis khusus dari graf ganda, yaitu graf semu, yang memiliki sisi gelang. [2], [3].

2.4. Penelitian Terdahulu

Berdasarkan Penelitian “Perbandingan Penerapan Relational Database dan Graph Database dalam Sistem Rekomendasi Film” Penelitian ini membandingkan kinerja database relasional (PostgreSQL) dan database graf (Neo4j) dalam sistem rekomendasi film, dengan fokus pada teknik content-based filtering, collaborative filtering, dan hybrid filtering. Hasilnya menunjukkan bahwa Neo4j memiliki waktu latensi yang lebih rendah dan penggunaan memori yang lebih efisien dibandingkan PostgreSQL. Yang menjadi pembeda ialah Penelitian yang digunakan merupakan pendekatan deep learning berbasis graf untuk memahami hubungan antar film dan preferensi pengguna secara lebih kompleks, sedangkan penelitian ini, lebih berfokus pada perbandingan efisiensi database relasional (PostgreSQL) dan graf (Neo4j) dalam sistem rekomendasi, tetapi tidak menggunakan deep learning atau GNN. [5].

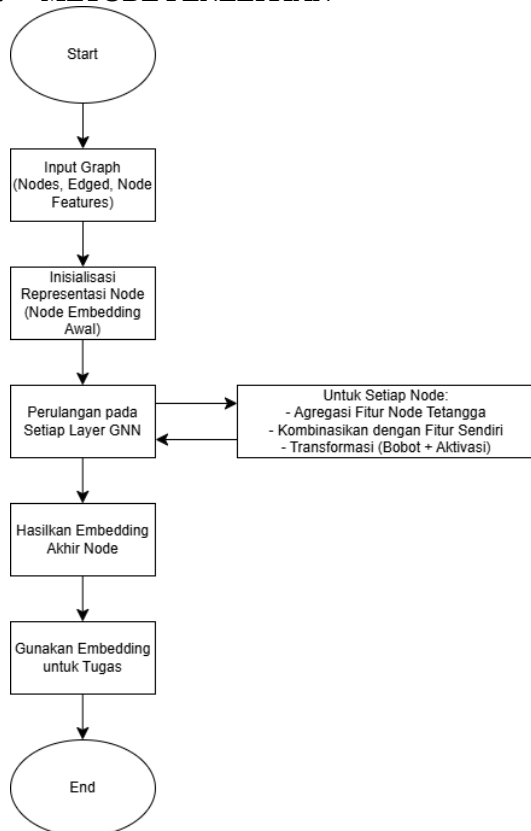
Berdasarkan Penelitian “Sistem Rekomendasi Film Menggunakan Metode K-Nearest Neighbors (K-NN) Sistem Rekomendasi Film Menggunakan Metode K-Nearest Neighbors (K-NN).” Penelitian ini mengeksplorasi implementasi metode K-NN dalam mengembangkan sistem rekomendasi film, dengan fokus pada peningkatan akurasi dan relevansi rekomendasi. Pengujian menunjukkan bahwa metode K-NN mencapai akurasi dan recall sebesar 45,4%, serta precision sebesar 100%, sehingga layak diterapkan dalam sistem rekomendasi film. Yang menjadi pembeda ialah Jurnal ini Menggunakan pendekatan berbasis kedekatan antar data, bukan hubungan berbasis graf. Dan juga Menggunakan metode *K-Nearest Neighbors (K-NN)*, yang merupakan algoritma berbasis kedekatan data, bukan deep learning atau model berbasis graf. [10].

Kemudian penelitian selanjutnya membahas analisis dataset Software Defined Network Intrusion (SDNI) menggunakan algoritma Deep Learning H2O. oleh Rizal, Martanto, dan Yudhistira Arie Wijaya (2022), mahasiswa Program Studi Sistem Informasi di STMIK IKMI Cirebon. Dalam penelitian ini, metode klasifikasi Deep Learning H2O dan aplikasi Rapid Miner digunakan untuk menganalisis dataset Software Defined Network Intrusion (SDNI). Hasil utama penelitian ini adalah algoritma Deep Learning H2O, yang memiliki kemampuan untuk menganalisa nilai recall sebesar 100.00% dan tingkat akurasi sebesar 99.66% pada dataset yang digunakan. [11].

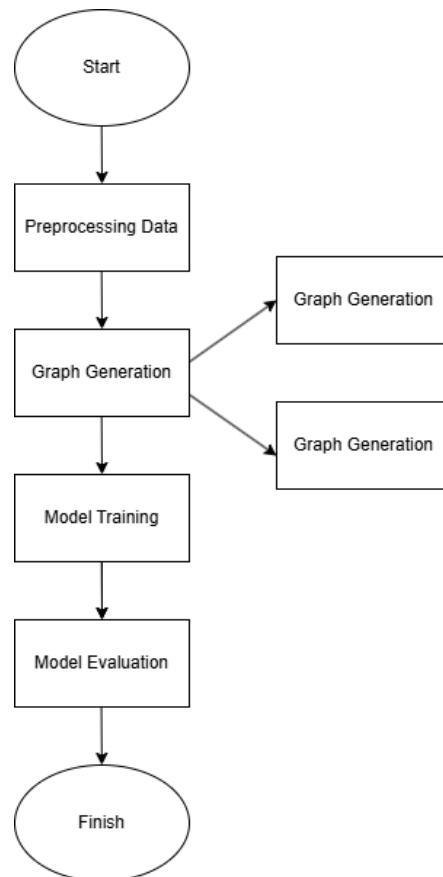
Dan penelitian dari Teknik Informatika di Universitas Muhammadiyah Riau dan menyelesaikan tesis berjudul “Klasifikasi Teks Berita Bahasa

Indonesia dengan Penggunaan Pembelajaran Mesin dan Pembelajaran Mendalam: Studi Literatur". Metode pembelajaran mesin dan pembelajaran mendalam digunakan dalam penelitian ini untuk mengklasifikasikan teks berita berbahasa Indonesia. Penelitian ini menggunakan berbagai algoritma pengajaran mesin dan pengajaran mendalam, termasuk Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Multinomial Naive Bayes, Long Short-Term Memory, Multilayer Perceptron, dan K-Nearest Neighbor (KNN). Hasil utama dari penelitian ini adalah bahwa Support Vector Machine (SVM) memberikan tingkat akurasi terbaik dari semua algoritma pengajaran mesin lainnya dengan nilai akurasi sebesar 94,24%. Sementara itu, Long Short-Term Memory mendapatkan nilai accuracy sebesar 95%. [12].

3. METODE PENELITIAN



Gambar 1. Tahapan Deep Learning GNN



Gambar 2. Tahapan Penelitian

3.1. Metode Pengumpulan Data

Data yang akan digunakan pada penelitian ini menggunakan dataset dari website MovieLens yang bernama MovieLens 32M. Ukuran dari dataset ini berada di sekitar 87 ribu data yang akan dipakai untuk penelitian ini.

Dataset yang dipakai pada penelitian ini berbentuk seperti ini:

Tabel 1. Tabel Dataset Ratings

userId	movieId	rating	timestamp
1	17	4	944249077
1	25	1	944250228
1	29	2	943230976
1	30	5	944249077
1	36	5	943228858

Tabel 2. Tabel Dataset Movie

movieId	title	genres
1	Toy Story (1995),	Adventure Animation Children Comedy Fantasy
2	Jumanji (1995)	Adventure Children Fantasy
3	Grumpier Old Men (1995)	Comedy Romance
4	Waiting to Exhale (1995)	Comedy Drama Romance
5	Father of the Bride Part II (1995)	Comedy

3.2. Preprocessing Data

Pada data yang diperoleh, dilakukan pembersihan dulu jika ada nilai yang kosong ataupun duplikat. Setelah itu ID pengguna dan ID film

dikonversi ke indeks numerik. Dan setelah itu dilakukan normalisasi pada data rating menggunakan standardScaler untuk menormalkan rating ke skala yang beragam

3.3. Pembangunan Graph

Karena GNN adalah Neural Network yang berbasis graph, maka data yang sudah disiapkan sebelumnya itu dikonversi kedalam bentuk graph dengan bantuan Deep Graph Library (DGL) dengan struktur:

- a. Node:
 - User (pengguna)
 - Movie (film)
- b. Edge:
 - User memberikan rating pada Movie (rates)
 - Movie dinilai oleh User (rated-by)

User Node merepresentasikan rata-rata rating yang diberikan oleh pengguna, dan Movie Node merepresentasikan rata-rata rating yang diterima oleh film.

3.4. Arsitektur Graph Neural Network

Model GNN yang dipakai pada penelitian ini adalah GraphSAGE yang terdiri dari dua lapisan graphSAGE convolution untuk menyebarkan informasi antar node. Lapisan pertama adalah lapisan fully connected (FC) untuk mengubah embedding node menjadi representasi akhir, dan lapisan output untuk menghasilkan prediksi rating berdasarkan representasi akhir.

3.5. Pelatihan Model

Pada saat melatih model, digunakan loss function yang menggunakan MSE untuk mengukur kesalahan prediksi rating, dan dilakukan sebanyak 20 epoch. Juga dengan in-feats 1, hidden feats 16, out feats = 8.

Rumus dari MSE dapat dilihat seperti berikut:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Dengan :

- n = jumlah data
- y_i = nilai aktual ke-i
- $\hat{y}_i$ = nilai prediksi ke-i
- $(y_i - \hat{y}_i)^2$ = selisih kuadrat antara nilai aktual dan prediksi

3.6. Evaluasi Model

Model yang dihasilkan dievaluasi menggunakan MSE untuk mengukur seberapa jauh prediksi rating dari nilai asli.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Training Model

Setelah dilakukan pelatihan model dengan GraphSAGE selama 20 epoch, diperoleh nilai loss terbaik sebesar 0.2780. Loss ini dihitung menggunakan Mean Squared Error (MSE) yang mengukur seberapa jauh prediksi rating terhadap rating asli pengguna. Semakin kecil nilai loss, semakin baik model dalam memprediksi rating film yang akan diberikan pengguna.

Tabel 3. Tabel Hasil Training Model

Epoch	Loss
1	0.4841
5	0.4019
10	0.3476
15	0.3094
20	0.2780

Grafik loss selama proses training menunjukkan bahwa model mengalami penurunan loss secara bertahap, menandakan bahwa model berhasil belajar dari data tanpa mengalami overfitting yang signifikan.

4.2. Hasil Rekomendasi Film

Setelah model dilatih, sistem rekomendasi diuji dengan mengambil beberapa sampel film secara acak dan menghitung kemiripan (similarity) antara film yang disukai pengguna dengan film lainnya dalam dataset.

Metode yang digunakan untuk menghitung kemiripan ini adalah dengan cosine similarity, yang membandingkan embedding film hasil dari Graph Neural Network. Film yang memiliki embedding yang mirip akan direkomendasikan kepada pengguna yang menyukai film tertentu.

Dari hasil pengujian, model menghasilkan rekomendasi film sebagai berikut:

Tabel 4. Tabel Hasil Training Model

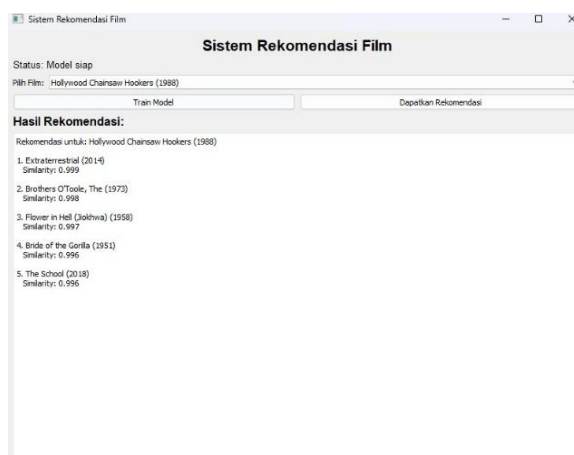
Stat	Title	Similarity
Movie title	Interstellar (2014)	
Movie recommendation	The Real Inglorious Bastards (2012)	0.9999
	Whiplash (2014)	0.9998
	Nina's Heavenly Delights (2006)	0.9998
	Inception (2010)	0.9997
	Zoey to the Max (2015)	0.9997

Hasil pelatihan model menunjukkan bahwa film yang direkomendasikan memiliki tingkat kemiripan yang sangat tinggi dengan *Interstellar (2014)*. Hal ini mengindikasikan bahwa berdasarkan embedding yang dihasilkan oleh model GNN, film-film tersebut memiliki pola preferensi pengguna yang serupa.



Gambar 3. GUI Latih untuk sistem

Sistem latih dalam penelitian ini dirancang menggunakan PyQt5, sebuah modul berbasis framework Qt untuk pengembangan antarmuka grafis. Antarmuka utama dibangun dengan QMainWindow sebagai wadah utama, sementara komponen interaktif seperti QPushButton digunakan untuk menginisiasi proses pelatihan model berbasis Deep Graph Library (DGL). QLabel digunakan untuk menampilkan status pelatihan model secara real-time, sehingga pengguna dapat memantau perkembangan proses. Selain itu, apabila terjadi kesalahan atau sistem memerlukan konfirmasi dari pengguna, QMessageBox digunakan untuk menampilkan pesan error maupun informasi penting lainnya. Sistem ini diintegrasikan dengan input.py sebagai skrip utama yang menghubungkan GUI dengan model DGL, sehingga mempermudah pengguna dalam melatih model secara langsung melalui antarmuka grafis tanpa perlu berinteraksi langsung dengan kode program.



Gambar 4. GUI sistem rekomendasi

Sistem rekomendasi dalam penelitian ini juga dikembangkan menggunakan PyQt5 dengan desain antarmuka yang intuitif. QComboBox digunakan sebagai dropdown untuk memilih film yang akan diproses oleh model rekomendasi, memungkinkan pengguna memilih input dengan mudah. Setelah film dipilih, pengguna dapat menekan tombol QPushButton untuk menjalankan proses rekomendasi. Hasil rekomendasi kemudian ditampilkan pada QLabel, memberikan informasi mengenai film yang disarankan oleh model. Untuk meningkatkan pengalaman pengguna, QMessageBox digunakan dalam menangani notifikasi atau peringatan terkait proses rekomendasi. Dengan pendekatan ini, sistem rekomendasi dapat dijalankan dengan efisien melalui antarmuka grafis yang sederhana dan mudah digunakan.

5. KESIMPULAN DAN SARAN

Penelitian ini membuktikan bahwa pengembangan sistem rekomendasi film menggunakan Graph Neural Network (GNN) berbasis arsitektur GraphSAGE mampu meningkatkan akurasi dan relevansi rekomendasi dengan memahami

hubungan kompleks antara pengguna dan film dalam bentuk graf. Hasil pelatihan model selama 20 epoch menunjukkan nilai loss terbaik sebesar 0,2780 menggunakan fungsi Mean Squared Error (MSE), dengan tren penurunan loss yang konsisten tanpa indikasi overfitting. Pada tahap pengujian, sistem mampu merekomendasikan film dengan tingkat kemiripan yang sangat tinggi, seperti "The Real Inglorious Bastards (2012)" dan "Whiplash (2014)" terhadap film "Interstellar (2014)", dengan nilai similarity mendekati 0,999, sebagaimana ditunjukkan dalam Tabel 4. Dengan demikian, pendekatan GNN dinilai efektif dalam membangun sistem rekomendasi film yang lebih personal dan akurat, meskipun tantangan terkait kebutuhan komputasi dan optimasi parameter model masih perlu diperhatikan dalam pengembangan selanjutnya.

DAFTAR PUSTAKA

- [1] R. Yunanto, A. P. Purfini, dan A. Prabuwisesa, "Survei Literatur: Deteksi Berita Palsu Menggunakan Pendekatan Deep Learning," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 11, no. 2, 2021.
- [2] M. G. A. Giffera, "Penerapan Graph Neural Network dalam Pembangunan Sistem Rekomendasi," *Makalah IF2120 Matematika Diskrit*, Institut Teknologi Bandung, 2021.
- [3] Z. N. S. Sudrajat, "Aplikasi Graph pada Graph Neural Network," *Makalah IF2120 Matematika Diskrit*, Institut Teknologi Bandung, 2020.
- [4] I. Hoeronis dan B. R. Trilaksono, "Node Classification on The Citation Network Using Graph Neural Network," *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, vol. 13, no. 1, pp. 96–105, Jun. 2023.
- [5] J. Zhou et al., "Graph Neural Networks: A Review of Methods and Applications," *AI Open*, vol. 1, pp. 57–81, Apr. 2021.
- [6] Y. Chen et al., "Graph-Based Global Reasoning Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.
- [7] F. Charli, H. Syaputra, M. Akbar, S. Sauda, dan F. Panjaitan, "Implementasi Metode Faster Region Convolutional Neural Network (Faster R-CNN) untuk Pengenalan Jenis Burung Lovebird," *Journal of Information Technology Ampera*, vol. 1, no. 3, pp. 185–197, Des. 2020.
- [8] R. Li et al., "A Graph Neural Network Model for the Diagnosis of Lung Adenocarcinoma Based on Multimodal Features and an Edge-Generation Network," *Quantitative Imaging in Medicine and Surgery*, vol. 13, no. 8, pp. 5333–5348, 2023.
- [9] T. I. Solihati, R. Kania, dan Rudianto, "Analisis Trend dan Pemetaan Penelitian Mahasiswa Teknik Informatika Menggunakan Graph Convolutional Network," *Jurnal Teknik Informatika dan Sistem Informasi (JATISI)*, vol. 10, no. 1, pp. 915–925, Mar. 2023.

- [10] L. Pujiastuti dan M. Wahyudi, “Advanced Graph Neural Networks for Dynamic Yield Optimization and Resource Allocation in Industrial Systems,” *Jurnal Teknik Informatika C.I.T Medicom*, vol. 16, no. 2, pp. 90–102, Mei 2024.
- [11] Rizal, Martanto, Y.A. Wijaya, “ANALISA DATASET SOFTWARE DEFINED NETWORK INTRUSION MENGGUNAKAN ALGORITMA DEEP LEARNING H2O”, *JATI: (Jurnal Mahasiswa Teknik Informatika)* vol. 6, no. 2, pp. 747-757, September 2022.
- [12] Alfando, R. Hayami, “KLASIFIKASI TEKS BERITA BERBAHASA INDONESIA MENGGUNAKAN MACHINE LEARNING DAN DEEP LEARNING: STUDI LITERATUR”, *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 681-686, Februari 2023.