

Pendayagunaan Relasi Wilayah dengan *Graph of Graph* untuk Klasifikasi Jenis Usaha Koperasi Merah Putih

Anthony Edbert Feriyanto, Vissuta Gunawan Lim, M. Arvin Wijayanto, Adila Alfa Krisnadhi

¹Fakultas Ilmu Komputer Universitas Indonesia, Depok, 16424, Indonesia, email: anthony.edbert@ui.ac.id

²Fakultas Ilmu Komputer Universitas Indonesia, Depok, 16424, Indonesia, email: vissuta.gunawan@ui.ac.id

³Fakultas Ilmu Komputer Universitas Indonesia, Depok, 16424, Indonesia, email: m.arvin@ui.ac.id

⁴Fakultas Ilmu Komputer Universitas Indonesia, Depok, 16424, Indonesia, email: adila@cs.ui.ac.id

Corresponding Author: Anthony Edbert Feriyanto

INTISARI — Program Koperasi Desa Merah Putih yang menargetkan pembentukan 80.000 koperasi di seluruh Indonesia menghadapi tantangan dalam menentukan jenis usaha yang tepat sesuai karakteristik dan potensi lokal. Dari 79.882 koperasi yang telah terbentuk per Juni 2025, hanya 47% yang berbadan hukum, mengindikasikan kesulitan dalam tahap penentuan usaha. Dataset penelitian mencakup 25.270 koperasi desa di enam provinsi Pulau Jawa yang diperkaya dengan data demografi, sosial-ekonomi, infrastruktur, dan geografis dari berbagai sumber terbuka. Penelitian ini menggabungkan pemodelan hierarkis desa-kecamatan melalui Graph of Graph Neural Network (GoGNN) sebagai *message passer* dengan TabPFN sebagai *classifier head* untuk menangani keterbatasan data tingkat desa. Evaluasi menggunakan *macro-averaged* F1-score menempatkan Graph-of-Graphs + TabPFN sebagai yang terbaik (0.93), melampaui CatBoost, TabPFN murni, dan varian GNN lainnya (0.91). Analisis *explainability* menggunakan GNNExplainer mengonfirmasi bahwa struktur hierarkis memungkinkan transfer informasi lintas-kecamatan yang memperkaya konteks spasial untuk prediksi yang lebih akurat. Sistem ini diharapkan dapat membantu pengambilan keputusan berbasis data dalam pengembangan koperasi desa dan mendukung pencapaian target program nasional.

KATA KUNCI — Kopdesa Merah Putih, Jenis Usaha Koperasi, *Graph Neural Network*, GoGNN, TabPFN, Klasifikasi Multi-label.

I PENDAHULUAN

A. LATAR BELAKANG

Koperasi Desa Merah Putih merupakan program nasional yang digagas Kementerian Koperasi dan UKM untuk menghadirkan koperasi di setiap desa di Indonesia. Dengan target ambisius membentuk lebih dari 80 ribu koperasi baru, setiap koperasi dirancang menjadi pusat kegiatan ekonomi warga setempat [1, 2]. Skala program ini terbilang luar biasa, didukung anggaran lebih dari Rp400 triliun, menjadikannya salah satu investasi pembangunan ekonomi desa terbesar dalam sejarah Indonesia [3].

Per Juni 2025, sekitar 79.882 koperasi desa Merah Putih telah terbentuk di hampir seluruh wilayah Indonesia, setara dengan $\pm 95\%$ dari target 80 ribu koperasi yang digagas oleh pemerintah [4]. Dari jumlah itu, hanya sekitar 37.300 unit ($\pm 47\%$) yang telah resmi berbadan hukum per awal Juni 2025 [5]. Sesuai skema pelaksanaan program, setiap desa terlebih dahulu menggelar musyawarah desa khusus untuk menentukan jenis usaha koperasi yang akan dijalankan, kemudian menyusun proposal, diverifikasi, dan dilegalisasi sebelum koperasi bisa operasional [4]. Artinya, meski mayoritas koperasi telah terbentuk secara administratif, banyak di antaranya yang masih berada pada tahap penentuan usaha atau proses legalisasi.

Menentukan jenis usaha koperasi desa bukanlah perkara mudah. Banyak desa menghadapi keterbatasan kapasitas SDM, minim literasi keuangan dan kelembagaan, serta potensi overlap dengan BUMDes lokal yang sudah ada [6]. Kondisi ini sering memaksa keputusan usaha dibuat hanya berdasarkan intuisi atau kebiasaan, sehingga memiliki risiko terhadap berkurangnya objektivitas dalam menentukan jenis usaha koperasi desa.

Pendekatan berbasis data menawarkan cara yang lebih objektif dan terukur untuk mengidentifikasi peluang usaha yang selaras dengan karakteristik desa. Sejumlah penelitian telah me-

nunjukkan bahwa data desa dapat dimanfaatkan untuk pemetaan dan pengambilan keputusan yang lebih akurat. Misalnya, analisis Indeks Pembangunan Desa (IDM) dengan algoritma klasifikasi terbukti efektif memetakan kategori desa seperti mandiri, berkembang, tertinggal [7, 8]. Studi lain memanfaatkan citra satelit penerangan malam untuk memprediksi indeks pembangunan desa dengan R^2 sekitar 0,5 [9]. Temuan-temuan ini mengindikasikan bahwa integrasi berbagai sumber data desa dapat menjadi fondasi kuat dalam merumuskan rekomendasi usaha koperasi yang lebih tepat sasaran dan berbasis bukti.

Melihat kompleksitas penentuan jenis usaha koperasi dan potensi besar data desa, penelitian ini mengusulkan pendekatan berbasis *Graph of Graph Neural Network (GoGNN)* sebagai inti solusi. GoGNN memungkinkan pemodelan hubungan hierarkis antar entitas, yaitu desa yang saling terhubung melalui kecamatan, sehingga informasi dapat dipertukarkan secara efektif di kedua tingkat wilayah tersebut. Pendekatan ini dipilih karena sebagian data penting, seperti indikator pembangunan tertentu dan data sektor ekonomi, hanya tersedia di tingkat kecamatan. Dengan memanfaatkan keterkaitan spasial dan sosial ini, model diharapkan mampu menangkap konteks yang sering terlewat jika hanya menggunakan data internal desa. Untuk mengatasi keterbatasan data tingkat desa, GoGNN diintegrasikan dengan TabPFN, sebuah *pretrained transformer-based model* yang tetap efektif pada skenario data pelatihan terbatas. Kombinasi kedua pendekatan ini diharapkan menghasilkan rekomendasi jenis usaha koperasi yang lebih akurat, kontekstual, dan relevan terhadap dinamika lokal maupun regional.

B. RUMUSAN MASALAH

Berdasarkan problema tersebut, tim peneliti merumuskan beberapa masalah yang akan dibahas dalam penelitian ini: **RQ1.** *Bagaimana TabPFN menghasilkan rekomendasi jenis usaha koperasi desa dibandingkan metode machine learning*

konvensional? **RQ2.** Bagaimana penggunaan pembelajaran berbasis graf memengaruhi kualitas rekomendasi dibandingkan pendekatan non-graf? dan **RQ3.** Bagaimana pemodelan graph dengan struktur hierarkis antarwilayah (desa-kecamatan) memengaruhi kualitas rekomendasi dibandingkan graf non-hierarkis, mengingat sebagian data hanya tersedia di tingkat kecamatan? Bagian III & IV membahas lengkap tentang metodologi yang digunakan untuk menjawab tiga pertanyaan tersebut.

C. TUJUAN PENELITIAN

Berdasarkan rumusan masalah, penelitian ini bertujuan untuk mengembangkan sistem rekomendasi jenis usaha koperasi desa yang akurat dan kontekstual melalui integrasi model TabPFN dan pembelajaran berbasis graf. Pendekatan ini memanfaatkan TabPFN sebagai model *pretrained transformer* yang efektif meski data tingkat desa terbatas, serta memanfaatkan keterkaitan antarwilayah melalui *Graph Neural Network* (GNN) dan *Graph of Graph Neural Network* (GoGNN) dengan struktur hierarkis desa-kecamatan. Dengan demikian, sistem ini diharapkan mampu mengevaluasi pengaruh pemodelan graph dan hierarki wilayah terhadap kualitas rekomendasi, sekaligus membandingkannya dengan pendekatan pembelajaran mesin konvensional dan non-graf.

D. KONTRIBUSI DAN MANFAAT

Penelitian ini sejalan dengan tema penguatan kemandirian ekonomi desa melalui pengembangan koperasi berbasis data. Secara praktis, penelitian ini diharapkan membantu pemerintah desa, pengelola koperasi, dan pihak terkait dalam menentukan jenis usaha yang sesuai dengan potensi lokal, sehingga mampu meningkatkan pendapatan dan kesejahteraan masyarakat. Hal ini mendukung pencapaian *Sustainable Development Goals* (SDGs), khususnya tujuan ke-8 (*Decent Work and Economic Growth*) dan ke-11 (*Sustainable Cities and Communities*).

Secara ilmiah, penelitian ini berkontribusi dengan memperkankan kerangka kerja yang memadukan TabPFN dan *Graph of Graph Neural Network* (GoGNN) untuk memproses data desa secara hierarkis, sehingga menghasilkan rekomendasi yang lebih akurat, kontekstual, dan relevan dibandingkan pendekatan konvensional.

E. BATASAN PENELITIAN

Metode yang dipaparkan dalam makalah penelitian ini memiliki beberapa batasan penelitian, antara lain: (1) Dataset yang digunakan terbatas pada wilayah enam provinsi di Pulau Jawa dan (2) Seluruh data yang digunakan dalam penelitian ini diperoleh per Juni 2025.

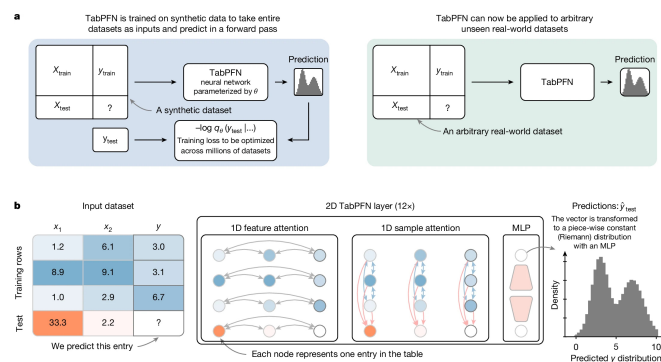
II STUDI LITERATUR

Kami meninjau metode yang digunakan dalam penelitian ini, yaitu TabPFN (II-A), Graph Neural Network (II-B) dan Graph of Graph Neural Network (II-C) yang dimanfaatkan sebagai mekanisme *message passing*.

A. TABPFN

TabPFN adalah model transformer yang dirancang khusus untuk data tabular berukuran kecil dengan pendekatan *prior-data fitted*, yaitu dilatih pada ribuan tugas sintesis beragam untuk memahami pola umum data tabular [10]. Berbeda dengan model konvensional, TabPFN menggunakan mekanisme *in-context learning* yang memungkinkan prediksi langsung pada data baru tanpa perlu pelatihan ulang.

Model ini menggunakan arsitektur *encoder transformer* dengan desain *set-invariant* dan skema *encoding* yang memisahkan data latih dan kueri [11]. Keunggulan utama TabPFN meliputi kecepatan inferensi tinggi dan performa kuat pada dataset kecil, seperti konteks data pelatihan tingkat desa. Hasil prediksi TabPFN seringkali lebih akurat dibandingkan metode konvensional seperti *gradient-boosted trees* [12, 13].



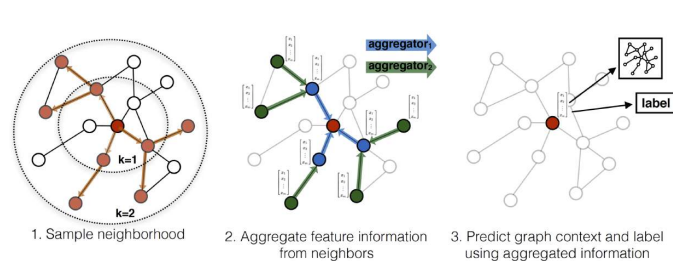
Gambar 1: Arsitektur TabPFN

B. GRAPH NEURAL NETWORK

Graph Neural Network (GNN) adalah pendekatan pembelajaran mesin untuk memproses data dengan keterkaitan antar entitas, seperti hubungan antar desa dalam wilayah administratif. GNN dapat memanfaatkan informasi dari "tetangga" untuk prediksi yang lebih akurat melalui mekanisme *message passing* [14].

Dalam konteks data desa, GNN memodelkan setiap desa sebagai titik (*node*) dengan hubungan antar desa sebagai garis penghubung (*edge*). Mekanisme *message passing* memungkinkan setiap node mengirimkan representasi fiturnya ke tetangga, kemudian mengagregasi informasi yang diterima untuk membentuk representasi yang lebih kontekstual [15].

GraphSAGE adalah salah satu varian induktif yang dapat melakukan generalisasi ke wilayah baru tanpa pelatihan ulang [16], sangat relevan untuk ekosistem koperasi desa yang terus berkembang. Pendekatan ini menangkap baik pola lokal maupun pola global [17].



Gambar 2: Arsitektur GraphSAGE dan Message Passing

C. GRAPH OF GRAPH NEURAL NETWORK

Graph of Graphs Neural Network (GoGNN) adalah pengembangan GNN untuk menangani struktur data hierarkis, seperti hubungan desa-kecamatan dalam sistem administratif Indonesia [18]. GoGNN dapat memodelkan interaksi pada multiple level secara simultan, baik hubungan *intra-level* (antar desa dalam kecamatan) maupun *inter-level* (antar kecamatan).

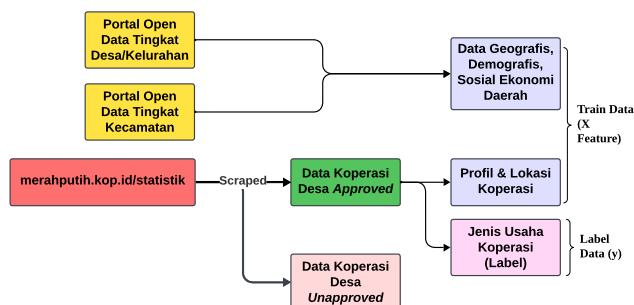
Arsitektur GoGNN terdiri dari *encoder* mikro yang memproses graf internal setiap kecamatan dan *encoder* makro yang memproses graf eksternal antar kecamatan. Kunci utama adalah mekanisme *dual-attention* yang memungkinkan informasi mengalir antar level secara adaptif melalui proses *coarsening* dan *pooling-unpooling* [19, 20].

Dengan memodelkan interaksi pada multiple level, arsitektur ini dapat menangkap dinamika pada skala yang berbeda. Dengan demikian, GoGNN dapat menghasilkan rekomendasi yang lebih komprehensif karena tidak hanya bergantung pada karakteristik lokal, tetapi juga tren regional yang lebih luas.

III METODOLOGI

Pada penelitian ini, kami membangun *dataset* (Bagian III-A), menyusun graf dan *Graph-of-Graphs* untuk konteks relational-spasial (Bagian III-C), lalu memodelkannya dengan GoGNN sebagai *message passer* dan TabPFN sebagai *classifier head* (Bagian III-D). Kinerja dinilai menggunakan *macro-averaged F1-score* (Bagian III-B).

A. KONSTRUKSI DATASET



Gambar 3: Konstruksi Dataset

Tim peneliti melakukan *web-scraping* pada portal merahputih.kop.id¹ untuk menghimpun data koperasi desa merah putih se-Indonesia. Proses *scraping* menghasilkan data dalam format JSON yang kemudian diolah menjadi format CSV untuk memudahkan analisis. Data yang diperoleh mencakup informasi profil koperasi, nama penanggung jawab, alamat lengkap, dan berbagai atribut koperasi lainnya untuk total 25.270 desa/kelurahan.

Dari hasil *scraping*, terdapat dua kategori koperasi: koperasi yang telah disetujui pemerintah dan berbadan hukum, serta koperasi yang masih dalam proses persetujuan. Penelitian ini memfokuskan pada koperasi yang telah disetujui karena memiliki informasi jenis usaha yang lengkap dan tervalidasi. Jenis usaha koperasi yang tercantum dalam data koperasi yang telah disetujui inilah yang dijadikan sebagai *ground truth* untuk proses klasifikasi. Dataset kemudian disaring khusus untuk enam provinsi di Pulau Jawa, yaitu Banten, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, dan Jawa Timur.

Pada tahap anotasi label, masalah ini dimodelkan sebagai klasifikasi multi-label karena satu koperasi dapat memiliki lebih dari satu jenis usaha/layanan. Distribusi proporsi jenis usaha koperasi yang menjadi target klasifikasi dapat dilihat pada Tabel 1.

Jenis Usaha Koperasi (Label)	Proporsi terhadap Total Data
gerai_sembako	98.5%
gerai_obat_murah	95%
kantor_koperasi	83.4%
unit_simpan_pinjam	96.6%
klinik_desa	84.4%
cold_storage	79%
logistik	77.7%
lainnya	70.9%

TABEL 1: Proporsi tiap jenis usaha koperasi (label) terhadap keseluruhan data

Setelah memperoleh dataset utama, data tersebut diperkaya dengan berbagai dataset pendukung. Dengan memanfaatkan informasi geografis koperasi meliputi lokasi desa, kecamatan, hingga kabupaten/kota, dilakukan proses *web-scraping* tambahan dan akses *open data* dari berbagai sumber terbuka. Fitur-fitur tambahan yang diperoleh mencakup aspek demografi, sosial-ekonomi, infrastruktur, dan geografis pada berbagai tingkat administratif (koperasi, kelurahan/desa, kecamatan, dan kabupaten/kota) yang relevan untuk wilayah enam provinsi tersebut. Rincian lengkap sumber, kategori, dan tingkat data yang digunakan dalam penelitian ini dapat dilihat pada Tabel 2.

¹ <https://merahputih.kop.id/>

Sumber Data	Kategori Data	Tingkat Spasial
bps.go.id, data.go.id, bi.go.id	Demografi; Sosial- Ekonomi	Desa/Kelurahan & Kecamatan
OpenStreetMap (OSM) API	Infrastruktur; Spasial	Desa/Kelurahan & Kecamatan
merahputih.kop.id	Unit Koperasi	Desa/Kelurahan
GeoPandas	Geolokasi	Desa/Kelurahan
Portal Open Data Prov/Kota (Jawa)	Demografi; Ketenagakerjaan	Desa/Kelurahan & Kota/Kabupaten

TABEL 2: Sumber dan kategori data beserta tingkat spasialnya

B. METRIK

Pada penelitian ini, pemilihan metrik evaluasi mempertimbangkan sifat *multilabel classification* dan distribusi label yang tidak seimbang sebagaimana ditunjukkan pada Tabel 1. Penggunaan metrik *accuracy* saja tidak direkomendasikan karena dapat memberikan hasil yang bias pada kelas dengan jumlah sampel dominan. Sebagai gantinya, penelitian ini menggunakan *macro-averaged F1-score* sebagai metrik utama. Pada pendekatan ini, F1-score dihitung secara terpisah untuk setiap label, kemudian dirata-ratakan tanpa mempertimbangkan proporsi jumlah sampel pada masing-masing label. Hal ini memastikan bahwa setiap label memiliki kontribusi yang sama terhadap skor akhir, terlepas dari distribusinya. Secara matematis, *macro-averaged F1-score* dihitung sebagai:

$$\text{Macro-F1} = \frac{1}{|\mathcal{L}|} \sum_{l \in \mathcal{L}} \frac{2 \text{Precision}_l \text{Recall}_l}{\text{Precision}_l + \text{Recall}_l} \quad (1)$$

C. KONSTRUKSI GRAPH DAN GRAPH OF GRAPHS

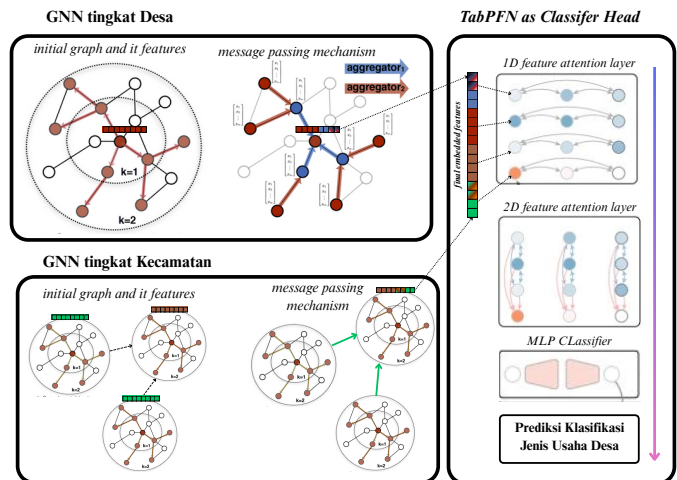
Berdasarkan data geolokasi dari Tabel 2, kami menggunakan Open Source Routing Machine di atas jaringan jalan OpenStreetMap untuk menghitung jarak rute d_{ij} dan waktu tempuh arah-spesifik ($t_{i \rightarrow j}, t_{j \rightarrow i}$) bagi setiap pasangan titik. Setiap edge menyimpan vektor bobot $\mathbf{w}_{ij} = [d_{ij}, t_{i \rightarrow j}, t_{j \rightarrow i}]$, dengan waktu tempuh sebagai sinyal utama karena lebih mewakili friksi mobilitas dibandingkan jarak geometris.

Dari proses ini terbentuk tiga graf: (i) graf antar-koperasi G^{coop} , (ii) graf antar-kecamatan G^{dist} , dan (iii) graf bipartit keterikatan koperasi-kecamatan G^{cont} dengan pemetaan $\phi : V^{\text{coop}} \rightarrow V^{\text{dist}}$. Asimetri waktu tempuh $t_{i \rightarrow j} \neq t_{j \rightarrow i}$ dipertahankan dalam representasi graf berarah.

Kami merakit *Graph-of-Graphs* (GoG) dua-level dengan level mikro memuat N node koperasi dan level makro memuat M node kecamatan. Untuk setiap kecamatan k , dibentuk graf terinduksi $G_k^{\text{micro}} := G^{\text{coop}}[V_k]$ dengan $V_k = \{i \in V^{\text{coop}} : \phi(i) = k\}$. Hubungan antarlevel direpresentasikan oleh matriks insidensi $C \in \{0, 1\}^{N \times M}$ yang memungkinkan aliran informasi *top-down* dan *bottom-up*. Struktur GoG global dapat dituliskan sebagai:

$$\mathcal{A} = \begin{bmatrix} A^{(\text{coop})} & C \\ C^\top & A^{(\text{dist})} \end{bmatrix}$$

D. ARSITEKTUR MODEL GOGNN-TABPFN



Gambar 4: Arsitektur Model GoGNN-TabPFN

Model yang diusulkan menggabungkan *Graph-of-Graphs Neural Network* (GoGNN) sebagai *message passer* dengan TabPFN sebagai *classifier head* untuk klasifikasi multi-label jenis usaha koperasi. Seperti ditunjukkan pada Gambar 4, arsitektur terdiri dari dua komponen utama yang bekerja secara terintegrasi.

GoGNN berperan sebagai *message passer* yang menangkap informasi struktural dan spasial melalui dua level: GNN tingkat kecamatan dan GNN tingkat desa/koperasi. Pada setiap level, dilakukan *message passing mechanism* dengan agregasi berbobot atensi berdasarkan koefisien $\alpha_{ij} \propto \exp(-\gamma c_{ij})$ yang memprioritaskan koneksi dengan mobilitas tinggi. Proses ini menghasilkan *embedded features* $\mathbf{z}_i \in \mathbb{R}^q$ yang merangkum konteks relasional dan spasial-operasional antar koperasi dan kecamatan.

TabPFN berfungsi sebagai *classifier head* yang menerima input gabungan antara fitur tabular asli x_i dan embedding dari GoGNN $\mathbf{z}_i$, yaitu $\tilde{x}_i = [\mathbf{z}_i; x_i]$. TabPFN kemudian memproses input ini melalui *feature attention layer* 1D dan 2D, diikuti oleh MLP classifier untuk menghasilkan prediksi multi-label.

Kombinasi ini memungkinkan model memanfaatkan kekuatan TabPFN dalam memodelkan fitur tabular numerik/kategorikal sekaligus memanfaatkan informasi kontekstual spasial-operasional dari GoGNN.

IV HASIL EKSPERIMEN DAN ANALISIS

Penelitian ini mengevaluasi pengaruh integrasi *Graph Neural Network* (GNN), *Graph-of-Graphs Neural Network* (GoGNN), dan TabPFN terhadap klasifikasi jenis usaha koperasi desa dengan menggunakan metrik Averaged F1-Macro sebagai evaluator. Seluruh eksperimen dilakukan menggunakan dataset koperasi desa di Pulau Jawa, sebagaimana dijelaskan pada Bagian III-A.

A. PERBANDINGAN MODEL

Tabel 3 menyajikan hasil rata-rata F1-Macro untuk berbagai konfigurasi model. CatBoost *Default Parameter* digunakan sebagai baseline pembelajaran mesin konvensional berbasis *gradient*

boosting. Hasilnya dibandingkan dengan TabPFN sebagai model transformer berbasis *amortized Bayesian inference*, serta varian model berbasis GNN dan GoGNN yang dikombinasikan dengan *classifier head* berbeda.

Model / Pendekatan	Averaged F1-Macro
CatBoost Default	0.9065
TabPFN	0.9129
SAGEConv + FFN Classifier	0.9184
SAGEConv + TabPFN	0.9280
Graph-of-Graphs + TabPFN	0.9321

TABEL 3: Perbandingan performa rata-rata F1-Macro berbagai pendekatan

Dari hasil pada Tabel 3, terlihat bahwa seluruh pendekatan yang memanfaatkan informasi relasional melalui GNN atau GoGNN mampu melampaui *baseline* CatBoost. Integrasi SAGEConv dengan *feed-forward network* sederhana sudah memberikan peningkatan dari 0.9065 menjadi 0.9184, mengindikasikan bahwa proses message passing antar node desa-koperasi mampu memperkaya representasi fitur.

Penggunaan TabPFN sebagai *classifier head* memberikan peningkatan tambahan, terlihat dari perbedaan antara SAGEConv + FFN (0.9184) dan SAGEConv + TabPFN (0.9280). Hal ini konsisten dengan karakteristik TabPFN yang efektif pada tabular data dengan distribusi kompleks. Pencapaian tertinggi diperoleh oleh *Graph-of-Graphs* + TabPFN dengan skor F1-Macro 0.9321. Keunggulan ini mengindikasikan bahwa pemodelan hierarki wilayah (desa–kecamatan) melalui GoGNN mampu menangkap konteks spasial dan sosial lintas tingkatan administratif secara lebih efektif dibandingkan GNN non-hierarkis.

B. GRAPH MODEL EXPLAINABILITY

Selain menghasilkan model dengan kemampuan prediksi yang baik, penting bagi penelitian ini untuk memastikan bahwa prediksi tersebut dapat dijelaskan secara transparan. Pendekatan interpretabilitas memungkinkan peneliti dan pemangku kepentingan memahami alasan di balik rekomendasi yang diberikan model sehingga meningkatkan kepercayaan dan memudahkan pengambilan keputusan berbasis data.

Untuk menelaah perbedaan perilaku antara graf non-hierarkis dan hierarkis, kami menerapkan *GNNExplainer* [21] pada dua varian model: SAGEConv (graf datar) dan GoGNN (graf hierarkis desa–kecamatan). *GNNExplainer* mengaproksimasi subgraf dan subset fitur paling berpengaruh bagi sebuah prediksi node target, sehingga kontribusi struktur (node/edge) dan fitur dapat dievaluasi secara bersamaan.

Gambar 5 menyajikan hasil interpretasi untuk satu kasus prediksi koperasi di Kecamatan West Karawang (node merah). Hasil untuk SAGEConv menunjukkan bahwa node-node berpengaruh (panah hijau/kuning) terutama berasal dari tetangga lokal dalam kecamatan yang sama, mengindikasikan ketergantungan pada konteks *ego-network*. Sebaliknya, hasil untuk GoGNN memperlihatkan bahwa desa-desa di kecamatan lain (mis. Telukjambe) juga tampil sebagai pemberi pengaruh bermakna. Hal ini dimungkinkan oleh struktur hierarkis desa–kecamatan

yang memungkinkan aliran informasi lintas-level, sehingga data yang hanya tersedia di tingkat kecamatan tetap berkontribusi pada prediksi desa dengan fitur terbatas.



Gambar 5: GNNExplainer pada SAGEConv (kiri) dan GoGNN (kanan). Node hijau: pengaruh besar; node kuning: pengaruh kecil; node merah: node yang akan diprediksi. GoGNN menangkap pengaruh lintas-kecamatan sehingga representasi lebih kaya dibandingkan graf datar yang didominasi tetangga intra-kecamatan.

Secara kualitatif, representasi GoGNN lebih informatif karena: (1) memanfaatkan konteks regional melalui koneksi antar-kecamatan sehingga cakupan ketetanggaan yang relevan melampaui *ego-network* lokal, (2) memungkinkan transfer sinyal dari wilayah dengan profil ekonomi/operasional serupa meski tidak berbatasan langsung, (3) mengompensasi kelangkaan fitur di level desa dengan meminjam informasi level kecamatan melalui mekanisme agregasi makro. Kualitas penjelasan ini selaras dengan peningkatan kinerja yang diamati, menunjukkan bahwa struktur hierarkis pada pemodelan graf tidak hanya menambah sumber informasi, tetapi juga memperluas ruang ketetanggaan efektif yang berkontribusi pada rekomendasi koperasi yang lebih akurat dan kontekstual.

V KESIMPULAN

Penelitian ini berhasil mengembangkan sistem klasifikasi jenis usaha koperasi desa berbasis Graph of Graph Neural Network (GoGNN) yang terintegrasi dengan TabPFN. Berdasarkan eksperimen pada dataset 25.270 koperasi desa di enam provinsi Pulau Jawa, diperoleh temuan penting yang menjawab ketiga *research questions*.

TabPFN terbukti superior dibandingkan metode konvensional (RQ1), menghasilkan peningkatan F1-Macro dari 0.9065 (CatBoost) menjadi 0.9129. **Pembelajaran berbasis graf konsisten meningkatkan kualitas rekomendasi** (RQ2), dengan SAGEConv + TabPFN mencapai 0.9280. **Struktur hierarkis GoGNN memberikan kontribusi signifikan** (RQ3), dengan pendekatan Graph-of-Graphs + TabPFN mencapai performa tertinggi 0.9321. Analisis *explainability* mengonfirmasi bahwa GoGNN mampu memanfaatkan konteks lintas-kecamatan untuk transfer informasi yang lebih efektif.

Kontribusi ilmiah penelitian ini adalah kerangka kerja novel yang memadukan TabPFN dan GoGNN untuk memproses data desa secara hierarkis, mengatasi tantangan keterbatasan data pada level granular. **Implikasi praktis** meliputi dukungan terhadap program Koperasi Desa Merah Putih melalui alat bantu pengambilan keputusan berbasis data yang objektif.

Penelitian masa depan dapat memperluas cakupan ke seluruh Indonesia, mengintegrasikan data temporal, dan mengeksplorasi *heterogeneous graph neural network* untuk menangani keragaman relasi yang lebih kompleks.

REFERENSI

- [1] Pemerintah Republik Indonesia, “Koperasi desa/kelurahan merah putih.” <https://merahputih.kop.id/>, 2025.
- [2] Pemerintah Republik Indonesia, “Peluncuran kelembagaan 80.000 koperasi desa merah putih.” <https://merahputih.kop.id/>, 2025. Pengumuman formal di Istana.
- [3] A. K. M. Simanjuntak, “Koperasi desa merah putih didukung uang pajak, ini pesan sri mulyani.” <https://news.ddtc.co.id/berita/nasional/1812098/>, 2025.
- [4] Suaramerdeka.com, “79.882 koperasi desa merah putih telah terbentuk per juni 2025.” kedu.suaramerdeka.com, 2025.
- [5] N. L. Anggela, “53% kopdes merah putih belum berbadan hukum per juni 2025,” *Bisnis Indonesia*, Juni 2025.
- [6] F. Arzewiniga and Zulkarnain, “Tantangan dan peluang kelembagaan koperasi desa merah putih,” *Jurnal Manajemen dan Inovasi*, 2025. Hasil kajian kelembagaan koperasi desa.
- [7] A. Rifai, S. E. Permana, and R. Hamonangan, “Optimalisasi klasifikasi indeks desa membangun menggunakan metode ensemble dan random forest,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 4, pp. 8226–8234, 2024.
- [8] M. R. Hendriawan and R. R. Marlina, “Pengelompokan desa di jawa barat berdasarkan indeks desa membangun (idm) menggunakan algoritma clara,” *Jurnal Pengabdian kepada Masyarakat (JPM)*, 2025.
- [9] C. D. Purnawanto, R. Nooraeni, and N. Widowati, “Estimating village development index based on satellite imagery using machine learning application,” 2023. Transfer learning from nighttime light imagery.
- [10] N. Hollmann, S. Müller, K. Eggensperger, and F. Hutter, “Tabpfn: A transformer that solves small tabular classification problems in a second,” in *Proceedings of the Eleventh International Conference on Learning Representations (ICLR)*, (Kigali, Rwanda), 2023.
- [11] S. Arik and T. Pfister, “Tabtransformer: Tabular data modeling using contextual embeddings,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, (Virtual Event), pp. 6679–6687, 2021.
- [12] Y. Gorishniy, I. Rubachev, V. Khrulkov, and A. Babenko, “Revisiting deep learning models for tabular data,” in *Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS)*, (Virtual Event), 2021.
- [13] S. Müller, M. Feurer, N. Hollmann, and F. Hutter, “PFNs4BO: In-context learning for bayesian optimization,” in *Proceedings of the 40th International Conference on Machine Learning (ICML)*, vol. 202 of *Proceedings of Machine Learning Research*, (Honolulu, HI, USA), pp. 25444–25470, PMLR, 2023.
- [14] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” in *Proceedings of the 5th International Conference on Learning Representations (ICLR)*, (Toulon, France), 2017.
- [15] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, “Neural message passing for quantum chemistry,” in *Proceedings of the 34th International Conference on Machine Learning (ICML)*, vol. 70 of *Proceedings of Machine Learning Research*, (Sydney, Australia), pp. 1263–1272, PMLR, 2017.
- [16] W. L. Hamilton, R. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*, (Long Beach, CA, USA), pp. 1024–1034, Curran Associates Inc., 2017.
- [17] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, “A comprehensive survey on graph neural networks,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, 2020.
- [18] H. Wang, D. Lian, Y. Zhang, L. Qin, and X. Lin, “Gognn: Graph of graphs neural network for predicting structured entity interactions,” in *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, (Yokohama, Japan), pp. 1317–1323, 2020.
- [19] R. Ying, J. You, C. Morris, W. L. Hamilton, and J. Leskovec, “Hierarchical graph representation learning with differentiable pooling,” in *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NeurIPS)*, (Montréal, Canada), pp. 4805–4815, Curran Associates Inc., 2018.
- [20] J. Baek, M. Kang, and S. J. Hwang, “Accurate learning of graph representations with graph multiset pooling,” in *Proceedings of the 9th International Conference on Learning Representations (ICLR)*, (Virtual Event), 2021.
- [21] R. Ying, D. Bourgeois, J. You, M. Zitnik, and J. Leskovec, “Gnnexplainer: Generating explanations for graph neural networks,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 32, 2019.