

Rekomendasi Produk Berbasis *Collaborative Filtering* Menggunakan *Factorization machine Graph Convolutional Networks*

Sherly Rosa Anggraeni¹, Diana Purwitasari^{2,*}, Chastine Fatichah³, Yoga Yustiawan⁴

^{1,2,3}Teknik Informatika, Institut Teknologi Sepuluh Nopember, Indonesia

⁴Bank Rakyat Indonesia, Indonesia

¹sherly.rosa@gmail.com; ²diana@if.its.ac.id, ³chastine@if.its.ac.id;

*corresponding author

INFO ARTIKEL

Sejarah Artikel

Diterima: 30 Juni 2023

Direvisi: 28 Juli 2023

Diterbitkan: 30 Agustus 2023

Kata Kunci

Faktorisasi Matriks

Graph Convolutional Networks

Rekomendasi Produk

ABSTRAK

Rekomendasi produk memiliki peran signifikan dalam berbagai industri, termasuk *e-commerce*, ritel, perhotelan, dan keuangan. Rekomendasi dapat meningkatkan kepuasan pelanggan dan penjualan dengan membantu pelanggan menemukan produk yang relevan. Pendekatan *collaborative filtering* digunakan dalam rekomendasi produk ini karena data yang tersedia hanya berfokus pada fitur pengguna. Pendekatan ini memanfaatkan data interaksi pengguna-produk untuk mengungkap pola dan kesamaan di antara para pengguna. Representasi graf digunakan untuk memodelkan hubungan interaksi pengguna-produk, yang memungkinkan pemodelan yang lebih komprehensif dari ketergantungan dan hubungan antara pengguna dan produk. Namun, dalam hal menyarankan produk yang mungkin diminati pengguna, beberapa metode seperti *Graph-based Neighborhood Methods* dan *Graph Laplacian-based* memiliki keterbatasan. Metode-metode ini cenderung mengandalkan pencarian produk serupa berdasarkan hubungan langsung mereka dengan produk lain, dan sering mengabaikan pola yang lebih rumit dan tren yang lebih besar di antara produk. Sehingga, rekomendasi yang dihasilkan dapat terasa generik dan gagal menangkap preferensi pengguna yang lebih spesifik dan kualitas unik dari setiap produk. Penelitian ini mengusulkan penggunaan GCN dalam kombinasi dengan FM untuk mengatasi keterbatasan-keterbatasan tersebut. GCN memanfaatkan struktur graf dan hubungan ketetanggaan antara pengguna dan produk, sehingga memungkinkan pemodelan yang lebih komprehensif dari ketergantungan dan hubungan antara pengguna dan produk. FM, di sisi lain, mempertimbangkan interaksi antara fitur pengguna dan fitur produk untuk memahami preferensi pengguna secara lebih mendalam. Hasil penelitian menunjukkan bahwa pendekatan FM-GCN memberikan performa yang lebih baik dalam evaluasi MAE pada data uji coba Cjamar group dengan 0.221 dan Amazon dengan 0.283. Hal ini menunjukkan bahwa metode FM-GCN lebih efektif dalam memberikan rekomendasi yang lebih akurat dan relevan dibandingkan dengan metode lain yang dievaluasi.

PENDAHULUAN

Rekomendasi produk memiliki peran yang signifikan bagi bisnis di berbagai industri, termasuk *e-commerce* [1], ritel [2], perhotelan [3], dan keuangan [4], karena dapat meningkatkan kepuasan dan penjualan pelanggan. Dalam industri *e-commerce*, rekomendasi produk dapat membantu pelanggan menemukan produk yang relevan, seperti ketika seorang pelanggan sedang melihat produk sepatu olahraga, rekomendasi dapat menyarankan produk-produk terkait seperti kaos kaki olahraga, alas kaki tambahan, atau aksesoris olahraga lainnya

[5]. Dalam industri ritel, rekomendasi produk dapat digunakan untuk meningkatkan penjualan melalui peningkatan *up-selling* dan *cross-selling*. Misalnya, di sebuah toko pakaian, saat seorang pelanggan memilih sebuah celana jeans, rekomendasi dapat menyarankan atasan yang cocok atau aksesoris yang sesuai dengan gaya dan warna celana tersebut [6]. Dalam industri perhotelan, rekomendasi produk dapat membantu meningkatkan pengalaman tamu dan mempersonalisasi layanan. Misalnya, ketika seorang tamu hotel memesan kamar, rekomendasi dapat menyarankan paket peningkatan kamar, layanan tambahan seperti spa atau restoran hotel, atau aktivitas lokal yang populer [7]. Dalam industri keuangan, rekomendasi produk dapat membantu pelanggan dalam mengelola keuangan mereka dengan lebih efektif. Misalnya, dalam aplikasi perbankan, rekomendasi dapat menyarankan produk dan layanan seperti investasi, asuransi, atau rencana pensiun berdasarkan profil risiko dan tujuan keuangan pelanggan [8].

Terdapat tiga pendekatan filtering untuk rekomendasi produk, yaitu *collaborative filtering* [9], *content based filtering* dan *hybrid filtering*. *Collaborative filtering* digunakan ketika data yang tersedia hanya fokus pada fitur pengguna dan tidak memiliki deskripsi produk yang eksplisit. Pendekatan ini memanfaatkan data interaksi pengguna-produk, bahkan tanpa adanya deskripsi produk yang terperinci, untuk mengungkap pola dan kesamaan di antara para pengguna [10]. Di sisi lain, *content-based filtering* sangat bergantung pada deskripsi produk yang komprehensif untuk mengekstrak fitur-fitur yang berarti dan memahami faktor-faktor produk yang unik. Tanpa deskripsi tersebut, *content-based filtering* menghadapi tantangan dalam memberikan rekomendasi yang akurat [11]. Meskipun pendekatan *hybrid* menggabungkan *collaborative filtering* dan *content-based filtering*, tanpa akses ke atribut produk yang mendetail, komponen *content-based filtering* mungkin kesulitan untuk mengekstrak fitur yang relevan, sehingga membatasi kemampuannya untuk meningkatkan kualitas rekomendasi secara keseluruhan [12].

Data memerlukan pemodelan hubungan interaksi pengguna-produk, oleh karena itu digunakan representasi berbasis graf, di mana pengguna dan produk adalah node, dan interaksi di antara keduanya direpresentasikan sebagai edge. Representasi graf ini memungkinkan pemodelan yang lebih komprehensif dari ketergantungan dan hubungan antara pengguna dan produk [13]. Sebaliknya, representasi lainnya, seperti metode berbasis *naïve bayes* [14], *deep learning* [15], dan berbasis klasifikasi [16], mungkin tidak secara eksplisit memodelkan hubungan antara pengguna-produk. Metode-metode ini biasanya bergantung pada ukuran kemiripan untuk membuat rekomendasi, yang mungkin mengabaikan keterkaitan dan ketergantungan yang ada di dalam data [17].

Setelah membangun representasi graf, berbagai pendekatan dapat diterapkan, termasuk *Graph Convolutional Network*, *Graph-based Neighborhood Methods*, dan *Graph Laplacian-based*. Metode berbasis *Graph Laplacian* dalam rekomendasi produk menggunakan analisis spektral graf untuk memodelkan hubungan antar produk. Dibangun representasi graf di mana produk berhubungan dengan node dan edge mewakili koneksi, matriks *Graph Laplacian* diturunkan untuk menangkap hubungan berpasangan. Sifat-sifat spektral dari matriks tersebut kemudian dianalisis untuk mengidentifikasi kelompok dan kemiripan di antara produk [18]. *Graph-based Neighborhood Methods* dalam rekomendasi produk melibatkan analisis hubungan antar produk dalam struktur graf. Dalam metode ini, setiap produk direpresentasikan sebagai sebuah node, dan sisi-sisi antar node mewakili koneksi atau hubungan. Memanfaatkan pertimbangan produk tetangga dari produk tertentu, metode ini bertujuan untuk mengidentifikasi produk yang serupa atau terkait [19]. GCN menerapkan konvolusi graf untuk menyebarkan dan memperbarui penyematanan node berdasarkan hubungan ketetanggaan mereka. Dalam konteks rekomendasi produk, setiap produk direpresentasikan sebagai node dalam graf, dan sisi-sisinya menunjukkan hubungan atau kesamaan antar produk. Secara iteratif mengumpulkan dan memperbarui embedding

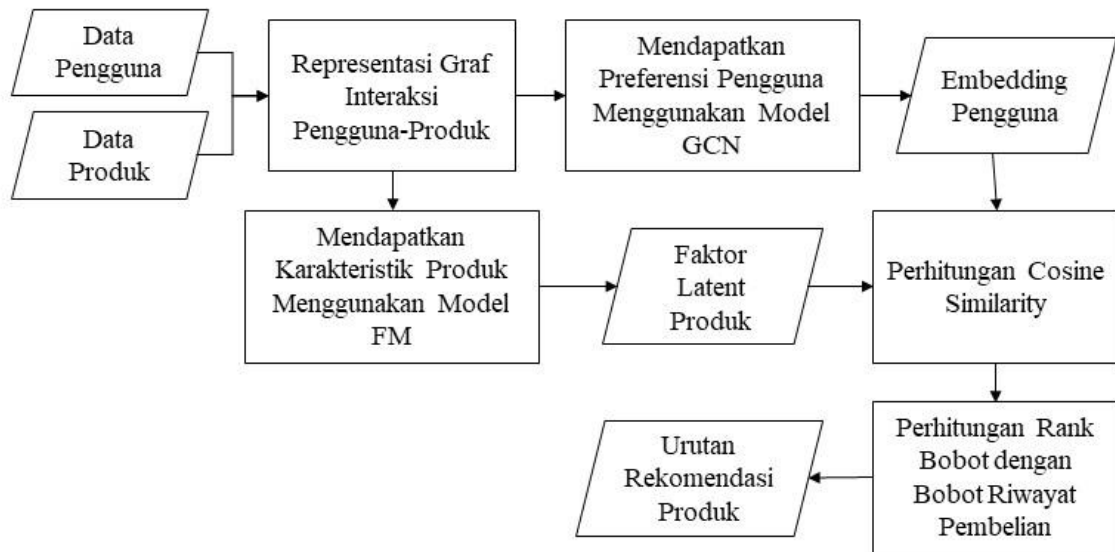
node menggunakan konvolusi graf, GCN menangkap ketergantungan tingkat tinggi dan pola global dalam graf. Hal ini memungkinkan GCN untuk menemukan fitur produk yang tersembunyi dan mengidentifikasi hubungan yang sangat penting untuk membuat rekomendasi yang akurat. GCN memanfaatkan informasi lingkungan sekitar dan struktur graf yang lebih luas, untuk meningkatkan pemahaman tentang preferensi pengguna dan menghasilkan rekomendasi yang dipersonalisasi yang mencerminkan interaksi yang kompleks antar produk, yang pada akhirnya meningkatkan pengalaman pengguna secara keseluruhan dalam sistem rekomendasi produk [20].

Dalam hal menyarankan produk yang mungkin diminati pengguna, beberapa metode mengandalkan pencarian produk serupa berdasarkan hubungan mereka dengan produk lain. Namun, metode-metode seperti *Graph-based Neighborhood Methods* [21], dan *Graph Laplacian-based* [22] memiliki keterbatasan. Salah satu keterbatasannya adalah mereka sering mengabaikan pola yang lebih rumit dan tren yang lebih besar di antara produk. Mereka mungkin hanya mengelompokkan produk berdasarkan hubungan langsung tanpa mempertimbangkan preferensi pengguna yang lebih spesifik dan kualitas unik dari setiap produk. Hal ini dapat menghasilkan rekomendasi yang terasa generik dan gagal menangkap beragam produk yang sebenarnya mungkin dinikmati pengguna.

Pada penelitian ini, diusulkan penggunaan Graph Convolutional Network (GCN) dalam kombinasi dengan Factorization Machine (FM) untuk meningkatkan personalisasi dan akurasi rekomendasi produk. Pendekatan FM-GCN memanfaatkan struktur graf dan hubungan ketetanggaan antara pengguna dan produk, sehingga memungkinkan pemodelan yang lebih komprehensif dari ketergantungan dan hubungan antara keduanya. GCN memperhitungkan hubungan antar produk dan fitur unik dari setiap produk. Dalam prosesnya, GCN dapat mengidentifikasi hubungan yang lebih rinci dan menangkap pola kompleks yang penting untuk rekomendasi. GCN melihat lebih dari sekadar hubungan langsung dan mempertimbangkan jaringan produk yang lebih besar, sehingga dapat mendeteksi tren dan kesamaan yang mungkin terlewatkan oleh metode lain [23]. Sementara itu, FM mempertimbangkan interaksi antara fitur pengguna dan fitur produk, sehingga memahami preferensi pengguna secara lebih mendalam. FM mampu menangkap interaksi kompleks antara fitur-fitur ini, bahkan tanpa adanya deskripsi produk yang eksplisit [24]. Secara keseluruhan, dengan memanfaatkan kekuatan GCN dan FM secara bersama-sama, sistem rekomendasi produk dapat memberikan pengalaman pengguna yang lebih menarik dan menyenangkan. Pengguna akan dapat menemukan produk baru dan relevan yang benar-benar sesuai dengan selera mereka, meningkatkan kepuasan dan keterlibatan mereka dengan platform rekomendasi.

METODE

Pada penelitian ini dilakukan pembentukan model rekomendasi seperti Gambar 1, dimana digunakan data pengguna dan data produk sebagai input, dimana kedua data akan direpresentasikan sebagai bentuk graf seperti pada Gambar 2. Node pada graf memrepresentasikan pengguna (P_n) dan produk (I_n), hubungan pengguna dengan produk yang dibeli direpresentasikan sebagai edge, dimana bobot edge merupakan frekuensi pembelian produk. Representasi graf melalui tahap pemodelan GCN, dilakukan iterasi untuk menggabungkan dan memperbarui informasi node dengan mempertimbangkan informasi node yang bersebelahan dan fitur node itu sendiri, untuk mendapatkan preferensi pengguna. Hasil pemodelan GCN disimpan dalam bentuk representasi vektor embedding pada setiap pengguna, seperti pada Tabel 1.



Gambar 1. Model rekomendasi Graph Convolutional Network

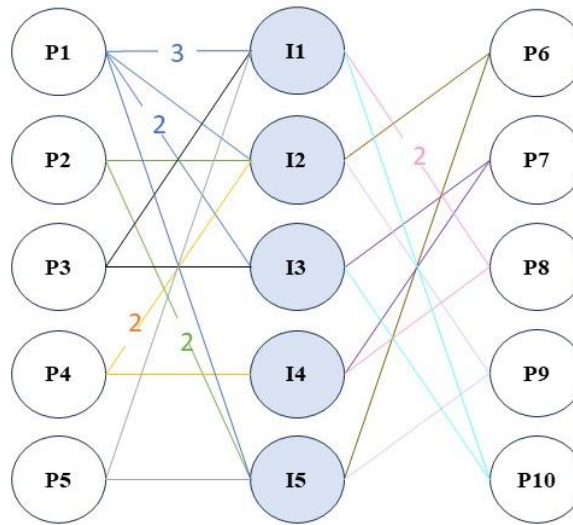
Selanjutnya, untuk melengkapi pemodelan preferensi pengguna, dilakukan perhitungan factor latent menggunakan model FM pada data pengguna dan produk yang direpresentasikan dalam bentuk matriks pengguna-produk. Model FM memanfaatkan teknik SVD (Singular Value Decomposition) [25] untuk menguraikan matriks pengguna-produk menjadi faktor-faktor latent yang merepresentasikan preferensi pengguna terhadap produk. Sebagai contoh, dalam Tabel 1, dilakukan perhitungan FM untuk mendapatkan faktor latent untuk setiap produk (I1, I2, I3, I4, I5). Dalam contoh ini, digunakan SVD dengan parameter $k=2$, yang menghasilkan dua faktor latent untuk setiap produk.

Setelah mendapatkan faktor latent, langkah selanjutnya adalah menghitung kesamaan antara vektor embedding pengguna (hasil dari GCN) dengan faktor latent setiap produk. Ini dapat dilakukan dengan menggunakan metode perhitungan kesamaan *cosine similarity*. Setelah menghitung kesamaan antara vektor embedding pengguna dan faktor latent setiap produk, tahapan selanjutnya adalah menerapkan penilaian berbobot (*weighted ranking*) untuk memperhitungkan faktor penting dari setiap produk dalam rekomendasi. *Weighted ranking* memungkinkan penulis memberikan bobot yang lebih besar pada faktor-faktor yang dianggap lebih signifikan atau penting. Pada penelitian ini faktor tersebut adalah frekuensi pembelian produk yang direpresentasi sebagai weight pada edge hubungan pengguna-produk.

Setelah mendapatkan nilai prediksi, MAE dapat dihitung dengan membandingkan nilai prediksi dengan nilai sebenarnya dalam dataset evaluasi [26]. MAE memberikan gambaran tentang kesalahan rata-rata antara prediksi model dan nilai sebenarnya dari pengguna. Semakin rendah nilai MAE, semakin baik kinerja model rekomendasi. Penggunaan MAE sebagai metrik evaluasi, penulis dapat mengukur kualitas dan akurasi dari sistem rekomendasi berbasis GCN. Tujuan penulis adalah untuk mengoptimalkan model sehingga MAE semakin mendekati nol, yang menunjukkan tingkat kesesuaian yang tinggi antara prediksi dan nilai sebenarnya. Berdasarkan hasil MAE, penulis dapat melakukan pembaruan dan perbaikan pada model GCN untuk meningkatkan kinerja dan kualitas rekomendasi yang dihasilkan.

Tabel 1. Vektor Embedding Pengguna dan Faktor Latent Produk

Id	Vektor Embedding Pengguna	Id	Faktor Latent
P1	[0.0832, 0.1141, 0.0807, 0.0661, 0.1006, 0.0837, 0.0483, 0.0589]	I1	[0.25, -0.37]
P2	[0.0197, 0.1467, 0.0306, 0.0843, 0.0527, 0.0241, 0.1095, 0.0003]	I2	[-0.37, -0.19]
P3	[0.0633, 0.0137, 0.0665, 0.0366, 0.0735, 0.0804, 0.0664, 0.0108]	I3	[0.02, -0.22]
P4	[0.0061, 0.0929, 0.0197, 0.1393, 0.0253, 0.0347, 0.0127, 0.0412]	I4	[-0.56, 0.45]
P5	[0.0621, 0.0364, 0.0584, 0.0256, 0.0676, 0.0744, 0.0622, 0.0168]	I5	[0.68, 0.74]
P6	[0.0515, 0.1059, 0.0476, 0.1002, 0.0455, 0.0534, 0.0253, 0.0738]		
P7	[0.0561, 0.0675, 0.0845, 0.0707, 0.0195, 0.0357, 0.0129, 0.0989]		
P8	[0.0735, 0.0202, 0.0408, 0.1447, 0.0253, 0.0386, 0.0218, 0.0637]		
P9	[0.0361, 0.0943, 0.0407, 0.0762, 0.0431, 0.0503, 0.0377, 0.0224]		
P10	[0.0629, 0.0175, 0.0632, 0.0354, 0.0734, 0.0802, 0.0662, 0.0106]		



Gambar 2. Representasi Graf Interaksi Pengguna-Produk

Skenario 1

Dimisalkan penulis ingin merekomendasikan produk pada pengguna dengan id P1 yang telah beberapa kali melakukan transaksi.

- Pertama, dilakukan perhitungan preferensi P1 menggunakan model GCN yang menghasilkan vector embedding P1 yaitu [0.0832, 0.1141, 0.0807, 0.0661, 0.1006, 0.0837, 0.0483, 0.0589].
- Selanjutnya, digunakan model FM untuk memahami karakteristik produk yang berbeda. Model ini membantu melihat fitur atau kualitas apa yang membuat suatu produk menarik bagi seseorang.
- Setelah mendapatkan gambaran tentang preferensi P1 dan karakteristik produk, kami membandingkannya menggunakan *cosine similarity* untuk membantu memahami seberapa mirip preferensi P1 dengan produk yang berbeda. Hasil dari perhitungan *cosine similarity* pada P1 dan setiap produk adalah sebagai berikut.

$$\begin{aligned}
 \text{Cosine similarity}(P1, I1) &= 0.79; & \text{Cosine similarity}(P1, I2) &= 0.62; \\
 \text{Cosine similarity}(P1, I3) &= 0.55; & \text{Cosine similarity}(P1, I4) &= 0.42; \\
 \text{Cosine similarity}(P1, I5) &= 0.91;
 \end{aligned}$$

- Dilakukan *weighted ranking* yang memperhitungkan skor *cosine similarity* dan riwayat frekuensi pembelian menggunakan persamaan 1.

$$\text{weighted ranking} = (0.7 * \text{cosine similarity}) + (0.3 * \text{frekuensi pembelian}) \quad (1)$$

- Hasil dari perhitungan *weighted ranking* selanjutnya diurutkan dari yang terbesar ke terkecil.
 - I1 - *Weighted ranking*: 1.46
 - I3 - *Weighted ranking*: 0.95
 - I2 - *Weighted ranking*: 0.72
 - I5 - *Weighted ranking*: 0.58
 - I4 - *Weighted ranking*: 0.49
- Berdasarkan hasil perhitungan, maka produk yang paling direkomendasikan kepada P1 adalah I1.

Skenario 2

Dimisalkan penulis ingin merekomendasikan produk pada pengguna dengan id P11 yang belum pernah melakukan transaksi.

- Pertama, dilakukan penghitungan preferensi P11 menggunakan model GCN. Model ini melihat produk yang dibeli orang lain dan mencoba memahami apa yang mungkin disukai P11. Dihasilkan vector embedding P11 sebagai [0.0154, 0.0218, 0.0396, 0.0112, 0.0273, 0.0345, 0.0251, 0.0187].
- Selanjutnya, digunakan model FM untuk memahami karakteristik produk yang berbeda.
- Setelah mendapatkan gambaran tentang preferensi P11 dan karakteristik produknya, dilakukan perbandingan menggunakan *cosine similarity*. Ini membantu kami memahami seberapa mirip preferensi P11 dengan produk yang berbeda.

Cosine similarity with I1: 0.4915;

Cosine similarity with I3: 0.0531;

Cosine similarity with I5: 0.5933

Cosine similarity with I2: 0.2528;

Cosine similarity with I4: -0.1121;

- Berdasarkan skor *cosine similarity* (dikarenakan pengguna belum memiliki riwayat frekuensi pembelian, maka tidak dilakukan perhitungan *weighted rank*) ini, kami memberi peringkat produk dan merekomendasikannya ke P11. Produk yang lebih mirip dengan preferensi P11 diberi peringkat lebih tinggi. Jadi, produk yang direkomendasikan untuk P11 adalah I5, I1, I2, I3, dan I4, dalam urutan tersebut.

Skenario 3

Dimisalkan penulis ingin merekomendasikan produk pada pengguna dengan id P10 yang telah beberapa kali melakukan transaksi.

- Pertama, dilakukan perhitungan preferensi P10 menggunakan model GCN yang menghasilkan vector embedding P10 yaitu [0.0629, 0.0175, 0.0632, 0.0354, 0.0734, 0.0802, 0.0662, 0.0106].
- Selanjutnya, digunakan model FM untuk memahami karakteristik produk yang berbeda. Model ini membantu melihat fitur atau kualitas apa yang membuat suatu produk menarik bagi seseorang.
- Setelah mendapatkan gambaran tentang preferensi P10 dan karakteristik produk, kami membandingkannya menggunakan *cosine similarity* untuk membantu memahami seberapa mirip preferensi P10 dengan produk yang berbeda. Hasil dari perhitungan *cosine similarity* pada P10 dan setiap produk adalah sebagai berikut.

Cosine similarity(P10, I1) = 0.83;

Cosine similarity(P10, I3) = 0.77;

Cosine similarity(P10, I5) = 0.69

Cosine similarity(P10, I2) = 0.50;

Cosine similarity(P10, I4) = 0.62;

- Berdasarkan hasil perhitungan (tidak dilakukan perhitungan *weight ranking* karena pengguna hanya pernah bertransaksi sekali pada kedua produk), maka produk yang paling direkomendasikan kepada P1 adalah I1, I3, I5, I4, dan I2, dalam urutan tersebut.

Skenario 4

Dimisalkan penulis ingin merekomendasikan produk pada pengguna baru dengan id P12, dimana P12 pernah bertransaksi dengan produk I3 sebanyak 1 kali, dan ingin direkomendasikan produk yang belum pernah dibeli.

- Pertama, dilakukan perhitungan preferensi P12 menggunakan model GCN yang menghasilkan vector embedding P12 yaitu [0.0231, 0.0356, 0.0872, 0.0143, 0.0415, 0.0726, 0.0508, 0.0219].
- Selanjutnya, digunakan model FM untuk memahami karakteristik produk yang berbeda. Model ini membantu melihat fitur atau kualitas apa yang membuat suatu produk menarik bagi seseorang.
- Setelah mendapatkan gambaran tentang preferensi P12 dan karakteristik produk, kami membandingkannya menggunakan *cosine similarity* untuk membantu memahami seberapa mirip preferensi P12 dengan produk yang berbeda. Hasil dari perhitungan *cosine similarity* pada P12 dan setiap produk adalah sebagai berikut.

$$\begin{aligned} \text{Cosine similarity}(P12, I1) &= 0.0181; & \text{Cosine similarity}(P12, I2) &= 0.0945; \\ \text{Cosine similarity}(P12, I3) &= 0.9811; & \text{Cosine similarity}(P12, I4) &= 0.1333; \\ \text{Cosine similarity}(P12, I5) &= 0.2007 \end{aligned}$$

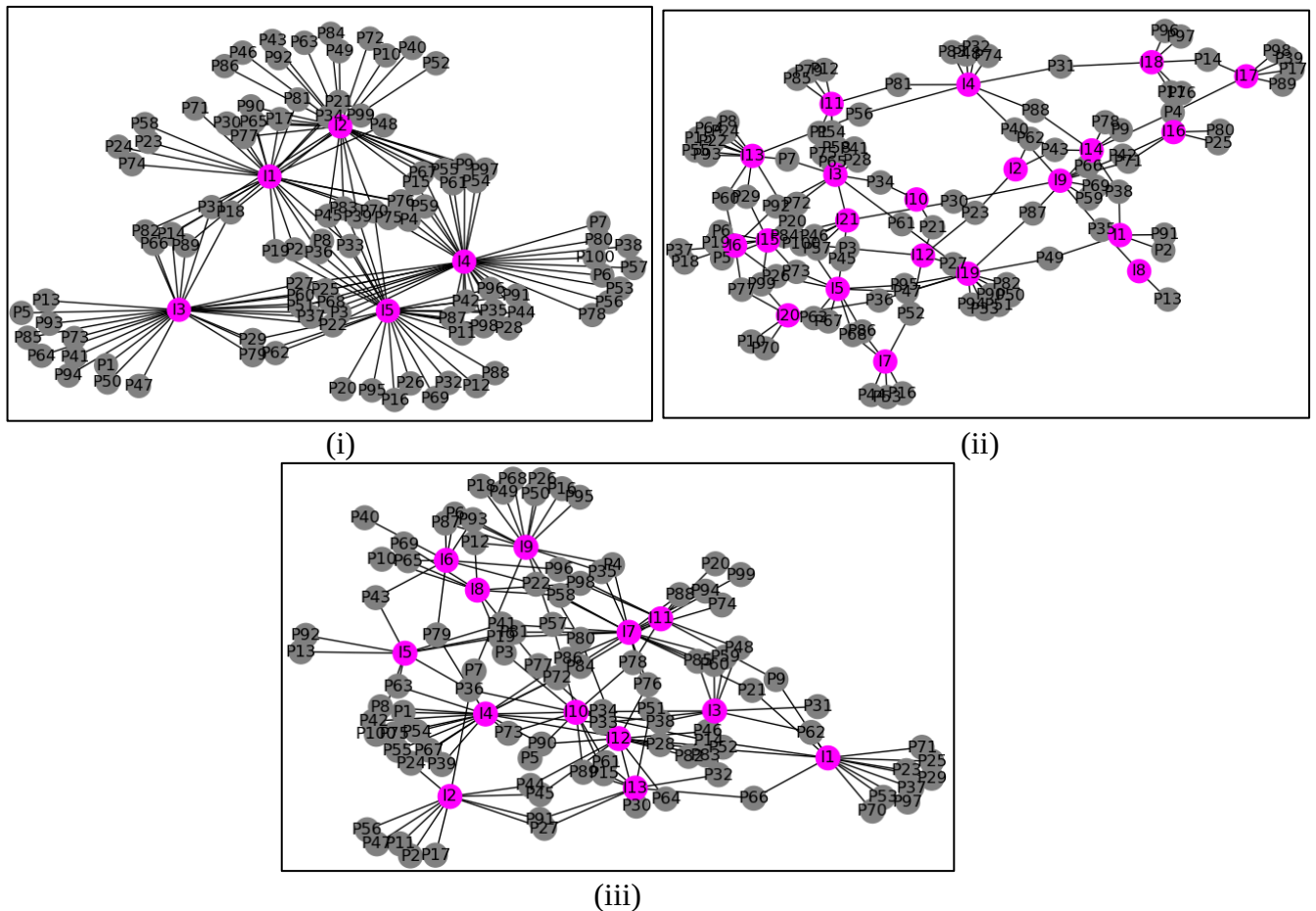
- Karena P12 sudah membeli I3, penulis tidak merekomendasikan produk tersebut. Sekarang penulis dapat mengurutkan produk-produk yang tersisa berdasarkan skor kemiripan kosinus. Produk yang direkomendasikan untuk P12, berdasarkan kemiripan yang terbesar, adalah I5, I4, I2, dan I1.

HASIL DAN PEMBAHASAN

Pada penelitian ini dilakukan uji coba pada 3 data seperti pada Gambar 3, yaitu data cajamar group, data amazon, dan data retailrocket. Data amazon terdiri dari 25.000 data pengguna dengan atribut id pengguna dan rating pengguna terhadap suatu produk dengan rentang 1-5, sedangkan data produk berfokus pada 25 produk kecantikan dengan atribut id produk dan kategori produk. Data retailrocket terdiri dari 212.182 data pengguna dengan atribut id pengguna, jenis kelamin, dan frekuensi bertransaksi, sedangkan data produk 31.968 memiliki atribut id produk, nama produk, dan kategori produk. Data cajamar group memiliki 676.340 data dan atribut pengguna berupa id pengguna, umur, pekerjaan, pendapatan, frekuensi bertransaksi, dan kepemilikan produk, sedangkan data produk memiliki 94 data dengan atribut id produk, nama produk, dan kategori produk. Ketiga data tersebut dilakukan uji coba model rekomendasi GCN dimana hasil evaluasi MAE bervariasi dengan nilai laten (k) seperti pada Tabel 2. Pada representasi graf, input node merupakan node pengguna (berdasarkan id pengguna) dan node produk (berdasarkan id produk), hubungan pengguna-produk merupakan riwayat pembelian produk oleh pengguna.

Pada hasil percobaan, nilai laten mempengaruhi hasil MAE. Semakin tinggi nilai laten maka hasil MAE semakin kurang bagus. Hal ini dikarenakan, dengan jumlah faktor laten yang lebih tinggi, model menjadi lebih kompleks dan membutuhkan lebih banyak parameter yang harus diestimasi. Kompleksitas yang meningkat ini dapat membuat model lebih rentan terhadap overfitting dan dapat menghambat kemampuannya untuk secara akurat menangkap pola dan hubungan yang mendasar dalam data. Faktor laten yang lebih tinggi dapat

memperbesar noise yang ada dalam data. Model mungkin mulai menangkap variasi yang tidak relevan dan noise daripada pola yang mendasarinya, sehingga menyebabkan akurasi prediksi yang lebih buruk dan MAE yang lebih tinggi.



Gambar 3. Representasi Graf (i) Data Cajamar (ii) Data Amazon (iii) Data Retailrocket

Tabel 2. Hasil evaluasi MAE pada ketiga data uji coba.

Nilai Laten (k)	Hasil MAE		
	Cajamar Group	Amazon	Retailrocket
2	0.2586	0.3142	0.2136
5	0.3721	0.3821	0.2708
10	0.3894	0.3981	0.3753
13	0.4183	0.4376	0.4198
15	0.5285	0.4931	0.4672
20	0.5729	0.5327	0.5521
23	0.5891	0.5781	0.5913
25	0.6032	0.5938	0.6382
30	0.6259	0.6194	0.6573

Pada peneitian ini, penulis membandingkan metode usulan, yaitu FM GCN, dengan metode pembanding Node2Vec dan GraphSAGE seperti pada Tabel 3. Dalam melakukan perbandingan ini, penulis memilih untuk melakukan variasi pada parameter *embedding dimension*. Node2Vec menggunakan teknik *random walk* untuk menjelajahi graf dan menghasilkan representasi embedding berdasarkan urutan langkah-langkah yang diambil [27]. Di sisi lain, FM GCN menggabungkan metode faktorisasi matriks dengan GCN untuk

memperoleh representasi embedding yang lebih kaya dengan memperhitungkan struktur dan atribut graf secara bersamaan. GraphSAGE adalah metode yang memperluas GCN dengan menggunakan pengambilan sampel tetangga yang adaptif. Metode ini memungkinkan penggabungan informasi dari berbagai jarak dan tingkat dalam graf [28]. Dalam membandingkan GraphSAGE dengan FM GCN, penulis ingin mengevaluasi apakah pengambilan sampel tetangga yang adaptif tersebut memberikan keunggulan dalam mendapatkan representasi embedding yang lebih baik dibandingkan dengan pendekatan yang dikombinasikan dengan faktorisasi matriks.

Berdasarkan hasil penelitian, dapat disimpulkan bahwa terdapat hubungan yang kompleks antara nilai *embedding dimension* dan performa model dalam memprediksi hasil MAE. Peningkatan nilai *embedding dimension* yang terlalu besar cenderung menghasilkan performa yang kurang, hal ini disebabkan semakin besar nilai *embedding dimension*, semakin banyak dimensi yang perlu dipelajari oleh model. Ketika dimensi menjadi sangat besar, jarak antara titik-titik data dalam ruang embedding dapat menjadi semakin tidak berarti. Ini dapat menyebabkan penyebaran data yang tidak merata atau kehilangan struktur asli dalam data. Akibatnya, performa model dapat menurun karena sulit untuk mempelajari pola-pola yang relevan dalam dimensi yang sangat tinggi.

KESIMPULAN

Pada penelitian ini digunakan metode FM GCN menggabungkan keunggulan dari *Factorization machine* (FM) dan Graph Convolutional Networks (GCN) untuk memberikan rekomendasi berdasarkan *collaborative filtering* dan pembelajaran representasi berbasis graf. GCN memanfaatkan interaksi antara pengguna dan produk, serta hubungan dalam graf produk-produk, model FM GCN menangkap pola-pola dan preferensi yang mendasar untuk memberikan rekomendasi yang personal. Dalam metode ini, komponen FM menangkap interaksi berpasangan antara pengguna dan produk, memungkinkan model untuk mempelajari preferensi individu dan interaksi di antara mereka. Komponen GCN memanfaatkan struktur graf produk-produk untuk menyebarkan informasi dan menangkap pengaruh serta kesamaan antara produk-produk. Menggunakan integrasi FM dan GCN, model FM GCN dapat menghasilkan rekomendasi dengan mempertimbangkan preferensi individu pengguna dan aspek *collaborative filtering*. Model ini memanfaatkan faktor laten yang dipelajari dan hubungan antara pengguna dan produk untuk membuat prediksi dan memberikan rekomendasi yang personal. Hasil evaluasi menunjukkan bahwa pendekatan FM-GCN memberikan performa yang lebih baik dalam evaluasi MAE pada dua data uji coba, yaitu Cajamar group dengan nilai 0.221 dan Amazon dengan nilai 0.283. Hasil ini menunjukkan bahwa metode FM-GCN lebih efektif dalam memberikan rekomendasi yang lebih akurat dan relevan dibandingkan dengan metode pembandingan yang dievaluasi.

Tabel 3. Hasil Evaluasi MAE pada Metode FM GCN, Node2Vec, dan GraphSAGE

Metode	Data Uji Coba	Embedding Dimension	Hasil MAE
FM GCN	Cajamar Group	10	0.221
		20	0.258
		50	0.460
Metode	Data Uji Coba	Embedding Dimension	Hasil MAE
FM GCN	Amazon	10	0.283
		20	0.315

	Retailrocket	50	0.352
		10	0.363
		20	0.439
		50	0.554
Node2Vec	Cajamar Group	10	0.448
		20	0.480
		50	0.504
		10	0.418
	Amazon	20	0.772
		50	0.784
		10	0.352
		20	0.809
	Retailrocket	50	0.678
		10	0.608
		20	0.341
		50	0.636
GraphSAGE	Cajamar Group	10	0.235
		20	0.374
		50	0.388
	Amazon	10	0.415
		20	0.577
		50	0.548
	Retailrocket	10	0.415
		20	0.577
		50	0.548
		10	0.415
		20	0.577
		50	0.548

REFERENSI

- [1] L. Jiang, "A trust-based collaborative filtering algorithm for E-commerce recommendation system," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, pp. 3023-3034, 2019.
- [2] H. Hwangbo, Y. Kim and K. Cha, "Recommendation system development for fashion retail e-commerce," *Electronic Commerce Research and Applications*, vol. 28, pp. 94-101, 2018.
- [3] M. Nilashi, A. Ahani, M. D. Esfahani, E. Yadegaridehkordi, S. Samad, O. Ibrahim, N. M. Sharef and E. Akbari, "Preference learning for eco-friendly hotels recommendation A multi-criteria collaborative filtering approach," *Journal of Cleaner Production*, vol. 215, pp. 767-783, 2019.
- [4] E. Hernandez-Nieves, G. Hernandez, A. Gil-Gonzalez, S. Rodriguez-Gonzalez and J. M. Corchado, "Fog computing architecture for personalized recommendation of banking products," *Expert Systems with Applications*, vol. 140, 2020.
- [5] M. Asep and Y. Siti, "APLIKASI E-COMMERCE DENGAN SISTEM REKOMENDASI BERBASIS COLLABORATIVE FILTERING," *JURTIK : JURNAL TEKNOLOGI INFORMASI DAN KOMUNIKASI*, vol. 7, no. 2, pp. 1-15, 2018.
- [6] P. Bayu, B. Indra and Y. Arlisa, "Product Recommendation in Offline Retail Industry by using Collaborative Filtering," *(IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 11, no. 9, pp. 635-643, 2020.
- [7] K. Tugba and K. Cihan, "A novel top-n recommendation method for multi-criteria collaborative filtering," *Expert Systems with Applications*, vol. 198, pp. 1-12, 2022.
- [8] B. Baptise and C. Laurent, "History-Augmented Collaborative Filtering for Financial Recommendations," in *RecSys '20: Proceedings of the 14th ACM Conference on Recommender Systems*, Brazil, 2020.
- [9] A. Iftikhar, M. A. Ghazanfar, M. Ayub, Z. Mehmood and M. Maqsood, "An Improved Product Recommendation Method for Collaborative Filtering," *IEEE Access*, vol. 8, pp. 123841-123857, 2020.
- [10] Z. Shahbazi, D. Hazra, S. Park and Y. C. Byun, "Toward improving the prediction accuracy of product recommendation system using extreme gradient boosting and encoding approaches," *Symmetry*, vol. 12, no. 9, pp. 1-20, 2020.
- [11] J. Son and S. B. Kim, "Content-based filtering for recommendation systems using multiattribute networks," *Expert Systems with Applications*, vol. 89, pp. 404-412, 2017.

- [12] S. Ismail, A. Abdul-Barik and S. Abdul-Mumin, "An enhanced item recommendation approach using the sigmoid function and jaccard similarity coefficient," *Journal of Mathematical Sciences & Computational Mathematic*, vol. 4, no. 3, pp. 98-120, 2023.
- [13] J. Sun, Y. Zhang, C. Ma, M. Coates, H. Guo, R. Tang and X. He, "Multi-graph Convolution Collaborative Filtering," in *IEEE International Conference on Data Mining*, Beijing, China, 2019.
- [14] H. Maariful and Y. Rizqi, "Kombinasi Naive Bayes dan Metode Time Series Sebagai Peramalan Pergerakan Harga pada Perdagangan Valuta Asing," *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 2, no. 2, pp. 151-155, 2020.
- [15] M. C. Untoro, R. Leonard, A. Perdana, N. Wijaya and N. Ferdiyanto, "Penerapan K-Means Clustering pada Imbalance Dataset Gejala Penyakit Jantung," *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 5, no. 1, pp. 1-7, 2023.
- [16] P. Edy and F. Charis, "Implementasi Algoritma Naive Bayes Classifier Untuk Penentuan Kelayakan Penerima Beasiswa Peningkatan Prestasi Akademik (PPA) Bagi Mahasiswa," *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 3, no. 2, pp. 165-177, 2021.
- [17] L. Yang, H. He, H. Lan and Y. Shi, "Personalized Recommendation Algorithm of E-commerce Based on Cloud Computing," in *Proceedings of the 2017 7th International Conference on Applied Science, Engineering and Technology*, Qingdao, China, 2017.
- [18] F. Fous, A. Pirotte and M. Saerens, "Random-Walk Computation of Similarities between Nodes of a Graph with Application to Collaborative Recommendation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, no. 3, pp. 355-369, 2007.
- [19] K. Reddy, K. Masaru, P. reekanth and S. Srinivasa, "A Graph Based Approach to Extract a Neighborhood Customer Community for Collaborative Filtering," in *Databases in Networked Information Systems*, Berlin, Heidelberg, 2002.
- [20] J. Zhang, X. Shi, S. Zhao and I. King, "GCN: Stacked and Reconstructed Graph Convolutional Networks for Recommender Systems," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, Macao, China, 2019.
- [21] S. Shaikh, S. Rathi and P. Janrao, "Recommendation System in E-Commerce Websites: A Graph Based Approach," in *IEEE 7th International Advance Computing Conference (IACC)*, Hyderabad, India, 2017.
- [22] Z. Yifei and H. Conor, "Graph-Based Diffusion Method for Top-N Recommendation," in *AICS 2022: Artificial Intelligence and Cognitive Science*, Munster, Ireland, 2022.
- [23] X. Fengli, L. Jianxun, H. Zhenyu and L. Yong, "Relation-Aware Graph Convolutional Networks for Agent-Initiated Social E-Commerce Recommendation," in *CIKM '19: Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, Beijing, China, 2019.
- [24] X. Junbo, "E-Commerce Product Recommendation Method Based on Collaborative Filtering Technology," in *2016 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, Chicago, 2016.
- [25] S. Akansha, G. S. Reddy and C. Kumar, "User Product Recommendation System Using KNN-Means and Singular Value Decomposition," in *2022 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON)*, Bengaluru, India, 2022.
- [26] T. Duy and H. Jun-Ho, "New machine learning model based on the time factor for e-commerce recommendation systems," *The Journal of Supercomputing*, vol. 79, pp. 6756-6801, 2023.
- [27] J. Chen, "N2VSCDNNR: A Local Recommender System Based on Node2vec and Rich Information Network," *IEEE Transactions on Computational Social Systems*, vol. 6, pp. 456-466, 2019.
- [28] A. Driss, R. Jamal, S. Abdelouahed and A. Badraddine, "Deep GraphSAGE-based recommendation system: jumping knowledge connections with ordinal aggregation network," *Neural Computing and Applications*, vol. 34, no. 14, pp. 11679-11690, 2022.