

Penerapan graph neural network dalam pembangunan sistem rekomendasi

Muhammad Gerald Akbar Giffera (13520143)
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10
Bandung 40132, Indonesia
133520143@std.stei.itb.ac.id

Abstrak—Sistem rekomendasi adalah hasil dari pengembangan ilmu Teknik informatika di bidang machine learning. Sistem ini mengolah data pengguna sedemikian rupa hingga akhirnya bisa memprediksi selera atau kesukaan pengguna berdasarkan layanan yang disediakan.

Bidang machine learning yang dimanfaatkan dalam pengembangan sistem rekomendasi adalah neural network. Neural network merupakan salah satu contoh dari pemanfaatan teori graf dalam ilmu Matematika Diskrit.

Kata kunci—sistem rekomendasi, graf, neural network, graph neural network

I. PENDAHULUAN

BIKIN JADI KAYAK ADA MASALAH

Seperti yang kita ketahui, perkembangan teknologi sekarang terjadi dengan sangat cepat, banyak keperluan kita sebagai manusia yang sebelumnya memerlukan waktu yang cukup banyak, sekarang bisa dipenuhi dengan memanfaatkan aplikasi atau situs tertentu. Hal ini pun sangat terlihat di masa pandemic covid-19 ini, karena terbatasnya akses mobilitas masyarakat sehingga layanan-layanan online menjadi pilihan utama bagi masyarakat di seluruh Indonesia maupun dunia. Layanan-layanan yang menjadi sangat krusial bagi masyarakat adalah *Online Shop* dan layanan *streaming*.

Jika diperhatikan layanan-layanan ini seringkali dapat merekomendasikan hal yang ingin kita beli/lihat, dan pada umumnya rekomendasi yang diberikan akan sesuai dengan keinginan kita. Sebagai contoh pada aplikasi *online shop*, di laman utama aplikasi akan muncul rekomendasi barang-barang yang diberikan oleh aplikasi tersebut, dan jika dibandingkan, antar user pasti mendapatkan rekomendasi yang berbeda-beda sesuai selera mereka masing-masing. Hal ini juga dapat dilihat pada layanan streaming seperti Youtube, Netflix, Disney Hotstar dan sebagainya, mereka bisa merekomendasikan tayangan-tayangan yang sesuai dengan selera kita. Lantas bagaimana sistem-sistem ini bekerja?

Layanan-layanan ini memanfaatkan cabang dari ilmu machine learning yaitu deep learning/neural network. Dengan memanfaatkan neural network, data yang diterima oleh layanan streamin atau online shop, bisa diolah sedemikian rupa sehingga

menghasilkan prediksi akan produk yang kemungkinan besar akan disukai oleh pengguna. Makalah ini akan membahas lebih lanjut mengenai aplikasi dari teori graf yang diajarkan di mata kuliah Matematika Diskrit kepada sistem rekomendasi yang didasarkan oleh neural network.

II. DASAR TEORI

A. Graf

Graf adalah suatu struktur yang digunakan untuk merepresentasikan objek diskrit dan hubungan antar objek-objek tersebut. Pada umumnya graf direpresentasikan sebagai *tuple* atau pasangan dua nilai yaitu:

$$G = (V, E)$$

V = Himpunan tidak kosong (vertices)

E = Himpunan sisi yang menghubungkan sepasang simpul

1) Jenis-jenis graf

Graf dapat diklasifikasi berdasarkan arah orientasi pada sisi graf dan juga berdasarkan ada atau tidaknya gelang pada sisi graf.

a) Graf sederhana

Graf sederhana adalah graf yang tidak mengandung sisi gelang pada sisi-sisinya.

TAMBAHIN GAMBAR

b) Graf tak sederhana

Graf tak sederhana adalah graf yang mempunyai sisi gelang pada sisi-sisinya, graf tak sederhana dibagi lagi menjadi 2 yaitu, graf ganda dan graf semu. Graf ganda adalah graf yang mempunyai sisi ganda, sedangkan graf semu (*pseudo-graph*), adalah graph yang mempunyai sisi gelang.

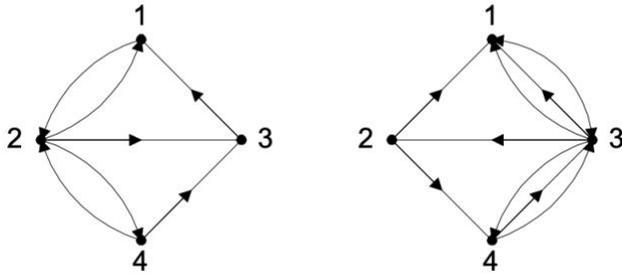
TAMBAHIN GAMBAR

c) Graf berarah

Graf berarah adalah graf yang memiliki orientasi arah pada sisi-sisinya, pada umumnya orientasi arah ditandai dengan adanya tanda panah pada sisi graf.

Sisi berarah pada graf disebut juga dengan busur (*arc*). Pada graf berarah urutan diperhatikan, sehingga untuk setiap sisi:

$$(u,v) \neq (v,u)$$



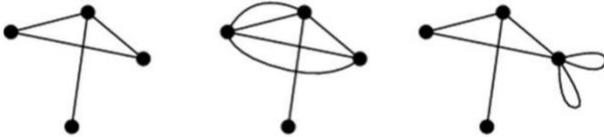
Gambar 1. Graf berarah

Sumber: Munir, Rinaldi. 2016. Matematika Diskrit. Edisi Revisi Keenam. Bandung: Informatika Bandung

d) Graf tak berarah

Graf tak berarah adalah graf yang tidak memiliki orientasi arah pada sisinya, sehingga urutan pada graf tak-berarah tidak diperhatikan. Sisi sisi pada graf berarah dapat direpresentasikan sebagai berikut :

$$(u,v) = (v,u)$$



Gambar 2. Graf tidak berarah

Sumber: Munir, Rinaldi. 2016. Matematika Diskrit. Edisi Revisi Keenam. Bandung: Informatika Bandung

2) Terminologi dalam graf

a) Ketetanggaan (*adjacency*)

Dua buah simpul dikatakan beretetangga jika keduanya berhubungan langsung.

b) Bersisian

Untuk sembarang sisi $e = (V_j, V_k)$ maka dapat dikatakan bahwa:

- e bersisian dengan V_j
- e bersisian dengan V_k

c) Simpul terpencil

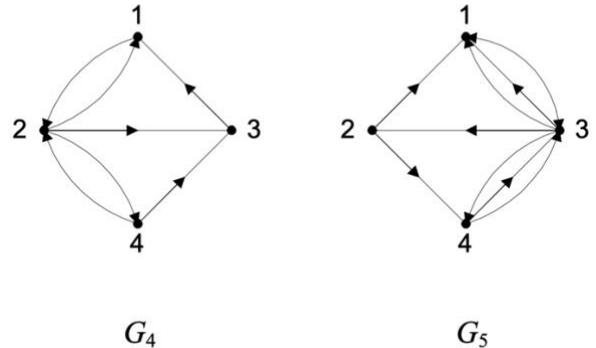
Simpul terpencil adalah simpul yang tidak punya sisi yang bersisian dengan dirinya sendiri.

d) Graf kosong

Graf kosong adalah graf yang sisi-sisinya merupakan himpunan kosong, karena sesuai definisi graf, E atau himpunan sisi bisa jadi kosong. Bisa juga diartikan sebagai graf yang terdiri dari simpul-simpul yang tidak saling terhubung.

e) Derajat

Derajat bisa didefinisikan sebagai jumlah sisi yang bersisian dengan suatu simpul pada graf, pada graf berarah derajat dibedakan menjadi derajat masuk (*din*) dan derajat keluar (*dout*).



Gambar 3. Graf berarah yang punya derajat berbeda

Sumber: Munir, Rinaldi. 2016. Matematika Diskrit. Edisi Revisi Keenam. Bandung: Informatika Bandung

Berdasarkan lemma jabat tangan, jumlah derajat semua simpul pada suatu graf adalah genap, yaitu sejumlah 2 kali dari sisi pada graf tersebut.

f) Lintasan

Lintasan yang memiliki panjang n , yang bergerak dari simpul awal (V_0) ke simpul akhir (V_n) pada sembarang graf adalah barisan selang-seling simpul-simpul dan sisi-sisi yang berbentuk $V_0, e_1, V_1, \dots, e_n, V_n$. Sedangkan panjang lintasan adalah jumlah sisi yang dilewati pada lintasan tersebut.

g) Siklus

Siklus adalah lintasan yang berawal dan berakhir pada simpul yang sama.

h) Keterhubungan

Dua buah simpul semisalnya V_1, V_2 , dikatakan terhubung jika terdapat sebuah lintasan yang menghubungkan keduanya. Pada graf berarah, tentunya lintasan mempertimbangkan arah orientasinya, v_1 dan v_2 dikatakan terhubung kuat jika ada lintasan yang menghubungkan v_1 ke v_2 dan v_2 ke v_1 , jika hanya terdapat 1 arah, maka disebut terhubung lemah.

i) Upagraf dan upagraf merentang

Upagraf bisa diartikan sebagai *subgraph* atau graf yang terdapat di dalam graf. Pada upagraf, semua simpul dan sisinya merupakan anggota dari graf utamanya. Komplemen upagraf adalah bagian dari graf utama yang tidak termasuk dari bagian upagraf. Sebuah upagraf dapat diklasifikasikan sebagai upagraf merentang jika upagraf tersebut memiliki semua simpul dari graf utama.

j) Cut-set

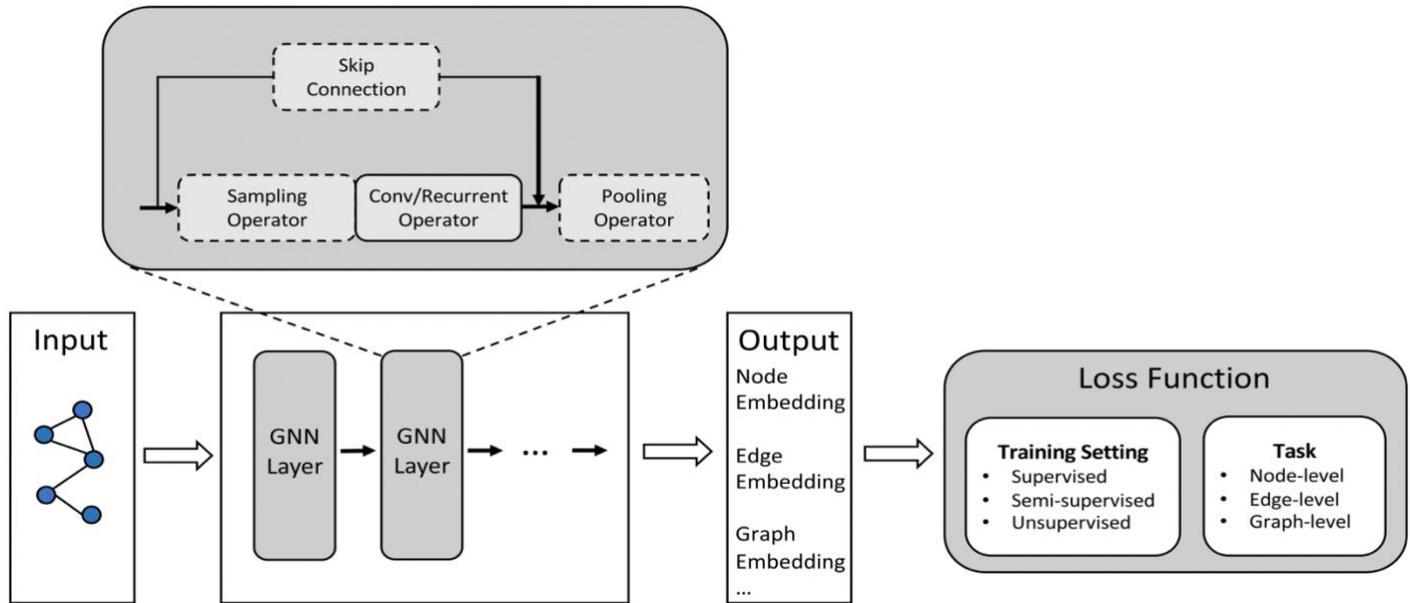
Cut set adalah sisi yang jika diputus/dibuang maka akan menyebabkan sebuah graf menjadi tidak terhubung.

Oleh karena itu jumlah cutset bisa lebih dari satu dan hasil cut-set dinyatakan dalam himpunan pasangan simpul-simpul.

k) *Graf berbobot*

Graf berbobot adalah graf yang pada sisi-sisinya diberi sebuah nilai.

Pada kasus structural, graf dideklarasikan secara eksplisit, dengan arti yaitu bahwa kita tidak perlu membuat graf dari awal. Contoh dari kasus ini bisa dilihat pada kasus-kasus seperti molekul, sistem fisik, dan Knowledge graph(Google).



Gambar 4. Alur pembuatan *Graph Neural Network*
 Sumber: *Graph neural networks: A review of methods and applications*

B. *Graph Neural Network (GNN)*

Graph Neural Network adalah jenis dari *neural network* yang memanfaatkan dan beroperasi dengan menggunakan struktur dari graf. Pengembangan GNN dimulai dari tahun 1990-an, dengan pengembangan awal Recursive Neural Network, Recurent Neural Network, dan FeedForward Neural Network, hal ini menjadi fondasi dasar untuk pengembangan jenis-jenis neural network lainnya. Hingga pengembangan terkini dari neural network yaitu CNN (Convolutional Neural Network), menghasilkan penemuan GNN.

Dalam membuat GNN, tentunya akan ada banyak hal yang perlu kita pertimbangkan untuk membuat suatu GNN yang optimal untuk menyelesaikan masalah kita punya. Hal-hal yang harus kita pertimbangkan sebelum mengembangkan GNN yaitu struktur graf, tipe graf, jenis *loss function*, dan pembangunan model.

1) *Struktur graf*

Pertama-tama. Kita harus mencari tahu struktur dari graf yang ingin kita gunakan, pada umumnya hal ini dibagi menjadi 2 kasus yaitu structural dan non-structural.

a) *Struktural*

Pada kasus non-struktural, graf terdefinisi secara implisit, sehingga kita harus membangun graf secara manual tergantung kasus yang kita miliki. Sebagai contoh, kita harus membuat *word-graph* untuk data text, dan *scene-graph* untuk memproses gambar.

2) *Tipe dan Skala graf*

Setelah memnetentukan struktur dari graf, maka kita harus menentukan tipe dari graf yang akan kita gunakan. Berikut adalah tipe-tipe dari graf yang bis akita pertimbangkan.

a) *Berarah/tidak berarah*

Setiap sisi dalam graf berearah mempunyai arah dari simpul asal ke simpul tujuan, hal ini berarti graf berarah akan menyimpan informasi yang lebih banyak dibandingkan dengan graf tidak berarah.

b) *Homogen/Heterogen*

Tipe-tipe data dalam graf juga berpengaruh, simpul-simpul yang menyimpan data dengan jenis yang sama disebut homogen, dan jika berebda disebut heterogeny, tentunya hal ini perlu dipertimbangkan karena akan ada perbedaan dalam proses mengakses data.

c) *Statis/Dinamis*

Graf dinamis adalah graf yang berubah seiring berjalannya waktu, sehingga graf ini mengandung parameter waktu di dalamnya. Parameter waktu dalam graf ini harus diolah secara teliti agar tidak menimbulkan kesalahan yang tidak diig

3) Jenis *loss function*

Dalam semua jenis neural network, termasuk GNN, tentunya prediksi yang dihasilkan tidak akan selalu tepat. Ketidaktepatan dalam sesuatu model neural network biasa disebut dengan *loss*. Pada dasarnya *loss* adalah ratio/jumlah data hasil prediksi yang tidak sesuai dengan data yang asli. *Loss function* adalah fungsi yang dapat menghitung nilai *loss* dari sebuah model. Pada umumnya, *loss function* disesuaikan dengan tugas yang harus dilakukan sebuah neural network atau dalam kasus ini GNN.

Pada proses *learning* sebuah model graph, pada umumnya dibagi menjadi 3 jenis tugas yaitu node-level, Edge-level, dan Graph-Level

a) *Node-Level*

Tugas *node-level* berfokuskan diri ke simpul-simpul pada graf. Tugas-tugasnya diantaranya *node-clustering, node-regression, node-classification* dll.

b) *Edge-Level*

Tugas-tugas pada edge-level yaitu *classification* dan *link-prediction*. *Classification* berarti membagi/mengklasifikasi sisi-sisi pada graf menjadi tipe-tipe tertentu, sedangkan *link-prediction* memprediksi apakah antar 2 simpul tertentu ada sebuah sisi yang bersisian dengan kedua simpul tersebut (ada hubungan).

c) *Graph-Level*

Graph-level melakukan tugas-tugas yang sama dengan *edge-level* dan *node-level*, akan tetapi ruang lingkupnya adalah seluruh graf. Sehingga pada graph-level, harus diperhatikan jenis representasi graf yang digunakan.

Selain dibagi berdasarkan tugas proses learning sebuah neural network juga bisa diklasifikasikan berdasarkan supervisi/pengawasan, yaitu *supervised, semi-supervised*, dan *unsupervised*. *Supervised* berarti bahwa data yang digunakan dalam proses training sudah mempunyai label yang jelas. Pada semi-supervised learning hanya Sebagian kecil data yang mempunyai label jelas, sedangkan sisanya digunakan untuk keperluan testing untuk memprediksi label yang tepat. *Unsupervised learning* menggunakan data tanpa label, sehingga pada *unsupervised learning*, yang dilakukan pada umumnya hanya sebatas proses clustering data.

4) *Pembangunan Model*

Lalu tahap terakhir dalam pembangunan GNN adalah pembuatan model. Pada umumnya model dikembangkan dengan menggunakan modul-modul komputasi, berikut adalah modul-modul yang paling umum digunakan. 3

a) *Propagation module*

Modul propagasi digunakan untuk mempropagasi/menyebarkan informasi, sehingga informasi tersebut dapat memperoleh informasi fitur dan topologis.

b) *Sampling module*

Pada graf yang berukuran besar, maka modul sampling digunakan untuk melakukan propagasi pada graf. Pada umumnya modul sampling digabungkan dengan modul propagasi.

c) *Pooling module*

Modul pooling digunakan untuk memperoleh informasi pada graf yang lebih rumit (*high-level graphs*).

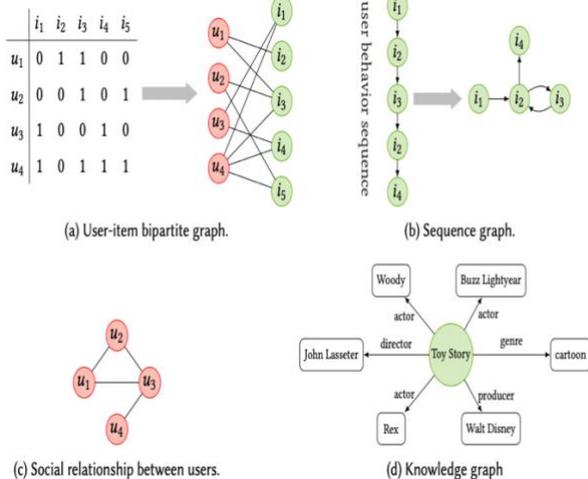
C. *Sistem Rekomendasi*

Sistem rekomendasi adalah aplikasi dari ilmu machine learning yang pada dasarnya mengolah data pengguna dan memunculkan hal yang sekiranya akan menarik bagi pengguna tersebut. Hal ini sudah sangat umum kita temukan, seperti pada layanan streaming, online shop, atau app store dll. Lantas bagaimana sistem ini bermanfaat? Seiring dengan perkembangan teknologi, tentunya jumlah semua data, produk, aplikasi, informasi akan ikut bertambah. Dengan adanya sistem rekomendasi para pengguna tidak akan kesulitan dalam mencari sesuatu diantara jutaan bahkan miliaran data. Dengan memanfaatkan sistem rekomendasi maka pengguna bisa dengan mudah menemukan barang, film, buku, atau aplikasi yang sesuai dengan selera dan kebutuhan mereka. Pada umumnya sistem rekomendasi dibagi menjadi 2 jenis yaitu home-page dan related-item recommendation. Sebagai contoh pada google play store home-page recommendation adalah rekomendasi yang muncul pada halaman utama aplikasi, biasanya berjudul "recommended for you", sedangkan item-related recommendation adalah rekomendasi yang didasarkan pada pencarian pengguna, misalnya pengguna sedang mencari aplikasi tentang sepakbola, maka sistem akan merekomendasikan aplikasi-aplikasi olahraga lainnya.

III. APLIKASI NEURAL NETWORK PADA SISTEM REKOMENDASI

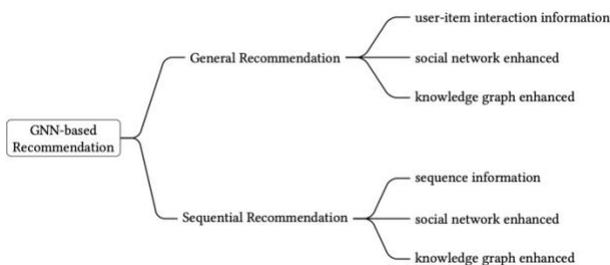
Mengapa Graph Neural Network (GNN) menjadi pilihan yang tepat untuk mengembangkan sistem rekomendasi? Karena telah diteliti bahwa GNN menghasilkan hasil yang memuaskan dalam proses learning dalam representasi graf. Selain itu, pada dasarnya data yang diperlukan sistem rekomendasi juga menggunakan representasi graf seperti graf bipartite, sequence, knowledge graph dan graf yang menggambarkan

hubungan sosial antar pengguna.



Gambar 5. Graf yang dimanfaatkan dalam pembentukan sistem rekomendasi
 Sumber: Graph Neural Networks in Recommender Systems

Sistem rekomendasi pada umumnya dibagi lagi menjadi 2 kategori yaitu *general recommendation* dan *sequential recommendation*. Pembagian ini didasarkan oleh beberapa aspek yaitu jenis informasi yang disimpan oleh graf dan tipe graf yang statis/dinamis.

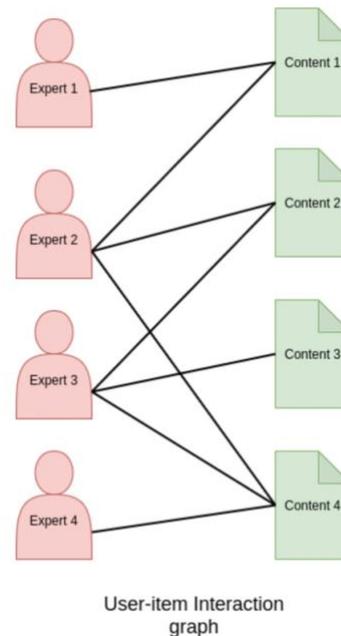


Gambar 6. Kategori dari sistem rekomendasi yang didasarkan oleh GNN

Sumber: Graph Neural Networks in Recommender Systems
 Pada kategori general, diambil asumsi bahwa preferensi pengguna/user tidak berubah seiring berjalannya waktu. Bisa juga ditentukan bahwa tipe graf yang digunakan adalah graf statis. Untuk kategori sekuensial, ide utamanya adalah menangkap pola-pola transisi dalam beberapa rangkaian untuk menentukan rekomendasi berikutnya.

Masing-masing dari kategori tersebut bisa dioptimalisasi/enhance lebih lanjut dengan menggunakan konsep user-item interaction, social network information, dan knowledge graph.

1) User item interaction

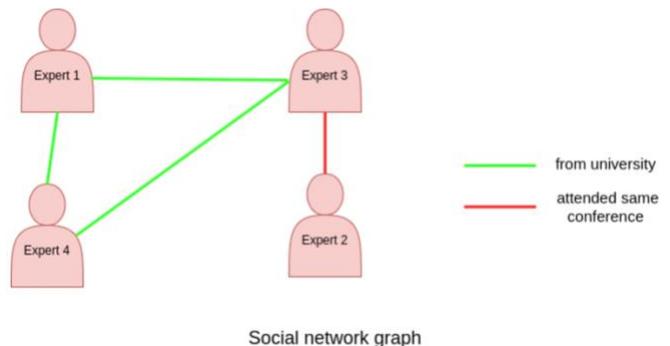


Gambar 7. Graf User-item interaction
 Sumber: towardsdatascience.com

Dalam pemodelan ini, data dimodelkan sebagai graf bipartit. Pada contoh diatas node expert menyatakan penulis yang membuat sebuah konten tertentu. Relasi/sisi yang ada pada graf tersebut menandakan apakah seorang expert merupakan salah satu penulis dari konten tersebut. Ada beberapa model yang umum digunakan untuk GNN dengan user-item interaction yaitu GCN, GraphSage, dan GAT.

2) Social network information

Dalam pemodelan ini, data bisa dimodelkan menjadi 2 tipe graf yaitu menjadi 2 graf yang terpisah atau 1 graf heterogen. Pada 2 graf terpisah, graf terdiri dari satu graf bipartite user-item interaction dan satu graf lainnya yang menyatakan hubungan antar pengguna. Pada graf heterogen maka hubungan user-user dan user-item direpresentasikan di dalam satu graf yang besar.

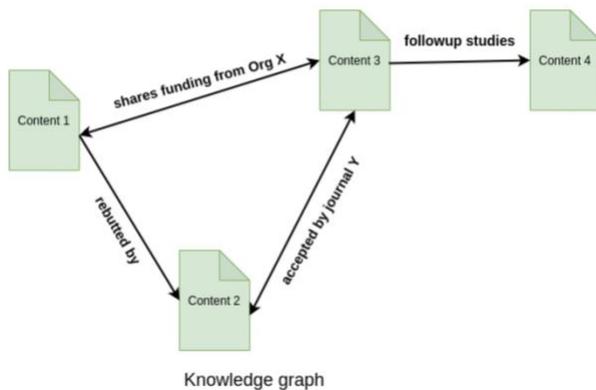


Gambar 8. Graf social-network information
 Sumber: towardsdatascience.com

Jika kita perhatikan contoh diatas, maka user itu sendiri memiliki keterkaitan/relasi antar satu sama lain, sehingga hal ini bisa memudahkan computer dalam mencari pola tentang referensi yang akan disukai oleh pengguna. Selain itu dengan hal ini maka bisa data bisa dianalisis secara lebih dalam karena bisa terlihat pola antara kesamaan pengguna dan preferensi mereka. Misalnya jika dilihat diatas, mungkin akan ada korelasi antara jumlah konten yang ditulis dan kesamaan alumni sebuah universitas.

3) Knowledge graph

Model dalam bentuk ini lebih memfokuskan diri kepada representasi item, hal ini menghasilkan rekomendasi yang lebih akurat yang didasarkan pada interaksi pengguna yang sudah dilakukan sebelumnya.



Gambar 8. Knowledge graph
Sumber: [torwardsdatascience.com](https://towardsdatascience.com)

IV. KESIMPULAN

Teori graf tentunya punya banyak aplikasi dalam penyelesaian masalah di dunia nyata, salah satunya adalah sistem rekomendasi pada platform-platform di era digital ini. Walaupun teori graf hanya membahas dasar dari aplikasi-aplikasi tersebut, akan tetapi penting untuk kita memahaminya, karena GNN hanyalah salah satu dari perkembangan yang ditemukan dan tidak menutup kemungkinan akan ada perubahan yang lebih baik dan lebih optimal kedepannya.

Selain itu, dengan dikembangkannya sistem rekomendasi seperti yang dicantumkan pada makalah ini, kita sebagai pelajar harus membuka pemikiran kita lebih jauh karena dengan pemikiran yang kreatif, kita akan bisa membuat hal-hal yang mungkin sebelumnya dianggap mustahil. Sama seperti computer yang dapat mengetahui kesukaan dan bahkan merekomendasikan hal untuk kita.

V. UCAPAN TERIMA KASIH

Pertama-tama, penulis mengucapkan puji syukur kepada Allah SWT, atas berkat dan rahmat-Nya, penulis bisa menulis makalah ini dengan maksimal. Penulis juga mengucapkan terimakasih kepada Ibu Nur Ulfa Maulidevi selaku dosen mata kuliah IF2120 Matematika Diskrit K03, yang selama ini sudah memberi ilmunya kepada penulis. Penulis juga mengucapkan terima kasih kepada Bapak Rinaldi Munir, yang selama satu semester ini selalu membantu penulis dengan menyediakan materi-materi, soal-soal Latihan yang selalu penulis gunakan untuk mempermudah proses pembelajaran IF2120 Matematika Diskrit.

REFERENSI

- [1] <https://towardsdatascience.com/graph-neural-network-gnn-architectures-for-recommendation-systems-7b9dd0de0856>, diakses pada 3 Desember 2021, pukul 22.13
- [2] <https://developers.google.com/machine-learning/recommendation/overview>, diakses pada 5 Desember 2021 pukul 09.00
- [3] <https://distill.pub/2021/gnn-intro/> diakses pada 13 Desember pukul 17.00
- [4] <https://towardsdatascience.com/a-gentle-introduction-to-graph-neural-network-basics-deepwalk-and-graphsage-db5d540d50b3>, diakses pada 13 Desember pukul 18.00
- [5] Shiwen Wu, Fei Sun, Wentao Zhang, and Bin Cui. 2021. Graph Neural Networks in Recommender Systems: A Survey. J. ACM 37, 4, Article 111 (April 2021)
- [6] Jie Zhou a,1 , Ganqu Cui, Graph neural networks: A review of methods and applications
- [7]

VI. PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 14 Desember 2021

Muhammad Gerald Akbar Giffera
13520143