**PROPOSAL PERMOHONAN DATA**

***RECOMMENDER SYSTEMS***



Diajukan Kepada:

**PT. TOKOBAGUS (OLX INDONESIA)**

Disusun oleh :

**Umiyati**

**0102514021**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS AL AZHAR**

**2018**

# **LEMBAR PENGESAHAN**

**PROPOSAL PERMOHONAN DATA**

**PT. TOKOBAGUS (OLX INDONESIA)**

Diajukan oleh :

Umiyati

0102514021

Jakarta, 08 Maret 2018

Menyetujui,

Dosen Pembimbing I

**Ir. Winangsari Pradani, M.T.**

Dosen Pembimbing II

**Ali Akbar Septiandri, S.T., M.Sc.**

Mengetahui,

Ketua Program Studi Teknik Informatika

Universitas Al Azhar Indonesia

**Riri Safitri, S.Si., M.T.**

# **DAFTAR ISI**

[LEMBAR PENGESAHAN i](#_Toc507108315)

[DAFTAR ISI ii](#_Toc507108316)

[BAB I 1](#_Toc507108317)

[PENDAHULUAN 1](#_Toc507108318)

[1.1. Latar Belakang 1](#_Toc507108319)

[1.2. Tujuan Penelitian 3](#_Toc507108320)

[1.3. Urgensi Penelitian 3](#_Toc507108321)

[1.4. Pemohon 3](#_Toc507108322)

[BAB II 5](#_Toc507108323)

[TINJAUAN PUSTAKA 5](#_Toc507108324)

[BAB III 9](#_Toc507108325)

[METODOLOGI 9](#_Toc507108326)

[DAFTAR PUSTAKA 11](#_Toc507108327)

# **BAB I**

## **PENDAHULUAN**

## **Latar Belakang**

Dewasa ini, hampir semua aplikasi atau Web yang menyajikan konten, seperti Google, Facebook, Twitter, Amazon, Spotify, Netflix, OLX dan sejenisnya telah memiliki fitur berupa rekomendasi konten yang mungkin relevan atau disukai oleh pengguna. Fitur tersebut dikenal dengan *recommender system*.

Bagi beberapa perusahaan, keberadaan *recommender system* menjadi penting untuk membantu menjalankan tugas dan mencapai tujuan perusahaan sebagaimana mestinya [1]. Kenyataannya, implementasi *recommender system* yang baik mampu menghasilkan *traffic* yang sangat tinggi bagi perusahaan. Misalnya, Spotify mengungkapkan bahwa lebih dari 8.000 musisi mendapat *listener* bulanan melalui fitur Discover Weekly [2]. Hal yang sama juga diungkapkan oleh Netflix bahwa 75% film yang disaksikan pengguna didapat dari *recommender system* [3]. Hal tersebut membuat penelitian seputar *recommender system* dalam dekade terakhir ini berkembang dengan pesat. Lebih dari 200 artikel penelitian tentang *recommender system* sudah terpublikasi [4]. Sebagian besar penelitian tersebut bertujuan untuk meningkatkan akurasi prediksi dan rekomendasi. Seiring dengan bertambahnya artikel-artikel tersebut, algoritma rekomendasi yang dikembangkan dan diinovasikan sesuai dengan kebutuhan masing-masing sistem pun semakin banyak. Sehingga banyak bidang yang telah memberikan kontribusi pada pengembangan *recommender system*, termasuk *Information System*, *Information Retrieval* (IR), *Machine Learning* (ML), *Human Computer Interaction* (HCI), dan bahkan disiplin ilmu yang jauh seperti *Marketing* dan *Physics* [1].

Pada penelitian ini, penulis bermaksud untuk mengembangkan *recommender system* pada situs iklan baris *online* terbesar di Indonesia yaitu **OLX Indonesia** [5] dengan tujuan untuk menghasilkan akurasi dan kecepatan rekomendasi yang lebih baik. Rekomendasi yang lebih baik ini nantinya diharapkan mampu meningkatkan *traffic* situs iklan baris *online* tersebut.

Pengembangan *recommender system* pada OLX ini dilakukan penulis karena proses bisnis yang dijalankan oleh OLX berbeda dengan aplikasi atau Web lain yang sama-sama menawarkan barang untuk dijual. Sebagaimana yang telah diketahui pada paragraf sebelumnya, bahwa OLX hanya menampilkan iklan untuk berbagai jenis *item*. Sedangkan untuk transaksi pemesanan bahkan sampai tahap pembayaran, situs tersebut tidak menyediakannya. Namun OLX menyediakan fitur *favorite* sebagai bentuk suka *user* terhadap *item* yang diiklankan.

Dalam *recommender system*, biasanya data konsumsi barang atau *rating* yang diberikan oleh *user* untuk setiap pembelian *item* menjadi penting sebagai dasar yang kuat untuk memberikan rekomendasi *item* yang mungkin akan disukai oleh *user*. Tidak tersedianya data tersebut menyebabkan kita tidak pernah tahu *item* apa saja yang telah dikonsumsi oleh *user*. Tetapi bukan berarti kita tidak bisa memberikan rekomendasi *item* yang mungkin akan dikonsumsi atau dibeli oleh *user*. Melainkan dapat dilakukan dengan memanfaatkan data *favorite* dan jumlah *view* yang diberikan oleh *user* terhadap *item* sebagai dasar pemberian rekomendasi. Dengan asumsi bahwa *item* yang dimasukan kedalam *favorite* dan banyaknya *view* terhadap *item* menandakan bahwa *user* berminat dengan *item* tersebut. Hal ini menjadi tantangan yang menarik untuk dilakukan.

## **Tujuan Penelitian**

Tujuan dari penelitian ini adalah untuk mengetahui apakah *recommender system* yang dibuat mampu meningkatkan akurasi dan kecepatan rekomendasi sehingga mampu menaikan *traffic* bagi OLX.

## **Urgensi Penelitian**

Penerapan fitur *recommender system* pada suatu aplikasi atau Web yang menawarkan konten menjadi hal yang penting dilakukan. Sebagaimana telah dibahas pada pemaparan sebelumnya, bahwa implementasi *recommender system* yang baik mampu meningkatkan *traffic* yang tinggi bagi aplikasi maupun Web dan tentunya mendatangkan keuntungan bagi perusahaan.

Pada tahun 2006, Netflix memberikan hadiah sebesar $1,000,000 untuk siapa saja yang mampu mengalahkan algoritma CineMath yang diterapkan pada *recommender system* miliknya sebesar 10% pada RMSE [3]. Lebih dari 5,000 tim yang mengikuti kompetisi tersebut dan hadiah tersebut akhirnya diberikan pada tahun 2009. Pemenangnya menggunakan gabungan beberapa algoritma dan pendekatan *machine learning* [1].

Uraian diatas menunjukan bahwa peningkatan akurasi suatu *recommender system* sekecil apapun sangatlah beharga dan tentunya membutuhkan kerja keras.

## **Pemohon**

Berikut adalah data diri pemohon :

Nama : Umiyati

NIM : 0102514021

Tempat/Tanggal Lahir : Bekasi, 06 Agustus 1996

Program Studi/Fakultas : Teknik Informatika/Sains dan Teknologi

Alamat : Desa Setia Asih RT002/RW011. Kec.

Tarumajaya, Kab. Bekasi, Jawa Barat.

Kontak : 0895395792421, umiyati1996@gmail.com

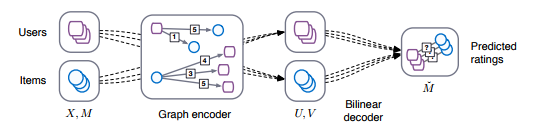
# **BAB II**

## **TINJAUAN PUSTAKA**

Secara Umum, algoritma rekomendasi yang sering digunakan dalam *recommender system*  dibagi menjadi dua kelompok, yaitu *Content-Based Recommender System* dan *Collaborative Filtering* (CF)*.* *Content-based* mendefiniskan profil untuk masing-masing *user* atau *item* [7]. Misalnya, profil film adalah genre, aktor, popularitas, dan lain-lain. Sedangkan untuk profil *user* adalah informasi geografis, isian kuisioner, dan lain-lain. Profil-profil ini memungkinkan sistem untuk mengasosiasikan *user* dengan *items* yang sesuai. Namun algoritma ini membutuhkan informasi eksternal yang boleh jadi susah untuk diperoleh. *Collaborative filtering* didasari pada matriks preferensi antara *user* dan *item.* Terdapat dua metode pada algoritma ini, yaitu *user-based collaborative filtering* (merekomendasikan *item* berdasarkan kesamaan antar pengguna) dan *item-based collaborative filtering* (merekomendasikan *item* berdasarkan kesamaan antar *item*) [7].

Penelitian terkait *recommender system* dengan algoritma CF telah banyak dilakukan, salah satunya [8]. Pada Penelitian tersebut *recommender system* dibangun dengan menerapkan dua algoritma dari *collaborative filtering* yaitu, *user-based* dan *item-based collaborative filtering* untuk melihat perbedaan *performance* dari kedua model tersebut dalam melakukan prediksi dan merekomendasikan *items* yang akan dibeli oleh *user* pada masa depan. *Dataset* yang digunakan *‘Gourmet\_Foods.txt.gz’* merupakan salah satu *dataset* dari Amazon. *Range* dari *rating* pada *dataset* tersebut adalah 1-5, sehingga dalam menghitung kedekatan antar *item* dan kedekatan antar *user*  berdasarkan *rating* yang diberikan oleh *user* terhadap *item,* metode yang dipilih adalah *correlation-based similarity***.** Dari penelitian tersebut, disimpulkan bahwa model *item-based collaborative filtering* menghasilkan kualitas rekomendasi yang lebih baik dibandingkan *user-based collaborative filtering*. Hal ini terjadi karena metode *item-based* bersifat lebih statis (penambahan jumlah *item* cenderung lebih sedikit daripada jumlah *user)* atau dapat diaggap bahwa metode tersebut terlepas dari jumlah *user* dan *item* dalam katalog produk [9], sehingga pada skala yang cukup besarpun metode tersebut dapat bekerja dengan baik. Kondisi tersebut menjadi alasan dipilihnya model *item-based collaborative filtering* untuk *recommender system* pada situs Amazon.

Penelitian terbaru dilakukan oleh [10]. Pada penelitian tersebut, dikenalkan sebuah model *graph convolutional matrix completion* (GC-MC) untuk membangun *recommender system*. Model tersebut merepresentasikan **graf bipartit** (dengan *user* dan *item* sebagai *node* sedangkan *edge* atau *link* yang menghubungkan kedua *node* tersebut menunjukan bahwa *user* memberikan *rating* kepada *item*)sebagai **matriks *user item***. Mengingat bahwa matriks yang dibentuk adalah *sparse matrix,* maka *recommender system* dibangun dengan mempertimbangkan *matrix completion*, yaitu melengkapi kekosongan matriks. Dalam implementasinya, matriks tersebut diolah dengan menggunakan *graph auto-encoder framework*. Pada gambar, pendekatan *auto-encoder* digunakan untuk menghilangkan *noise* untuk mendapatkan representasi gambar yang lebih baik. Dalam *recommender system*, pendekatan tersebut akan menemukan variabel laten atau variabel yang tersembunyi dari *user* dan *item* pada matriks tersebut dengan bantuan ReLu sebagai *activation function*. Variabel laten tersebut diolah dengan model *bilinear decoder* untuk dihasilkan prediksi *items* yang mungkin relevan atau disukai oleh *user*.



Gambar 1. 1 *Graph Convolutional Matrix Completion*

*Dataset* yang digunakan adalah MovieLens 100K(skala sebesar 100,000), 1M (skala sebesar 1,000,029) dan 10M (skala sebesar 10,000,054). Matriks yang dibentuk merupakan matriks *user item* yang berisi *rating* dengan range 1-5 (pada MovieLens 100K dan 1M**)** dan 0.5-5 (pada *dataset* MovieLens 10M). Penelitian tersebut berhasil membuktikan bahwa *performance* dari model yang digunakan mampu menyaingi standar *collaborative filtering benchmark*. Selain itu model yang dihasilkan dapat dilatih pada dataset dengan skala yang lebih besar dan bahkan jika diatur atau diaplikasikan pada *structured data* seperti *social network*, model tersebut mampu melebihi metode terbaru yang telah digunakan.

Berdasarkan penelitian yang telah dilakukan sebelumnya, penulis bermaksud untuk menawarkan beberapa alternatif algoritma atau model untuk membangun *recommender system,* yaitu *user-based collaborative filtering, item-based collaborative filtering,* dan *graph convolutional matrix completion*. Mengingat bahwa barang yang diiklankan oleh OLX.co.id sebagian besar **“barang bekas yang sifatnya sekali habis”**, maka pada model *user-based collaborative filtering*, matriks yang akan digunakan bukanlah matriks *user-item*melainkan matriks *user* dan kategori *item* dengan jumlah *view* yang akan mengisi matriks tersebut. Sehingga *items* yang nantinya akan direkomendasikan bukan *item* yang sama persis melainkan *item* dengan kategori yang sama dengan *item* yang dicari/dilihat/disukai oleh *user*. Misalnya, *user* U mencari atau menyukai *item* berkategori A, B, C dan D. Sedangkan *user* V mencari atau meyukai *item* berkategori A, B dan D. Maka rekomendasi untuk *user* V adalah *item* berkategori C.

Pada model *item-based collaborative filtering,* matriks yang digunakan adalah matriks kategori *item*. Nantinya *items* yang direkomendasikan adalah *items* yang memiliki kedekatan atau kesamaan kategori dengan *item* yang disukai oleh *user*. Misalnya, *user* mencari atau menyukai *item* berkategori A, maka *item*s lain yang akan ditawarkan adalah *items* dengan kategori A atau *items* dengan kategori lain yang berdekatan dengan kategori A.

Pada model *graph convolutional matrix completion,* penulis akan melakukan hal yang sama dengan yang dilakukan oleh [10]. Namun penulis akan mencoba untuk mengaplikasikan matriks *user-*kategori *item* dan matriks kategori *item* secara bergantian untuk melihat *performance* terbaik yang dihasilkan.

Alternatif-alternatif tersebut diajukan oleh penulis sebagai bahan pertimbangan bagi pihak OLX.co.id agar bersedia memberikan dukungan kepada penulis untuk menjalani penelitian tersebut sebagaimana mestinya.

# **BAB III**

## **METODOLOGI**

Penulis menawarkan beberapa alternatif algoritma rekomendasi pada penelitian ini. Alternatif tersebut adalah:

1. *User-Based Collaborative Filtering*
2. *Item-Based Collaborative Filtering*
3. *Graph Convolutional Matrix Completion*

Berdasarkan alternatif yang ditawarkan diatas, maka data yang relevan yang sekiranya dapat diperoleh dari subjek penelitian untuk menunjang penelitian ini adalah sebagai berikut :

Tabel 3. 1 Daftar Data Yang Dibutuhkan

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Data** | **Variabel** | **Tahun** |
| 1 | Iklan | * ID Iklan * ID penjual * Judul Iklan * Kategori Iklan * Harga Iklan | 2017 |
| 2 | Kontak | * ID Pengontak * ID Iklan yang dikontak | 2017 |
| 3 | *View* | * ID Pengunjung yang melihat * ID Iklan yang dikontak | 2017 |

# 

# **DAFTAR PUSTAKA**

|  |  |
| --- | --- |
| [1] | Jannach, D., Resnick, P., Tuzhilin, A., Zanker, M. (2016). *Recommender Systems-Beyond Matrix Completion*. Communications of the ACM |
| [2] | Buskirk, E.V. (2016). *The Most Streamed Music from Spotify Discover Weekly*. [Online] Diakses dari https://insights.spotify.com/us/2016/07/07/top-music-discover-weekly/ (Diakses tanggal 03 Februari 2018) |
| [3] | Amatriain, X., Basilico, J. (2012). *Netflix Recommendation Beyond 5 Stars (Part 1).* [Online] Diakses dari https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429 (Diakses tanggal 20 Februari 2018) |
| [4] | Beel, J., Langer, S., Genzmehr, M., Gipp, B. and Nürnberger, A. (2013). A Comparative Analysis of Offline and Online Evaluations and Discussion of Research Paper Recommender System Evaluation. *Proceedings of the Workshop on Reproducibility and Replication in Recommender Systems Evaluation (RepSys) at the ACM Recommender System Conference (RecSys)* (2013), 7–14 |
| [5] | OLX Blog. [Online] Diakses dari http://blog.olx.co.id/about-us/ (Diakses tanggal 03 Februari 2018) |
| [7] | Leskovec, J., Rajaraman, A., & Ullman, J. D. (2014). *Mining of massive datasets*. Cambridge University Press. (Chapter 9) |
| [8] | Yao, G., & Cai, L. *User-Based and Item-Based Collaborative Filtering Recommendation Algorithms Design*. University of California, San Diego. |
| [9] | Linden, G., Smith, B., & York, J. (2003). *Amazon.com Recommendations Item-to-Item Collaborative Filtering*. IEEE Internet Computing. |
| [10] | Berg van den, R., Kifp, T., Welling, M. (2017). *Graph Convolutional Matrix Completion*. University of Amsterdam. |