

LAPORAN PRAKTEK KERJA LAPANGAN

**RANCANG BANGUN SISTEM REKOMENDASI PRODUK E-COMMERCE I-
SHOP DENGAN METODE COLLABORATIVE FILTERING**



Disusun oleh:

MUHAMMAD ILHAM MAULANA

F1D017059

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MATARAM

2020

LEMBAR PENGESAHAN
LAPORAN PRAKTEK KERJA LAPANGAN

RANCANG BANGUN SISTEM REKOMENDASI PRODUK E-COMMERCE I-SHOP DENGAN METODE COLLABORATIVE FILTERING

Disusun oleh:
MUHAMMAD ILHAM MAULANA
F1D 017 059

Telah disetujui oleh :

Tanggal :

1. Dosen Pembimbing

1. 21 Juni 2020




Andy Hidayat Jatmika, ST., M.Kom.
NIP. 19831209 201212 1 001

2. Pembimbing Lapangan

2. 22 Juni 2020



Robert Silas Kabanga, S.Kom., M.Eng
NIP: 19860513 201101 1 015

Mengetahui,
Sekretaris Program Studi Teknik Informatika
Fakultas Teknik
Universitas Mataram



Andy Hidayat Jatmika, ST., M.Kom.
NIP. 19831209 201212 1 001

KATA PENGANTAR

Puji syukur penulisan panjatkan ke hadirat Allah SWT, Tuhan semesta alam, karena dengan berkat, rahmat dan limpahan karunianya sehingga penulis dapat menyelesaikan Laporan Praktek Kerja Lapangan ini sebagaimana mestinya.

Adapun Praktek Kerja Lapangan (PKL) ini penulis laksanakan di Dinas Komunikasi Informatika dan Statistik Pemerintah Provinsi Nusa Tenggara Barat dengan judul “Sistem Rekomendasi Produk *E-commerce* I-Shop Dengan Metode *Collaborative Filtering*”. Sistem rekomendasi produk *e-commerce* digunakan untuk mempermudah pengguna situs I-shop yang untuk mendapatkan produk yang relevan dan sesuai dengan apa yang dicari pengguna. Dalam pembuatan laporan ini penulis berpedoman pada bahan kuliah, petunjuk dari pembimbing lapangan, dosen pembimbing, referensi dan literatur yang terkait dengan penulisan laporan.

Penulis menyadari laporan ini masih jauh dari kesempurnaan. Oleh karena itu, penulis mengharapkan kritik serta saran yang bersifat membangun agar dapat menghasilkan karya yang lebih baik dimasa mendatang. Semoga laporan ini dapat bermanfaat bagi penulis dan pembaca sekalian.

Mataram,

2020

Penulis

DAFTAR ISI

DAFTAR ISI.....	iv
DAFTAR GAMBAR.....	v
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan.....	2
1.5 Manfaat.....	3
BAB II TINJAUAN INSTANSI TEMPAT PKL.....	4
2.1 Profil Dinas Komunikasi Informatika dan Statistik Provinsi NTB.....	4
2.2 Visi dan Misi Dinas Komunikasi Informatika dan Statistik Provinsi NTB	5
2.4 Struktur Organisasi Dinas Komunikasi Informatika dan Statistik Provinsi NTB	5
BAB III LANDASAN TEORI	7
3.1 Sistem Rekomendasi.....	7
3.2 Metode-Metode pada Sistem Rekomendasi	9
3.3 Metode Evaluasi.....	11
3.4 Faktorisasi Matriks.....	11
BAB IV PEMBAHASAN.....	13
4.1 Analisa Sistem.....	13
4.2 Analisa Masalah.....	13
4.3 Deskripsi Kebutuhan Sistem (Dataset).....	14
4.4 <i>Collaborative Filtering</i>	14
4.5 Proses <i>Collaborative Filtering</i>	14
BAB V PENUTUP	16
5.1 Kesimpulan.....	16
5.2 Saran.....	16
DAFTAR PUSTAKA	17
LAMPIRAN	

DAFTAR GAMBAR

Gambar 2. 1 Struktur Organisasi Dinas Komunikasi Informatika dan Statistik NTB.....	6
Gambar 3. 1 Taksonomi sumber informasi pada sistem informasi	8
Gambar 3. 2 Matriks <i>User x Item</i> pada Sistem Rekomendasi <i>Collaborative Filtering</i> ..	10
Gambar 3. 3 Faktorisasi Matriks.....	11
Gambar 4. 1 <i>Representasi data menggunakan matriks</i>	14
Gambar 4. 2 Hasil evaluasi <i>popularity based recommendation</i>	16
Gambar 4. 3 Hasil evaluasi sistem rekomendasi <i>model collaborative filtering</i>	16
Gambar 4. 4 Pengujian jumlah <i>latent factor</i> pada model collaborative filtering	17

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sistem rekomendasi didefinisikan sebagai aplikasi pada website e-commerce untuk mengusulkan informasi dan menyediakan fasilitas yang diinginkan pengguna dalam membuat suatu keputusan. Sistem ini diasumsikan seperti penggambaran kebutuhan dan keinginan pengguna melalui pendekatan metode rekomendasi dengan mencari dan merekomendasikan suatu *item* dengan menggunakan *rating* berdasarkan kemiripan dari karakteristik informasi pengguna.

Semakin banyaknya informasi produk yang ada di internet menghadirkan tantangan baik pembeli maupun pebisnis online dalam lingkungan e-commerce. Pembeli sering mengalami kesulitan saat mencari produk di internet karena banyaknya produk yang dijual di internet. Pebisnis online sering mengalami kesulitan karena memiliki data mengenai produk, pembeli, dan transaksi yang sangat banyak, sehingga menyebabkan pebisnis online mengalami kesulitan untuk mempromosikan produk yang tepat pada target pembeli tertentu. Sistem rekomendasi secara otomatis dapat menganalisis penggunaan data calon pembeli untuk menyaring konten halaman web, mengkategorisasi pesan newsgroup, dan merekomendasikan informasi. Sistem rekomendasi menganalisis data mengenai produk atau interaksi pengguna dan produk untuk menemukan hubungan antara produk dan pengguna. Hasil yang diterima akan ditampilkan sebagai rekomendasi [1].

Sistem rekomendasi bisa memanfaatkan teknik *Collaborative Filtering* atau yang biasa disebut dengan *crowd-wisdom* adalah salah satu metode rekomendasi yang menggunakan data *rating* dari seorang pengguna, dan pengguna lain untuk menghasilkan rekomendasi. *Collaborative filtering* menganggap bahwa selera pengguna terhadap suatu *item* atau barang akan cenderung sama dari waktu ke waktu. Ditambah lagi, pengguna yang menyukai suatu *item* biasanya juga akan menyukai *item* lain yang disukai oleh pengguna lain yang juga menyukai *item* yang sama dengan pengguna tersebut.

I-Shop adalah aplikasi web yang sedang dikembangkan oleh Kominfo NTB yang menyediakan penjualan produk-produk UKM yang ada Nusa Tenggara Barat. Saat ini sudah banyak orang yang memanfaatkan internet untuk memesan produk secara *online*. Dengan semakin banyaknya pengguna yang melakukan transaksi *online*, maka diperlukan sistem yang dapat mempermudah pengguna dalam mencari produk yang sesuai dengan

apa yang dicari. Saat ini pengguna yang ingin mencari sebuah produk, sistem akan merekomendasikan produk-produk terpopuler yang didapatkan dari seberapa banyak orang yang membeli produk tersebut. Masalah terjadi ketika produk yang ditampilkan tidak relevan dengan apa yang dicari oleh pengguna.

Penelitian ini bertujuan mendesain sistem rekomendasi dengan metode *Collaborative Filtering* dan menguji kinerja sistem rekomendasi tersebut kembali menggunakan *precision* dan *recall*. Selanjutnya dibuat aplikasi dengan algoritma yang efektif dalam merekomendasikan produk dengan metode *Collaborative Filtering*, sehingga pengguna mudah mencari produk yang diinginkan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka dapat dirumuskan permasalahan yaitu “Bagaimana merancang dan menguji Sistem Rekomendasi Produk I-Shop pada Dinas Komunikasi, Informatika dan Statistik NTB?” sehingga dapat merekomendasikan produk yang relevan dengan keinginan pengguna.

1.3 Batasan Masalah

Berdasarkan rumusan masalah yang ada, berikut adalah batasan-batasan masalah yang akan dibahas dalam laporan ini :

1. Bagaimana merancang Sistem Rekomendasi Produk I-Shop menggunakan metode *Collaborative Filtering*.
2. Bagaimana menguji hasil rekomendasi dari Sistem Rekomendasi Produk I-Shop.

1.4 Tujuan

Berdasarkan rumusan masalah yang telah diuraikan, tujuan Praktek Kerja Lapangan (PKL) yang dilakukan yaitu:

1. Membuat perancangan dan menguji menguji Sistem Rekomendasi Produk I-Shop pada Dinas Komunikasi, Informatika dan Statistik NTB sehingga dapat merekomendasikan produk yang relevan kepada pengguna.
2. Mengetahui apakah dengan menggunakan sistem rekomendasi *Collaborative Filtering*, akan lebih mempermudah pengguna dalam mendapatkan produk yang relevan dan sesuai dengan yang dicari pengguna.

1.5 Manfaat

Manfaat yang diperoleh dari pelaksanaan Praktek Kerja Lapangan (PKL) ini antara lain adalah sebagai berikut :

1. Memberikan kemudahan bagi pengguna I-shop untuk mendapatkan produk yang relevan dan sesuai dengan apa yang diinginkan.
2. Memudahkan UKM sebagai penjual produk dalam memasarkan produknya.

BAB II

TINJAUAN INSTANSI TEMPAT PKL

2.1 Gambaran Umum Instansi

Dinas Komunikasi dan Informatika adalah Dinas yang mempunyai tugas melaksanakan kewenangan daerah di bidang pengelolaan Teknologi Informasi dan Komunikasi serta melaksanakan tugas pembantuan yang diberikan oleh pemerintah dan pemerintah provinsi dimana dalam setiap kegiatannya selalu berhubungan dengan pembangunan dan pengembangan system informasi, pengembangan dan pemeliharaan jaringan computer antar bidang, pengelolaan produksi informasi dan publikasi, pengelolaan dan pengembangan komunikasi publik, yang mana pada setiap kegiatan-kegiatan tersebut terbagi menjadi tiga bidang serta satu Sekretariat dan dikepalai oleh kepala bidang dari setiap bidangnya

Sebagai lembaga pemerintahan yang mempunyai tanggung jawab besar dan bergerak di dalam lingkungan Pemerintah NTB, maka KOMINFO mempunyai tugas pokok dan fungsi yang besar dalam membangun Teknologi Informasi dan Komunikasi (TIK) di Provinsi NTB. Dinas Komunikasi dan Informatika Provinsi NTB saat ini berkedudukan dan menempati kantor dengan alamat Jalan Udayana No. 14 Mataram - NTB , Telp. (0370) 644264; Fax (0370) 7509831.

2.2 Visi Misi

a. Visi

Membangun Nusa Tenggara Barat Gemilang.

b. Misi

Misi dari Program Dinas Komunikasi, Informatika dan Statistik NTB yaitu:

1. NTB Tangguh dan Mantap

Melalui penguatan mitigasi bencana dan pembangunan infrastruktur serta konektivitas wilayah.

2. NTB Bersih dan Melayani

Melalui transformasi birokrasi yang berintegritas, berkinerja tinggi, bersih dari KKN dan berdedikasi.

3. NTB Sehat dan Cerdas

Melalui peningkatan kualitas sumber daya manusia sebagai pondasi daya saing daerah.

4. NTB Asri dan Lestari

Melalui pengelolaan sumber daya alam dan lingkungan yang berkemajuan.

5. NTB Sejahtera dan Mandiri

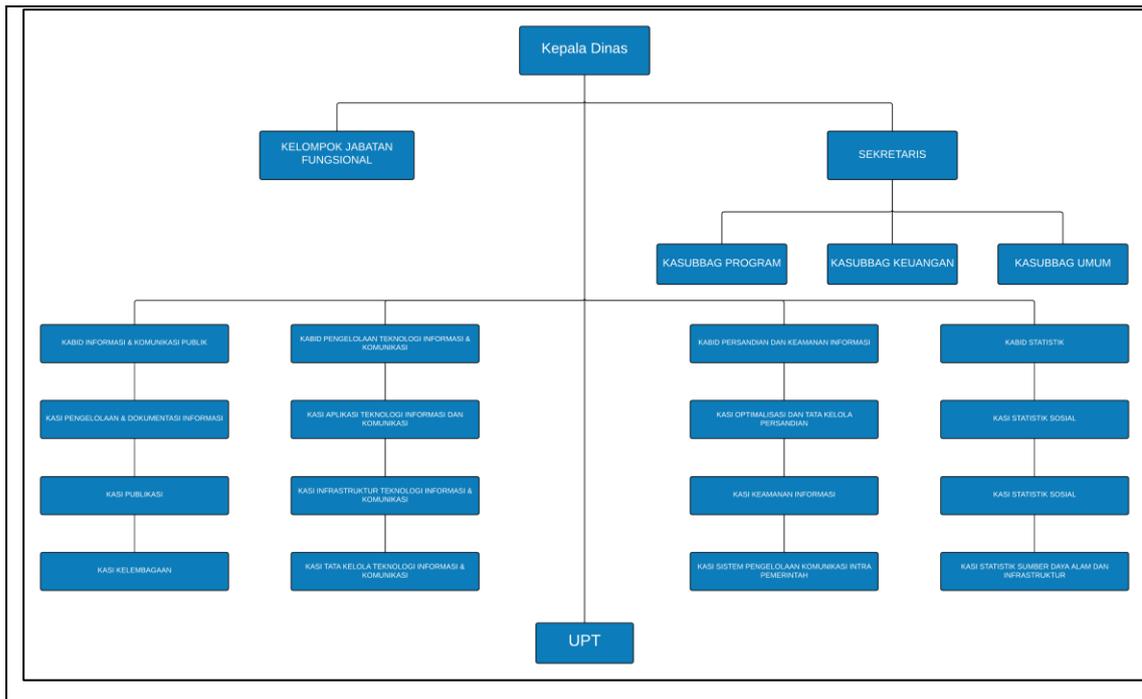
Melalui penanggulangan kemiskinan, mengurangi kesenjangan, pertumbuhan ekonomi inklusif bertumpu pada pertanian, pariwisata dan industrialisasi.

6. NTB Aman dan Berkah

Melalui perwujudan masyarakat mandani yang beriman, berkarakter dan penegakan hukum yang berkeadilan.

2.3 Struktur Organisasi

Struktur Organisasi Dinas Komunikasi Informatika dan Statistik Pemerintah Provinsi Nusa Tenggara Barat dibentuk berdasarkan Peraturan Gubernur Nusa Tenggara Barat Nomor 50 Tahun 2016 tentang Kedudukan, Susunan Organisasi, Tugas dan Fungsi serta Tata Kerja Dinas Daerah Provinsi Nusa Tenggara Barat (Sebagaimana telah diubah dengan Peraturan Gubernur Nomor 46 Tahun 2018 tentang Perubahan Kedua Atas Peraturan Gubernur Nusa Tenggara Barat Nomor 50 Tahun 2016 tentang Kedudukan, Susunan Organisasi, Tugas dan Fungsi serta Tata Kerja Dinas Daerah Provinsi Nusa Tenggara Barat) dapat dilihat pada Gambar 2.1.



Gambar 2.1 Struktur Organisasi Dinas Komunikasi Informatika dan Statistik NTB

BAB III

DASAR TEORI

3.1 Sistem Rekomendasi

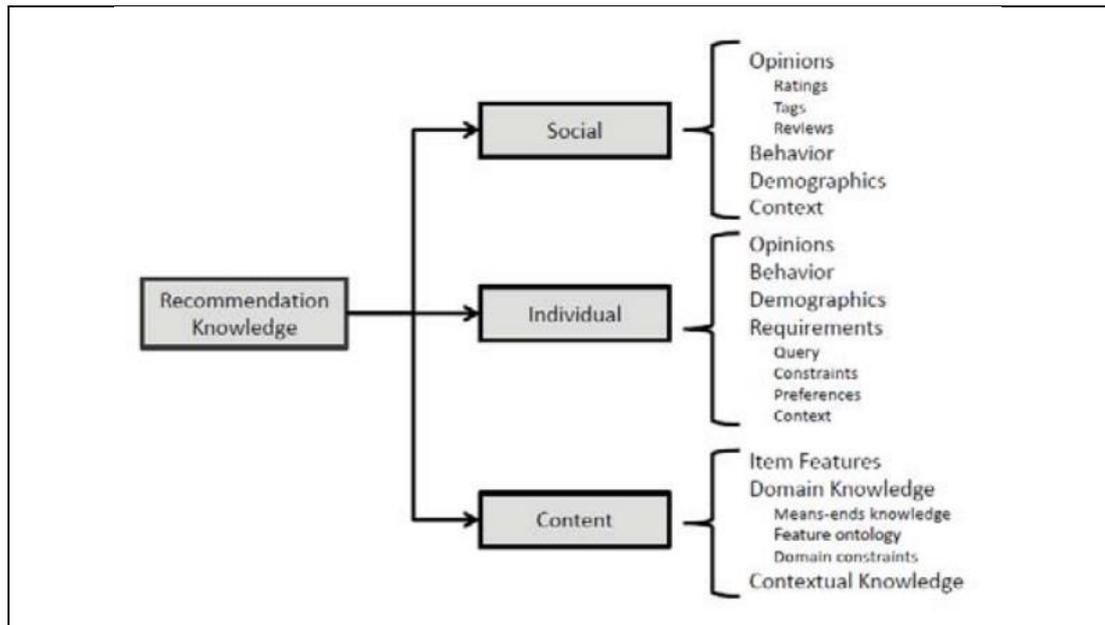
Konsep sistem rekomendasi telah digunakan secara luas oleh hampir semua area bisnis dimana seorang konsumen memerlukan informasi untuk membuat suatu keputusan. Sistem rekomendasi paket wisata atau perjalanan menggunakan konsep ini dapat menolong para wisatawan untuk memutuskan tujuan perjalanan mereka, akomodasi dan aktivitas di tempat tujuan [2].

Sistem rekomendasi adalah sistem yang berfungsi untuk membantu pengguna dalam mengidentifikasi produk yang sesuai dengan kebutuhan, kesenangan, dan keinginan pengguna. Sistem rekomendasi akan membimbing pengguna untuk menemukan produk yang relevan dan berguna dari banyaknya produk yang tersedia [3].

Sistem rekomendasi mengandalkan dua tipe informasi yang berbeda, yaitu explicit input dan implicit input. Explicit merupakan informasi didapat dari hasil penilaian yang diberikan langsung oleh pengguna, misalnya pemberian *rating* atau thumbs-up/down untuk *item* tertentu. Sedangkan implicit merupakan informasi didapat dengan cara mengamati kebiasaan pengguna, misalnya catatan pembelian, catatan penelusuran, pola pencarian, atau bahkan *click-stream* [3].

Sebuah sistem rekomendasi juga dapat diklasifikasi berdasarkan sumber informasi. Gambar 3.1 menunjukkan taksonomi sumber informasi yang digunakan dalam sistem informasi. Ada 3 jenis sistem rekomendasi berdasarkan sumber informasinya :

- a. Informasi sosial tentang pengguna secara umum.
- b. Informasi individual tentang pengguna tertentu yang mana rekomendasi untuk pengguna tersebut dicari.
- c. pengetahuan konten tentang *item* yang sedang direkomendasikan, mulai dari daftar sederhana hingga yang lebih kompleks yang memungkinkan sistem rekomendasi untuk mengetahui tentang bagaimana suatu barang dapat memenuhi kebutuhan pengguna [4].



Gambar 3.1 Taksonomi sumber informasi pada sistem informasi.

Tantangan utama sebuah sistem rekomendasi adalah:

- *Data Sparsity*: Hal ini sangat jarang bahwa dua pengguna menilai *item* yang sama berkali-kali. Hal ini membuat lebih sulit untuk menghitung kesamaan.
- *Cold start for user*: Pengguna yang belum melakukan *rating* terhadap beberapa barang dapat menyebabkan ia mendapatkan hasil yang tidak akurat. Hal ini mungkin terjadi karena pengguna lain yang memiliki kemiripan dengan pengguna ini tidak dapat ditemukan
- *Cold start for item*: *Item* yang baru diperkenalkan memungkinkan tidak mendapatkan *rating* yang cukup, yang menyebabkan tidak direkomendasikan kepada pengguna.
- *Attacks*: Sistem rekomendasi memiliki ancaman terhadap serangan, seperti menyalin keseluruhan profil pengguna dan membuat sistem berpikir bahwa si penyerang dan pengguna tersebut sangat mirip. Hal ini membuat penyerang dapat membodohi sistem dan membuatnya menyarankan *item* apapun yang di *rating* oleh penyerang kepada pengguna [5].

3.2 Metode-Metode pada Sistem Rekomendasi

Ada beberapa metode yang biasa digunakan dalam sistem rekomendasi yang dibagi menjadi empat yaitu :

a. Sistem rekomendasi berbasis popularitas (*popularity based recommendation*)

Adalah metode rekomendasi yang biasanya sulit dikalahkan dan menjadi standar metode yang lain.. Sistem rekomendasi dengan metode ini tidak terpersonalisasi untuk setiap user. Metode ini simpel dimana sistem ini hanya merekomendasi *item* paling populer yang belum pernah di konsumsi oleh pengguna.

b. Sistem rekomendasi berbasis logika *fuzzy*

Adalah metode rekomendasi yang menggunakan logika *fuzzy* untuk melakukan filtrasi terhadap fitur-fitur yang berisi profil pengguna. Fitur-fitur tersebut biasanya bersifat tidak pasti dan kabur. Hal tersebut membuat logika *fuzzy* cocok digunakan karena metode logika *fuzzy* terkenal dapat mengakomodir data bukan biner dan bersifat non linier [5].

c. Sistem rekomendasi berbasis konten (*content based recommendation*)

Ada berbagai macam metode pendekatan yang digunakan untuk menyelesaikan permasalahan pada sistem rekomendasi, antara lain *user-based collaborative filtering*, *content based filtering* dan *hybrid*. *Content-based filtering* memberikan suatu rekomendasi berdasarkan hasil analisa kemiripan *item* yang telah dinilai oleh penggunanya. *Content-Based Filtering* membentuk profil penggunanya berdasarkan atribut pembentuk suatu *item*. *Cosine Similarity* yang tertinggi yang akan menjadi hasil rekomendasi.

Content-Based Filtering membentuk profil penggunanya berdasarkan atribut pembentuk suatu *item*. Sebagai contoh untuk suatu dokumen, atribut pembentuknya adalah katakata yang terdapat pada dokumen tersebut. Parameter pembentuk profil pengguna ini juga diberi nilai bobot berdasarkan kriteria tertentu. Adapun langkah – langkah algoritmanya :

- 1) Suatu *item* barang dibagi-bagi berdasarkan suatu vektor komponen pembentuknya.
- 2) Sistem akan membuat profil pengguna berdasarkan bobot vektor komponen pembentuk suatu *item*. Pembuatan profil pengguna dapat menggunakan algoritma TF-IDF (*term frequency-invers document frequency*). TF adalah jumlah term dalam suatu dokumen.

3) Berdasarkan profil pengguna tersebut, sistem akan memperkirakan penilaian suka atau tidak suka suatu *item* berdasarkan analisis kemiripan profil pengguna dengan vektor komponen pembentuk *item*. Jika sistem memperkirakan bahwa *item* tersebut akan disukai oleh pengguna maka *item* tersebut akan direkomendasikan ke pengguna.

d. Sistem rekomendasi berbasis kolaborasi pengguna (*collaborative filtering*)

Ide utama dalam sistem rekomendasi *collaborative filtering* adalah untuk memanfaatkan opini pengguna lain yang ada untuk memprediksi *item* yang mungkin akan disukai/diminati oleh seorang pengguna [6]. Teknik ini menggunakan asumsi mendasar dimana opini pengguna lain dapat dipilih dan di agregasikan untuk memberikan prediksi dari preferensi pengguna aktif. Intinya, diasumsikan bahwa apabila beberapa pengguna mempunyai minat yang sama terhadap suatu buku, maka besar kemungkinan mereka mempunyai minat yang sama juga untuk buku yang lain [7]. Sistem rekomendasi *collaborative filtering* beroperasi di dalam sebuah ruang dua dimensi *User x Item*. *Rating* yang dapat diberikan oleh seorang pengguna terhadap sebuah *item* dapat direpresentasikan sebagai R (bilangan bulat tidak negatif atau bilangan real dengan jarak tertentu), dan sistem rekomendasi mencoba untuk memprediksi *rating* yang akan diberikan seorang user untuk sebuah *item* yang belum pernah ia beri *rating* sebelumnya. Misalnya, terdapat lima pengguna $u_1, u_2, u_3, \dots, u_5$ dan lima *item* $i_1, i_2, i_3, \dots, i_5$. Sebuah sistem rekomendasi ingin memprediksi berapa *rating* yang akan diberikan oleh u_1 kepada i_5 .

	Item i_1	Item i_2	Item i_3	Item i_4	Item i_5
User u_1	5	7	5	7	?
User u_2	5	7	5	7	9
User u_3	5	7	5	7	9
User u_4	6	6	6	6	5
User u_5	6	6	6	6	5

Gambar 3.2 Matriks User x Item pada Sistem Rekomendasi *Collaborative Filtering*

Terdapat sebuah cara yang sering digunakan dalam menghitung / memprediksi *rating* pengguna *u* terhadap *item* *i* dengan mengkalkulasikan rata-rata *rating* pengguna lain (*neighbor*), yaitu rumus *weighted sum*:

$$R'(u, i) = \frac{1}{\sum_{u' \in N(u)} |sim(u, u')|} \sum_{u' \in N(u)} sim(u, u') * R(u', i)$$

Dimana:

- $R'(u, i)$ adalah prediksi *rating* pengguna *u* terhadap *item* *i*.
- $R(u', i)$ adalah *rating* pengguna u' (pengguna selain *u*) terhadap *item* *i*.
- $sim(u, u')$ adalah kemiripan antara pengguna *u* dengan u' .
- $N(u)$ adalah satu set pengguna lain yang mirip dengan pengguna *u*.

3.3 Metode Evaluasi

Metode evaluasi sangatlah penting dalam proyek pembelajaran mesin, karena evaluasi memungkinkan perbandingan secara objektif algoritma-algoritma yang berbeda beserta parameter-parameternya. Salah satu aspek penting dalam metode evaluasi adalah memastikan bahwa model pembelajaran mesin yang dipakai *men-generalisasi* untuk data yang tidak pernah dilatih.

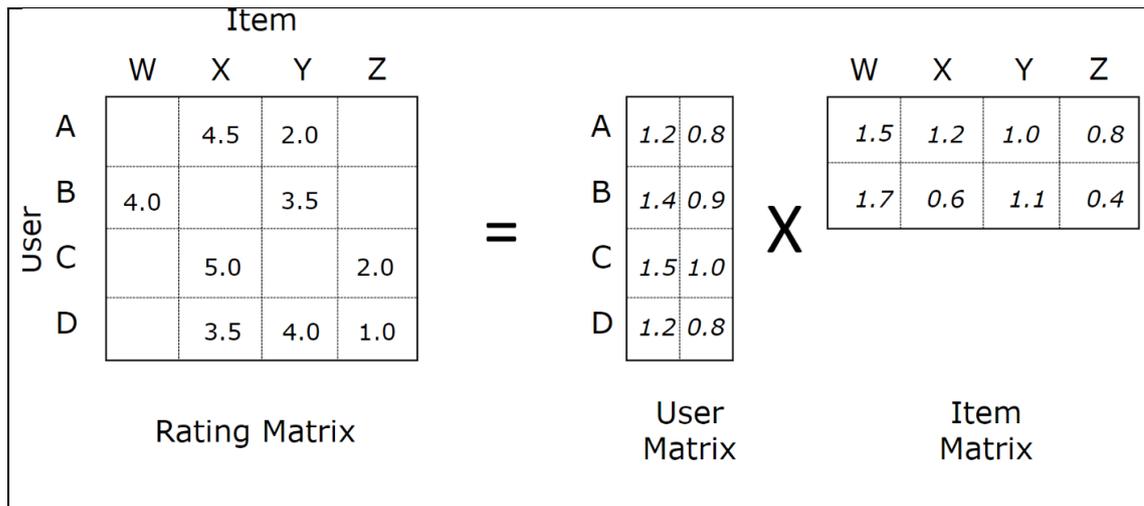
3.3.1 Recall@N

Analisa performa sistem dihitung menggunakan metode *recall@n*. Dimana *recall* adalah persentase semua *item* yang relevan yang dikembalikan dari hasil rekomendasi. perhitungan *recall* dilakukan menggunakan rumus [8]:

$$Recall = \frac{\text{item relevan yang direkomendasi}}{\text{total item yang disukai oleh pengguna}}$$

3.4 Faktorisasi Matriks

Faktorisasi Matriks, merupakan salah satu metode untuk memfaktorkan sebuah matriks yang besar menjadi matriks yang lebih kecil, yang mewakili matriks original.



Gambar 3.3 Faktorisasi Matriks

3.4.1. Singular Value Decomposition

Suatu proses dekomposisi dalam matriks berarti memfaktorkan sebuah matriks menjadi lebih dari satu matriks. Salah satu teknik dekomposisi yaitu SVD yang berkaitan dengan nilai singular (*singular value*) suatu matriks, dimana nilai *singular* mencerminkan karakteristik matriks tersebut. SVD didasarkan pada teori aljabar linier, bahwa suatu matriks persegi panjang dimensi $M \times N$ dapat dipecah atau difaktorkan menjadi perkalian dari 3 buah matriks, yaitu matriks ortogonal U , matriks diagonal Σ dan *transpose* dari matriks ortogonal V .

BAB IV

HASIL DAN PEMBAHASAN

4.1 Analisis Sistem

Diskominfo NTB berencana akan membuat *e-Commerce* I-Shop yang nantinya akan digunakan sebagai tempat UKM-UKM yang ada di NTB untuk menjual barang-barangnya. Untuk itu, dibutuhkan sistem rekomendasi yang dapat merekomendasikan produk-produk yang ada pada *3-commerce* tersebut kepada pengguna secara personalisasi.

4.2 Deskripsi Kebutuhan Sistem (Dataset)

Dataset yang digunakan dalam melatih sistem rekomendasi ini adalah *Articles sharing and reading from CI&T DeskDrop*. Penggunaan dataset tersebut dilakukan karena dataset tersebut memiliki banyak karakteristik yang sama dengan apa yang dibutuhkan dalam dataset *E-Commerce* seperti adanya beberapa tipe interaksi yang dilakukan pengguna. Dataset *Articles sharing and reading from CI&T DeskDrop* memiliki data yang akan digunakan sebagai landasan sistem sebagai berikut :

- a. Histori data sampel *real* selama 12 bulan (maret 2016 – februari 2017) dari *CI&T's Internal Communication platform (DeskDrop)*.
- b. Dataset ini memiliki feedback implisit yang banyak (*comment, like, dan view*).

4.3 Implementasi Sistem

Implementasi sistem pada sistem rekomendasi terdiri atas beberapa langkah. Langkah pertama yang dilakukan adalah *data cleaning* dan *data munging*. Kemudian langkah selanjutnya adalah mengubah data menjadi matriks. Matriks yang didapatkan kemudian akan di faktorisasi dengan menggunakan metode *Singular Vector Decomposition*. Dari faktor-faktor matriks yang didapatkan inilah nantinya rekomendasi ke pengguna akan dilakukan.

4.5.1 Data Munging

Pada langkah ini, dilakukan perubahan pada data yang ada pada dataset karena terdapat beberapa interaksi implisit dari pengguna seperti *comment, bookmark, view* dan lain-lain. Sehingga perlu dilakukan perubahan data tersebut menjadi data numerik agar dapat dikalkulasi. Karena terdapat beberapa interaksi implisit antara pengguna dan produk, maka setiap interaksi tersebut dapat diasosiasikan dengan berat (*weight*) dan kekuatan (*strength*). Contohnya, pengguna yang menyukai sebuah produk, memiliki daya

tarik yang lebih besar kepada produk tersebut daripada pengguna yang hanya melihat produk tersebut. Sehingga dapat dilakukan perubahan data menjadi numerik sebagai berikut :

1. *View* dengan nilai *strength* 1.0 .
2. *Like* dengan nilai *strength* 2.0 .
3. *Bookmark* dengan nilai *strength* 2.5 .
4. *Follow* dengan nilai *strength* 3.0 .
5. *Comment* dengan nilai *strength*: 4.0 .

4.5.2 Data Cleaning

Sistem rekomendasi menggunakan metode *collaborative filtering* memiliki masalah yang disebut dengan *user cold start* dimana terjadi kesulitan untuk merekomendasi yang terpersonalisasi untuk pengguna yang memiliki jumlah interaksi sedikit, yang disebabkan oleh kurangnya informasi untuk memodelkan preferensi pengguna tersebut. Oleh karena itu, dilakukan *filtering* data untuk menyaring pengguna yang memiliki minimal 5 interaksi dengan produk.

4.5.3 Representasi matriks

Pada langkah ini, data yang telah dibersihkan kemudian akan direpresentasikan dalam bentuk matriks. Salah satu cara merepresentasikan data *rating* pengguna kedalam sebuah matriks adalah dengan memperlakukan data pengguna sebagai kolom dan data *rating* produk sebagai baris dalam matriks atau sebaliknya.

contentId	-9222795471790223670	-9216926795620865886	-9194572880052200111	-9192549002213406534
personId				
-9223121837663643404	0.0	0.0	0.0	0.0
-9212075797126931087	0.0	0.0	0.0	0.0
-9207251133131336884	0.0	2.0	0.0	0.0
-9199575329909162940	0.0	0.0	0.0	0.0
-9196668942822132778	0.0	0.0	0.0	0.0
-9188188261933657343	0.0	0.0	0.0	0.0
-9172914609055320039	0.0	0.0	0.0	0.0
-9156344805277471150	0.0	0.0	0.0	0.0
-9120685872592674274	0.0	0.0	0.0	0.0
-9109785559521267180	0.0	0.0	0.0	0.0

Gambar 4.1 Representasi data menggunakan matriks

Seperti yang dapat dilihat pada Gambar 4.1, terlihat bahwa matriks yang dihasilkan masih banyak yang kosong (*sparse*). Nilai nol pada matriks yang didapatkan tersebut merepresentasikan produk yang belum pernah di-*rating* oleh pengguna. Hal yang akan dilakukan oleh sistem rekomendasi adalah mengisi memprediksi *rating* yang masih kosong tersebut dan merekomendasikan produk-produk tersebut kepada pengguna berdasarkan *rating* yang didapat.

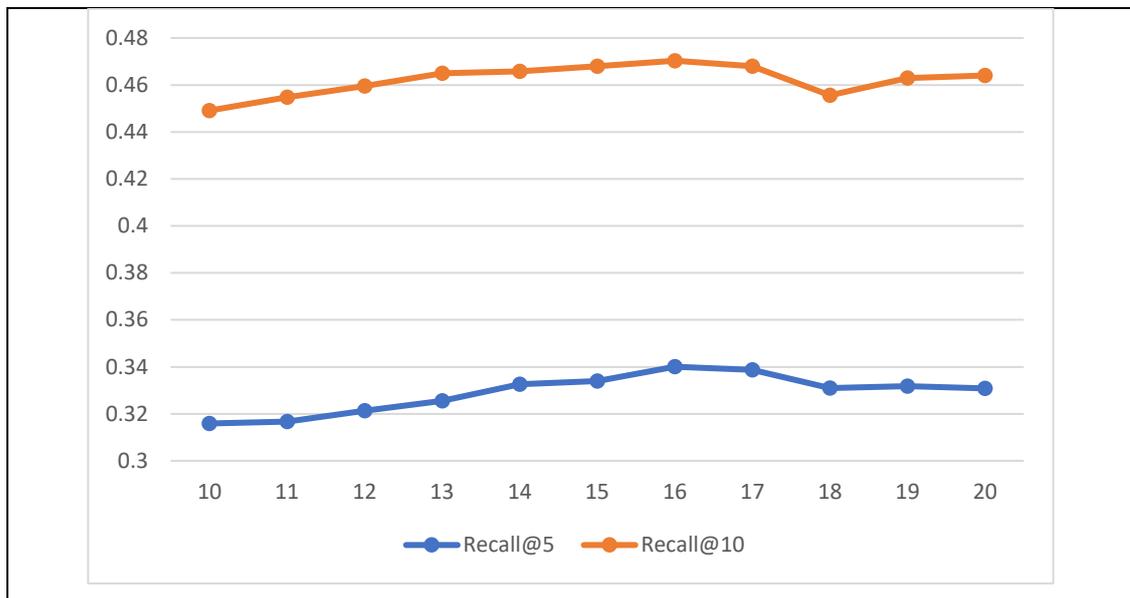
4.5.4 Faktorisasi Matriks

Faktorisasi dari matriks pengguna-produk-*rating* akan didapatkan dengan menggunakan metode *singular value decomposition*. Metode *singular value decomposition* adalah metode faktorisasi matriks menjadi 3 matriks berbeda. Ketiga matriks tersebut merepresentasikan *latent factor* atau karakteristik dari data dimana dalam kasus ini, adalah kategori dari produk. Banyaknya *latent factor* yang akan digunakan dalam melakukan faktorisasi akan mempengaruhi performa dari sistem rekomendasi. Oleh karena itu, untuk mengetahui jumlah *latent factor* yang optimal untuk digunakan, dilakukan pengujian dengan menggunakan 10-20 *latent factor*.

4.5.5 Optimalisasi Sistem

Pada langkah ini, dilakukan optimalisasi sistem rekomendasi model *collaborative filtering* dilakukan dengan menentukan jumlah *latent factor* optimal yang akan digunakan dalam faktorisasi matriks, sesuai dengan yang dijelaskan pada langkah faktorisasi matriks. Jumlah *latent factor* yang digunakan dalam faktorisasi matriks mempengaruhi akurasi dalam merekomendasikan produk kepada pengguna. untuk itu

dilakukan pengujian untuk jumlah *latent factor* 10 sampai dengan 20 untuk mengetahui jumlah *latent factor* yang optimal seperti yang diperlihatkan pada Gambar 4.4.



Gambar 4.4 Pengujian jumlah *latent factor* pada model *collaborative filtering*

Pada gambar 4.4, dapat dilihat bahwa akurasi dari model *collaborative filtering* meningkat dengan bertambahnya jumlah *latent factor* hingga *latent factor* berjumlah 16. Setelah melebihi jumlah tersebut, akurasi dari model *collaborative filtering* menurun. Sehingga jumlah *latent factor* optimal yang digunakan dalam model *collaborative filtering* ini berjumlah 16.

4.5.6 Evaluasi Sistem

Pada langkah ini, dilakukan evaluasi terhadap performa sistem rekomendasi. Salah satu aspek penting dalam evaluasi adalah memastikan bahwa hasil yang didapatkan, dapat digeneralisasi untuk data baru yang belum pernah dilihat oleh sistem. Metode yang digunakan dalam pengujian ini adalah metode *cross-validation* yang mana, dari data yang sudah didapatkan dari langkah-langkah sebelumnya, data-data tersebut akan dipisah menjadi dataset *training*, dan dataset *test*. Semua hasil evaluasi yang dilakukan pada pengujian kali ini dilakukan pada dataset *test* yang sama untuk setiap metode. Selain metode *cross-validation* untuk evaluasi sistem rekomendasi, digunakan juga metode *recall@n* untuk menghitung akurasi sistem rekomendasi.

Untuk menguji sistem rekomendasi menggunakan metode *collaborative filtering*, diperlukan sebuah garis dasar atau *baseline* untuk mengetahui seberapa efektif sistem rekomendasi menggunakan *collaborative filtering* dibandingkan dengan cara rekomendasi tradisional atau yang biasa disebut dengan *popularity based*

recommendation. Dengan menggunakan *popularity based recommendation* didapatkan hasil seperti Gambar 4.2.

```
Global metrics:
{'namaModel': 'Popularity', 'recall@5': 0.242, 'recall@10': 0.373}
```

	_person_id	hits@10_count	hits@5_count	interacted_count	recall@10	recall@5
76	3609194402293569455	50	28	192	0.260417	0.145833
17	-2626634673110551643	25	12	134	0.186567	0.089552
16	-1032019229384696495	23	13	130	0.176923	0.100000
10	-1443636648652872475	9	5	117	0.076923	0.042735
82	-2979881261169775358	40	26	88	0.454545	0.295455
161	-3596626804281480007	18	12	80	0.225000	0.150000
65	1116121227607581999	34	20	73	0.465753	0.273973
81	692689608292948411	23	17	69	0.333333	0.246377
106	-9016528795238256703	18	14	69	0.260870	0.202899
52	3636910968448833585	28	21	68	0.411765	0.308824

Gambar 4.2 Hasil evaluasi *popularity based recommendation*

Pada Gambar 4.2, hasil evaluasi dari model *popularity* mencapai 24.2% pada *recall@5* yang berarti, dari 5 produk yang belum berinteraksi dengan pengguna, terdapat 24.2% produk yang direkomendasikan akan berinteraksi dengan pengguna ketika dilakukan *cross-validation* dengan dataset *test*. Selain itu, hasil evaluasi dari model *popularity* ini pada *recall@10* mencapai 37.3%.

Setelah mendapatkan hasil evaluasi pada model *popularity*, dilakukan pengujian terhadap model *collaborative filtering* menggunakan metode yang sama dan dataset *test* dataset *training* yang sama dan didapatkan hasil evaluasi seperti pada Gambar 4.3.

```
Global metrics:
{'NamaModel': 'Collaborative Filtering', 'recall@5': 0.334, 'recall@10': 0.468}
```

	_person_id	hits@10_count	hits@5_count	interacted_count	recall@10	recall@5
76	3609194402293569455	46	21	192	0.239583	0.109375
17	-2626634673110551643	56	30	134	0.417910	0.223881
16	-1032019229384696495	34	16	130	0.261538	0.123077
10	-1443636648652872475	51	38	117	0.435897	0.324786
82	-2979881261169775358	48	39	88	0.545455	0.443182
161	-3596626804281480007	34	22	80	0.425000	0.275000
65	1116121227607581999	32	24	73	0.438356	0.328767
81	692689608292948411	21	16	69	0.304348	0.231884
106	-9016528795238256703	28	20	69	0.405797	0.289855
52	3636910968448833585	30	23	68	0.441176	0.338235

Gambar 4.3 Hasil evaluasi sistem rekomendasi model *collaborative filtering*

Hasil evaluasi model *collaborative filtering* pada Gambar 4.3 menggunakan 15 *latent factor* dan dari hasil yang didapatkan model *collaborative filtering* mendapatkan hasil *recall@5* sebesar 33.4%, dan *recall@10* sebesar 46.8%.

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil rancangan Sistem Rekomendasi Produk I-Shop yang diusulkan dalam PKL ini, maka dapat diambil beberapa kesimpulan yaitu:

- a. Setelah dilakukan pengujian, rancangan Sistem Rekomendasi Produk I-Shop menggunakan metode *collaborative filtering* berhasil memberikan sistem rekomendasi yang relevan kepada pengguna sesuai dengan hasil pengujian yang menghasilkan akurasi *recall@5* sebesar 33.4%, dan *recall@10* sebesar 46.8%.
- b. Rancang bangun sistem rekomendasi produk I-Shop menggunakan metode *collaborative filtering* yang diusulkan berhasil mengalahkan model *popularity* sesuai dengan hasil pengujian yang dilakukan dimana sistem yang diusulkan mencapai tingkat akurasi *recall@5* sebesar 33.4%, dan *recall@10* sebesar 46.8%. sedangkan model *popularity* mendapat hasil 24.2% pada *recall@5* dan 37.5% pada *recall@10*.

5.2 Saran

Adapun saran yang dapat diberikan dalam rancang bangun sistem rekomendasi produk I-Shop menggunakan metode *collaborative filtering* :

- a. Sistem Rekomendasi Produk I-Shop masih membutuhkan data awal sebagai basis dalam rekomendasi yang akan diberikan pengguna. Hal ini dapat menyebabkan pengguna mendapatkan hasil yang tidak akurat
- b. Produk yang baru dimasukkan kedalam sistem memungkinkan tidak mendapatkan *rating* yang cukup, yang menyebabkan tidak direkomendasikan kepada pengguna..

DAFTAR PUSTAKA

- [1] C. S. D. Prasetya, "Sistem Rekomendasi Pada E-Commerce Menggunakan K-Nearest Neighbor," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 3, p. 194, 2017.
- [2] B. T. W. Utomo and A. W. Anggriawan, "Sistem Rekomendasi Paket Wisata Se-Malang Raya Menggunakan Metode Hybrid Content Based Dan Collaborative," *J. Ilm. Teknol. Inf. Asia*, vol. 9, no. 1, pp. 6–13, 2015.
- [3] M. Rahmadhani, Dina Maulana, Firman Nurul Agnia, "Collaborative Filtering Yang Memanfaatkan Implicit Feedback Datasets Song Recommendation System With Collaborative Collaborative Filtering," *Politek. Negeri Bandung*, 2012.
- [4] Y. Park, "Recommender systems: An overview," *Encycl. E-bus. Dev. Manag. Glob. Econ.*, vol. 3, 2010.
- [5] A. J. Kuswinta, I. G. P. W. Wedashwara W, and I. W. A. Arimbawa, "Implementasi IoT Cerdas Berbasis Inference Fuzzy Tsukamoto pada Pemantauan Kadar pH dan Ketinggian Air dalam Akuaponik," *J. Comput. Sci. Informatics Eng.*, vol. 3, no. 1, pp. 65–74, 2019.
- [6] F. Ricci, L. Rokach, and B. Shapira, "Recommender Systems Handbook," in *Recommender Systems Handbook*, 2010.
- [7] M. I. Fathurrahman, D. Nurjanah, and R. Rismala, "Sistem Rekomendasi Pada Buku Dengan Menggunakan Metode Trust-Aware Recommendation Recommendation System For Book By Using Trust-Aware Recommendation Method," *e-Proceeding Eng.*, vol. 4, no. 3, pp. 4966–4977, 2017.
- [8] A. Bayu and A. Priyono, "Performa Apriori dan Collaborative Filtering untuk Sistem Rekomendasi," *J. Ilm. Inform. dan Komput.*, vol. 21, no. 100, pp. 51–59, 2016.