

Implementation of Collaborative Filtering in the Salted Fish Recommendation Process

Moh Taufiq Rizky^{1*}, Rinianty², Deny Wiria Nugraha³, Amriana⁴, Nouval Trezandy Lapatta⁵

* Program Studi Teknologi Informasi, Fakultas Teknik Informatika, Universitas Tadulako, Palu
mohtaufiqrizky194@gmail.com¹, riniantyinformatika@gmail.com², deny.wiria.nugraha@untad.ac.id³, amriana@untad.ac.id⁴,
nouval@untad.ac.id⁵

Article Info

Article history:

Received 2025-10-24

Revised 2025-11-22

Accepted 2025-11-15

Keyword:

*Collaborative Filtering
E-Commerce
Sales Recommendation
Website.*

ABSTRACT

The development of e-commerce in the current era has been so rapid that buying and selling transactions are carried out online through various media, including websites and applications. With so many products available in the application, users often feel confused when choosing the product they want to buy, so it takes a long time to choose a product to avoid regret after purchasing it. In this study, a web-based recommendation system was created for the process of recommending salted fish with the aim of making it easier for customers to choose the type of salted fish. The Collaborative Filtering method was used, employing Pearson Correlation as a tool to calculate the similarity value between users, then using Weighted Sum to calculate the prediction value. Collaborative Filtering often experiences the cold start problem, where the system has difficulty providing recommendations to users who do not yet have a transaction history. Therefore, the author proposes a popularity-based strategy as a measure to overcome this problem. Based on testing, the author obtained results of MAE = 0.63 and RMSE = 0.81 based on train-test split results with a data distribution of 80:20, 80% of the dataset for training and 20% of the dataset for testing with an accuracy of 70-80%, indicating that this system works well. This system has been tested using the Blackbox method.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. INTRODUCTION

Dalam industri bisnis saat ini yang semakin berkembang, beberapa wilayah yang ada di Indonesia telah memunculkan ide-ide yang dapat meningkatkan hasil pendapatan pemasaran melalui aplikasi ataupun *website*. Seperti yang kita ketahui bahwa Indonesia merupakan salah satu negara dengan hasil laut yang cukup tinggi, sehingga salah satu profesi pekerjaan yang banyak dilakoni oleh masyarakat Indonesia adalah nelayan. Dengan banyaknya hasil laut yang didapatkan, ikan asin merupakan salah satu hasil produk yang dipasarkan dipasar yang dimana ikan asin ini bukan hasil dari ikan yang tidak layak pakai melainkan ikan segar yang di keringkan menggunakan garam sebagai bahan pengawet agar kualitas tetap terjaga. Proses penjualan produk ikan asin ini pelanggan yang ingin membeli maupun mencari informasi tentang produk baru dari pasar ini harus datang secara langsung. Melalui internet/*website*, baik pelaku bisnis maupun konsumen dapat dengan mudah melakukan transaksi secara

online dimana pun dan juga kapan pun. Sistem *web e-commerce* merupakan suatu alat yang dibuat agar kegiatan jual beli antara pelaku usaha dan juga konsumen dapat dilaksanakan secara luas sehingga konsumen tidak perlu repot harus pergi berkunjung ke wilayah tempat barang yang ingin dia beli akan tetapi bisa melakukannya dari jarak yang jauh. Dengan cara ini konsumen akan menghemat biaya transportasi karena jika barang yang ingin dibeli cukup jauh maka biaya transportasinya akan mahal karena harus pergi pulang dan dengan solusi ini pelaku usaha juga bisa mendapatkan pelanggan tidak hanya dari wilayah tempat usahanya berada. Pada sistem ini ditambahkan metode *Collaborative Filtering* sebagai alat rekomendasi [1][2][3]. Sistem rekomendasi merupakan sistem pembelajaran mesin yang bertujuan untuk memberikan prediksi berupa nilai atau suatu tindakan pada suatu item yang diberikan oleh sejumlah pengguna, sehingga pengguna dapat menemukan item baru yang sesuai dengan apa yang disukai berdasarkan preferensi masing-masing pengguna [4][5]. *Collaborative Filtering*

adalah teknik rekomendasi yang digunakan untuk memprediksi preferensi pengguna berdasarkan perilaku pengguna lain yang memiliki kesamaan dengan pengguna tersebut [6]. *Collaborative Filtering* ini sangat berguna bagi pengguna baru yang ingin dan bingung dalam memilih suatu produk yang akan di cobanya. Dengan bantuan teknik ini sistem akan merekomendasikan produk yang memiliki preferensi paling tinggi berdasarkan pengalaman dari pengguna-pengguna sebelumnya[7][8][9]. pada kasus ini adalah ikan asin yang menjadi ciri khas dari wilayah Kecamatan Pagimana. Dalam prosesnya nilai rating merupakan suatu komponen penting dalam sistem rekomendasi ini, nilai rating diperoleh dari konsumen yang pernah mengkonsumsi ikan asin.

Berdasarkan publikasi Pemerintah Sulawesi Tengah pada tahun 2022, produksi ikan tangkap sebesar 15.456 ton dengan nilai produksi sebesar 346,05 miliar rupiah. Volume produksi selama periode tersebut menurun sebesar 3,7 ton atau sebanyak 20 persen dari nilai pendapatan pada tahun 2019 yang nilai produksinya mencapai 19.174 ton dengan total nilai produksi sebesar 347,28 miliar rupiah yang disebabkan oleh kurangnya dalam sektor pengelolaan(sumber:BPS Kabupaten Banggai). Pemilihan ikan asin sebagai domain dikarenakan ikan asin ini merupakan produk unggulan lokal dari wilayah kecamatan pagimana dengan pasar *e-commerce* yang berkembang pesat namun menghadapi sparsity data akibat variasi regional seperti ikan asin kerapu dan kakap merupakan ikan asin yang populer dari wilayah ini. Hal ini berbeda dari domain umum seperti film, dimana data lebih dens dan mudah di prediksi. Urgensi akademiknya terletak pada pengembangan sistem rekomendasi untuk produk niche, yang dapat meningkatkan penjualan berdasarkan studi serupa pada *e-commerce* lokal.

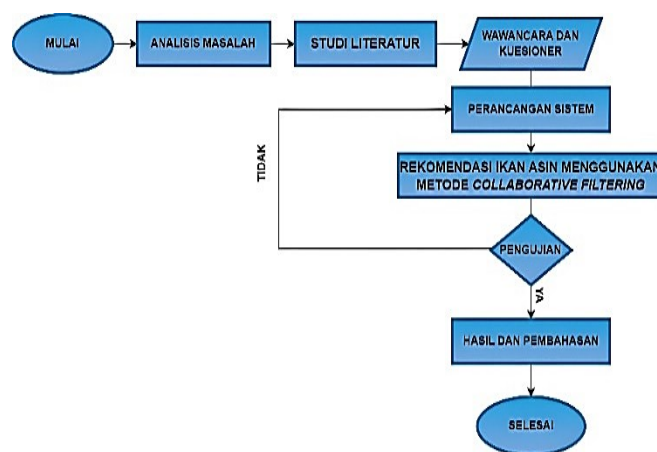
Dalam kasus ini penulis menambahkan sebuah logika dalam menyempurnakan proses rekomendasi ikan asin, dimana *collaborative filtering* akan mengalami *cold start problem* yaitu sistem akan kesulitan dalam memahami pola preferensi pengguna dalam memberikan rekomendasi ketika terdapat pengguna baru yang belum memiliki riwayat transaksi, sehingga penulis memberikan logika tambahan ketika pengguna belum memiliki riwayat transaksi maka pengguna tersebut akan diberikan rekomendasi berdasarkan ikan asin yang populer dikalangan pengguna-pengguna sebelumnya. Hal ini diambil oleh penulis karena lebih efektif dalam memberikan rekomendasi dimana rekomendasi tetap diberikan berdasarkan ikan asin yang banyak disukai oleh pengguna melalui rata-rata rating dari pengguna-pengguna sebelumnya.

Tujuan dari penelitian ini ialah mengimplementasikan program penjualan produk ikan asin berbasis *web* menggunakan metode *collaborative filtering*. Data rating yang telah dikumpulkan dari beberapa pengguna yang pernah mengkonsumsi ikan asin khususnya dari daerah pagimana akan dikumpulkan dan dimasukkan kedalam perhitungan similaritas, setelah itu akan ditentukan nilai prediksi berdasarkan nilai similaritas yang telah dikumpulkan

menggunakan metode *Weighted Sum* guna menghasilkan rekomendasi yang lebih akurat dan relevan bagi pengguna dimana kontribusi yang penulis masukkan meliputi implementasi hybrid antara *user-based collaborative filtering* dan rekomendasi popularitas untuk *cold start*, yang dimana berbeda dari penelitian sebelumnya yang hanya menggunakan *collaborative filtering murni* tanpa solusi praktis untuk data sparse. Dalam prosesnya alur penelitian dapat dilihat pada gambar berikut:

II. METHOD

Pada penelitian ini, metode yang digunakan dalam proses pengembangan sistem adalah metode *waterfall*. metode *waterfall* merupakan suatu metode pengembangan sistem yang menggunakan pendekatan sequensial dan linier, dimana setiap tahap harus diselesaikan sebelum melanjutkan ketahap berikutnya[10]. Metode ini akan menggambarkan alur dari sistem yang dibuat dengan ketentuan yang telah di tetapkan.



Gambar 1. Tahapan Penelitian

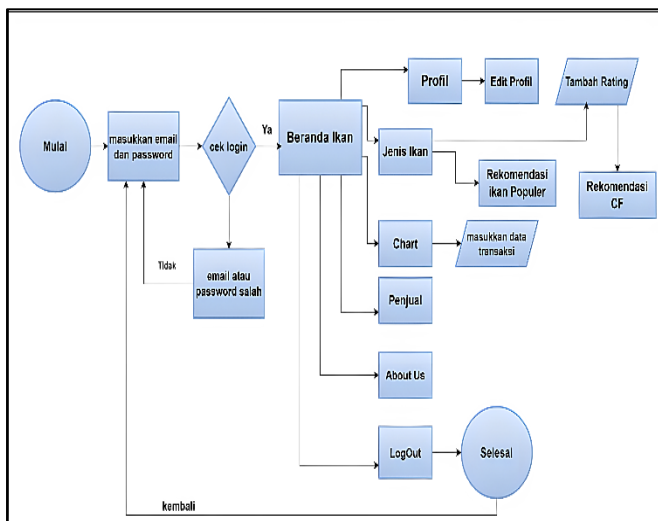
A. Analisis Kebutuhan

Pada tahapan ini merupakan tahapan pengumpulan kebutuhan dan persyaratan dari sistem yang akan dibuat agar dapat berfungsi sesuai dengan tujuan dari sistem *collaborative filtering* itu sendiri dimana data yang dikumpulkan berupa dataset yang terdiri dari 50 pengguna, 12 jenis ikan asin, dan 600 interaksi rating dari skala 1-5. Densitas data sekitar 10% menunjukkan sparsity yang mempengaruhi akurasi *collaborative filtering*, berbeda dari dataset Movielens yang lebih dens (5-10%). Dalam hal ini diketahui bahwa pada sistem penjualan ikan asin di Kecamatan Pagimana masih menggunakan sistem penjualan manual yang dimana prosesnya masih menunggu pembeli datang ke lokasi pasar ikan asin tersebut untuk melakukan transaksi jual beli dengan demikian akan memakan waktu dan biaya terlebih lagi untuk peminat ikan asin yang berada diluar kecamatan pagimana, selain kurangnya informasi yang cukup terkait jenis-jenis ikan asin yang akan dibeli pastinya akan membutuhkan waktu dan biaya yang cukup besar jika harus datang langsung ketempat penjualan ikan asin ini. Kurangnya pemanfaatan sistem informasi akan menyebabkan

pendapatan usaha menurun. Sistem informasi adalah suatu sistem yang mengumpulkan data, memproses, serta menyimpan dan mendistribusikan informasi untuk mendukung pengambilan keputusan, operasional, dan strategi organisasi [11][12][13]. Sehingga penulis berinisiatif membangun sebuah sistem informasi dalam hal ini sistem penjualan berbasis *web* yang berfungsi sebagai media penyaluran informasi terkait jenis-jenis ikan asin dan menggabungkannya dengan metode *collaborative filtering* sebagai alat rekomendasinya [14][15][16]. Sistem rekomendasi merupakan suatu sistem yang dirancang untuk merekomendasikan produk atau jasa kepada pengguna berdasarkan data perilaku pengguna dan juga karakteristik produk atau jasa [17]. Penulis mengumpulkan data-data yang diperlukan dalam pembuatan sistem melalui wawancara dengan penjual serta mengumpulkan data rating terhadap jenis-jenis ikan dari individu-individu yang pernah mengkonsumsinya melalui kuesioner. Mengumpulkan informasi terkait penelitian yang serupa melalui studi literatur menjadi salah satu langkah penting dalam membantu proses pembuatan sistem.

B. Desain Sistem

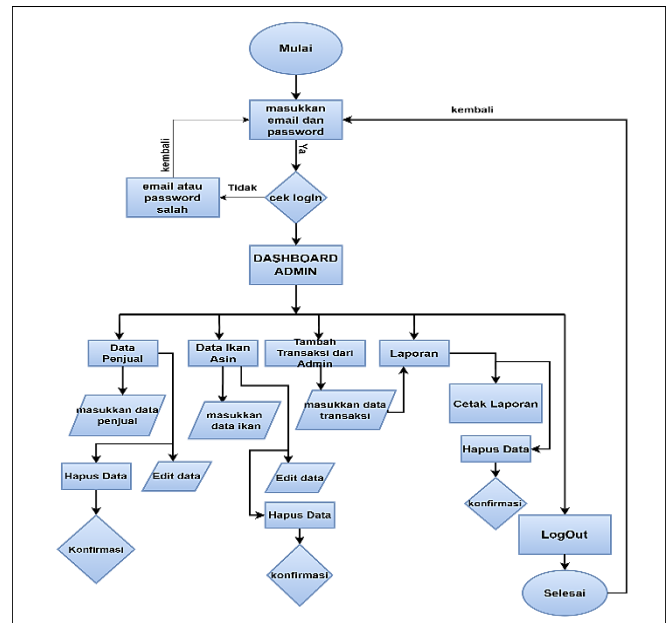
Dalam prosesnya, alur dari sistem terbagi menjadi 2 yaitu alur berjalannya sistem pada pengguna dan juga admin yang dapat dilihat pada gambar 2 dan 3 berikut:



Gambar 2. Alur Sistem pada user

gambar diatas menjelaskan tentang bagaimana proses dari sistem yang dibangun, dimana sebelum mengakses website pengguna terlebih dahulu melakukan proses login menggunakan akun yang telah didaftarkan, kemudian sistem memverifikasi apakah akun yang dimasukkan kedalam sistem telah sesuai jika tidak akan dikembalikan kehalaman login untuk memasukkan kembali akun yang sesuai, jika berhasil user akan diarahkan kehalaman awal dari sistem yang dimana didalamnya user dapat mengakses menu jenis ikan, dapat melihat daftar penjual, melakukan transaksi dimenu chart dan dapat melihat persyaratan yang harus dipenuhi dalam mengakses website penjualan ini pada menu about us dan

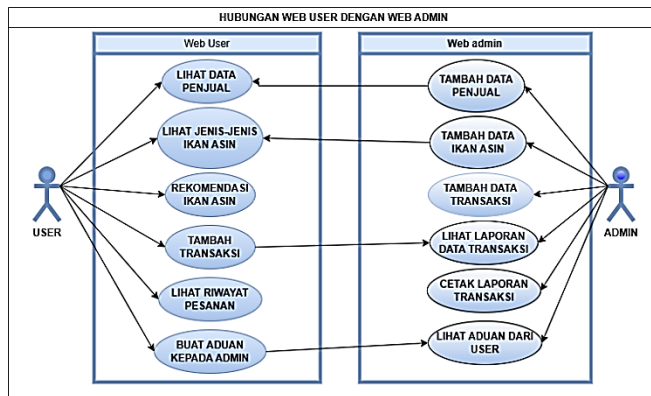
juga di dalamnya terdapat kontak admin yang berfungsi sebagai tempat pengaduan. Dalam menu jenis ikan terdapat rekomendasi dari sistem berdasarkan metode *collaborative filtering* yang merupakan tujuan pada penelitian ini.



Gambar 3. Alur Sistem Pada Admin

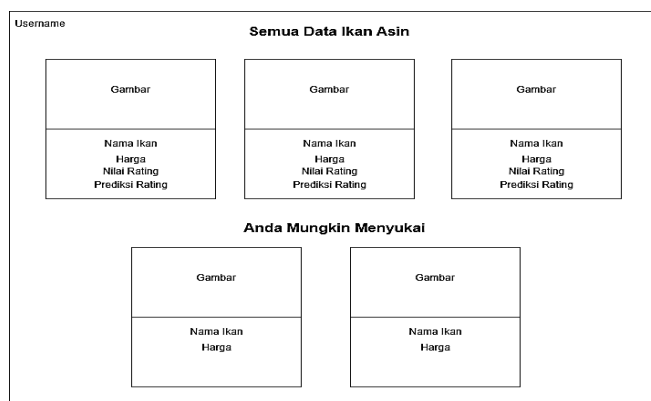
Berdasarkan gambar 3, setelah admin terverifikasi pada proses login admin dapat mengakses menu-menu yang diantaranya mengakses menu data ikan yang dimana admin dapat mengedit dan menambahkan data ikan serta dapat menghapus jenis-jenis ikan tertentu. Kemudian admin dapat menambahkan data penjual pada menu data penjual serta mengakses menu laporan transaksi yang didalamnya tersedia menu cetak laporan dan juga hapus transaksi.

Setelah merancang alur sistem dari pengguna dan admin, selanjutnya masuk kedalam proses hubungan antara pengguna dan juga admin yang digambarkan melalui diagram use case dan juga desain antar muka dari sistem yang akan dibuat meliputi desain menu tampilan jenis ikan beserta rekomendasinya di menu jenis ikan pada halaman pengguna yang dapat dilihat pada gambar 4.



Gambar 4. Use Case Diagram

Gambar 4 dapat dijelaskan hubungan yang terjadi antara pengguna dan juga admin, dimana user dapat melihat daftar penjual dan juga jenis-jenis ikan asin yang diinputkan oleh admin, admin dapat melihat data transaksi yang dilakukan oleh user pada menu laporan dan mencetak laporan tersebut sebagai bukti transaksi yang dilakukan oleh user yang nantinya laporan tersebut dapat dikirimkan kepada penjual yang bersangkutan, lalu bisa mendapatkan aduan tentang sistem dari user.



Gambar 5: Tampilan Halaman Jenis Ikan dan Rekomendasinya

C. Implementasi

Pada tahapan implementasi, penulis akan membuat kode program dan pengembangan sistem rekomendasi *collaborative filtering* dalam merekomendasikan ikan asin kepada pengguna dengan menghitung nilai similaritas antar pengguna menggunakan metode *pearson correlation* yang dimana cocok untuk rating *eksplisit*, dimana dapat memberikan korelasi linear yang akurat. Akan tetapi sensitif terhadap *sparsity* sehingga penulis menerapkan *shrinkage* dengan *threshold* Prosesorasi >0.1 dan normalisasi *mean-centering* untuk mengurangi bias pengguna aktif. pernghitungan nilai prediksi menggunakan metode *weigted sum* dimana kesederhanaannya dalam menghitung prediksi berbobot, penulis bandingkan dengan baseline model (rata-rata global) dalam evaluasi. Pada tahapan implementasi ini penulis merancang arsitektur dimana sistem dibangun dengan

pendekatan *client-server*, di mana *server-side* menangani perhitungan algoritma dan penyimpanan data, sedangkan pada *client-side* mengelola antarmuka pengguna.

1) Server-Side

Menggunakan PHP untuk menghitung *collaborative filtering real-time*, termasuk pembentukan matriks kesamaan *pearson* dan prediksi rating. Data disimpan kedalam database MySQL, yang diakses menggunakan *query SQL* untuk mengambil data rating dan juga data item. Perhitungan dilakukan *on-demand* saat pengguna meminta rekomendasi, sehingga memungkinkan *respons* cepat tanpa *batch processing*.

2) Client-Side

Antarmuka pengguna dibangun dengan HTML untuk struktur halaman, CSS untuk *styling*, dan *JavaScript* untuk interaktivitasnya misalnya *me-refresh* rekomendasi populer setiap 1 menit. Halaman *web* menampilkan daftar ikan asin, rekomendasi *collaborative filtering*, dan fitur *login/admin*.

3) Skalabilitas

Sistem dirancang untuk mendukung hingga 100 pengguna aktif dengan lantency <2 detik per rekomendasi, tergantung pada ukuran dataset (50 pengguna, 12 item). Untuk peningkatan, *caching* sederhana dapat ditambahkan menggunakan *session PHP* atau ekstensi seperti *redis* jumlah pengguna bertambah. Arsitektur ini memungkinkan sistem dapat diakses dari berbagai perangkat tanpa aplikasi *native*.

Pengembangan dilakukan menggunakan *Visual Studio Code* sebagai editor kode, yang mendukung integrasi HTML, *JavaScript*, CSS, dan PHP untuk mengembangkan *full-stack web*.

Pada dasarnya proses implementasi ini dibagi menjadi:

- 1) Pengumpulan data: Data yang dikumpulkan disini adalah data rating dari 50 pengguna terhadap 12 jenis ikan asin.
- 2) Menganalisis data: metode *collaborative filtering* yang digunakan dalam penelitian ini berfokus pada *user-based collaborative filtering* yang dimana mencari pengguna lain yang memiliki preferensi serupa dan merekomendasikan ikan asin yang mereka sukai [18]

Algoritma yang digunakan pada penelitian ini ialah algoritma *Pearson Correlation* yang digunakan untuk mengukur hubungan antara dua pengguna untuk menentukan seberapa mirip mereka kemudian setelah didapatkan nilai kemiripannya maka nilai yang mendekati 1 menandakan adanya hubungan yang positif dan nilai -1 menunjukkan terdapat hubungan yang negatif dimana ketika user 1 memberikan nilai tinggi maka yang terjadi pada user 2 justru kebalikannya, sedangkan angka 0 menunjukan tidak ada kemiripan antara kedua user [19]. Rumus yang digunakan dalam perhitungan *pearson correlation* adalah:

$$r = \frac{\sum (xi - x')(yi - y')}{\sqrt{\sum (xi - x')^2 \sum (yi - y')^2}} \quad [1]$$

Dimana:

r = pearson correlation

xi dan yi adalah nilai individu dari X dan Y

x' dan y' adalah nilai rata-rata dari X dan Y

Setelah mendapatkan nilai kemiripan, sistem akan memberikan nilai rekomendasi kepada target berdasarkan nilai rating yang diberikan oleh pengguna yang memiliki tingkat kesamaan yang tinggi dengan target. Kemudian penulis akan menghitung nilai prediksi ratingnya menggunakan metode *weighted sum* dengan tujuan agar pemberian rekomendasi kepada target bisa lebih akurat dengan menggunakan rumus:

$$r'u, i = \sum [v \in N(u)] (rv, i \times \frac{sim(u, v)}{\sum [v \in N(u)] |sim(u, v)|}) \quad [2]$$

Dimana:

$r'u, i$ = prediksi rating pengguna u dari item i

rv, i = rating dari pengguna v terhadap item i

$sim(u, v)$ = bobot kesamaan pengguna u dan v

$N(u)$ = himpunan kesamaan dari pengguna

dengan pengguna u

D. Pengujian

Tahapan ini melibatkan pengujian sistem rekomendasi untuk memastikan bahwa sistem berfungsi dengan benar dan sesuai dengan kebutuhan, termasuk pengujian fungsional dan pengujian performa. Pada pengujiannya, performa sistem diuji menggunakan metode *blackbox* yang befokus pada kinerja sistem apakah masih terdapat bug ataupun error yang terjadi sehingga sistem dapat bekerja dengan baik. Dalam kasus menguji tingkat keakuratan nilai prediksi dan kinerja sistem, penulis disini menggunakan matrik *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE) untuk mengetahui tingkat perbandingan keakuratannya dimana evaluasi algoritmanya menggunakan *train-test split* 80-20 dimana 80% data untuk *training* dan 20% data untuk *testing*, diulang 5 kali untuk rata-rata. Penggunaan strategi popularitas dievaluasi dengan simulasi *cold start* dimana MAE pengguna baru 0,75 dibandingkan dengan 0,65 *non-cold start* menandakan adanya peningkatan 15% akurasi. Deversitas dengan *score* 0,4 rendah, karena berfokus pada item populer akan tetapi cukup untuk *angepement* awal.

E. Pemeliharaan

Pada tahapan ini, penulis akan melakukan pemeliharaan sistem yang telah dibuat untuk memastikan bahwa sistem dapat beradaptasi dengan perubahan kebutuhan yang dimana kasusnya telah tergabung didalam langkah pengujian sistem.

III. RESULTS AND DISCUSSION

Pada bagian ini penulis akan menjelaskan cara kerja serta hasil yang diperoleh dari proses implementasi *collaborative filtering* dalam merekomendasikan ikan asin kepada pengguna berdasarkan data yang telah penulis kumpulkan dimana jenis-jenis ikan asinnya dilambang dengan huruf yaitu IA sampai dengan IK dan pengguna dilambangkan dengan P1 sampai dengan P 50 seperti tabel 1 berikut:

TABEL I
DATA RATING, SERTA NILAI RATA-RATA RATINGNYA

user	I A	I B	I C	I D	I E	I F	I G	I H	I I	I J	I K	Rata-rata rating
P1	4	5	5	5	5	5	5	5	5	5	5	4,90
P2	5	5	5	3	5	3	5	5	5	2	2	4,09
P3	3	5	5	5	5	5	5	5	5	5	4	4,72
P4	3	4	5	3	4	0	3	5	4	0	0	2,81
P5	4	3	3	4	4	4	5	4	4	4	3	3,81
P6	3	1	3	4	4	5	3	4	5	5	3	3,63
P7	5	5	1	5	5	5	1	5	5	1	3	3,72
P8	5	5	5	5	5	5	3	4	4	5	4	4,54
P9	0	5	5	5	5	5	5	5	5	5	5	4,54
P10	3	4	2	2	4	2	5	4	5	2	2	4,81
P11	0	4	0	4	0	0	0	0	0	0	0	0,72
P12	3	2	5	5	4	3	3	4	5	2	2	3,45
P13	5	5	0	3	4	0	5	3	0	0	0	2,27
P14	5	4	5	3	5	5	5	5	5	5	3	4,54
P15	5	5	5	4	4	5	3	5	4	3	4	4,27
P16	0	5	5	5	5	5	5	5	5	0	0	3,63
P17	5	3	4	5	5	2	5	3	3	2	1	3,45
Pn=50	3	0	2	1	4	3	5	5	0	3	0	2,63

Setelah memiliki data rating, selanjutnya masuk kedalam proses perhitungan nilai kemiripan untuk setiap pengguna guna mengetahui korelasi atau hubungan antara setiap pasang pengguna yang dapat dilihat pada perhitungan dibawah ini:

$$r = \frac{\sum (xi - x')(yi - y')}{\sqrt{\sum (xi - x')^2 \sum (yi - y')^2}} \quad [3]$$

Data rating p5 (x) dan p6 (y):

P5 (x): 4, 3, 3, 4, 4, 4, 5, 4, 4, 4, 3

P6 (y): 3, 1, 3, 4, 4, 5, 3, 4, 5, 5, 3

Rata-rata x' dan $y' = x' = 3,81818$, $y' = 3,63636$

Setelah mengetahui nilai dari x' dan nilai y' selanjutnya menghitung nilai deviasi dari rata-rata dan kuadratnya dimana indikatornya merupakan nilai rating dari Ikan A sampai dengan ikan K:

xi (P5 IA-IK): 4, 3, 3, 4, 4, 4, 5, 4, 4, 4, 3

$xi - x' = 4 - 3,8182 = 0,1818$

$xi - x' = 3 - 3,8182 = -0,8182$

$xi - x' = 5 - 3,8182 = 1,1818$

yi (P6 IA-IK): 3, 1, 3, 4, 4, 5, 3, 4, 5, 5, 3

$yi - y' = 3 - 3,6364 = -0,6364$

$yi - y' = 1 - 3,6364 = -2,6364$

$$yi - y' = 4 - 3,6364 = 0,3636$$

$$yi - y' = 5 - 3,6364 = 1,3636$$

Maka nilai dari P5 IA-IK ($xi - x'$) dan P6 IA-IK ($yi - y'$) dapat dilihat pada tabel dibawah ini:

TABEL II
NILAI $xi - x'$ DAN $yi - y'$

Ikan(i)	$xi - x'$	$yi - y'$	$(xi - x') - (yi - y')$	$(xi - x')^2$	$(yi - y')^2$
IA	0,1818	-0,6364	-0,1157	0,0330	0,4050
IB	-0,8182	-2,6364	2,1587	0,6694	6,9504
IC	-0,8182	-0,6364	0,5204	0,6694	0,4050
ID	0,1818	0,3636	0,0661	0,0330	0,1322
IE	0,1818	0,3636	0,0661	0,0330	0,1322
IF	0,1818	1,3636	0,2479	0,0330	1,8595
IG	1,1818	-0,6364	-0,7521	1,3967	0,4050
IH	0,1818	0,3636	0,0661	0,0330	0,1322
II	0,1818	1,3636	0,2479	0,0330	1,8595
IJ	0,1818	1,3636	0,2479	0,0330	1,8595
IK	-0,8182	-0,6364	0,5204	0,6694	0,4050

Setelah mendapatkan nilai-nilainya kemudian hitung korelasinya dimana:

$$\sum(xi - x')(yi - y') = 3,2547$$

$$\sum(xi - x')^2 = 3,6365$$

$$\sum(yi - y')^2 = 14,5455$$

$$\sqrt{3,6365 \times 14,5455} = \sqrt{52,9234} \approx 7,2748$$

$$r = \frac{3,2547}{7,2748} \approx 0,4474 \quad [4]$$

Hasil $r = 0,4474$ menunjukkan adanya hubungan yang positif moderat antara pola rating dari pengguna 5 dan juga pengguna 6 yang berarti pengguna 5 dan pengguna 6 cenderung memiliki prefensi rating yang cukup mirip, jika pengguna 5 memberikan rating tinggi pada suatu ikan maka pengguna 6 juga memberikan nilai rating yang tinggi pada item yang sama. Untuk korelasi antar setiap user dapat dilihat pada tabel berikut:

TABEL III
KORELASI ANTAR TIAP PENGGUNA

pengguna	P1	P2	P3	P4	Pn=50
P1	1,000	-0,100	0,456	0,218	-0,128
P2		1,000	0,456	-0,038	0,332
P3			1,000	0,370	0,456
P4				1,000	0,218
P5	0,510	0,088	0,528	0,510	0,510
P6	0,245	0,076	0,484	0,435	0,245
P7	-0,088	0,449	0,176	0,218	-0,088
P8	0,414	0,240	0,686	0,602	0,414
P9	0,508	0,287	0,640	0,638	0,508
P10	-0,342	0,301	0,247	0,281	-0,342
P11	0,279	-0,140	0,106	0,605	0,279

P12	0,000	0,575	0,490	0,518	0,000
P13	0,065	0,165	0,289	0,276	0,065
P14	0,468	0,093	0,603	0,449	0,468
P15	0,316	0,282	0,689	0,484	0,316
P16	0,485	-0,004	0,640	0,709	0,485
P17	-0,117	0,032	0,237	0,252	-0,117
Pn=50	-0,128	0,332	0,456	0,218	1,000

Setelah mendapatkan nilai korelasi antar tiap user selanjutnya menghitung nilai prediksi rating untuk ikan yang belum dinilai, dalam hal ini kita mengambil contoh pada P11 yang hanya memberikan rating pada IB dan ID.

Hitung rata-rata rating aktual dimana:

$$R'P11 = \frac{(4+4)}{2} = 4,0 \quad [5]$$

Temukan pengguna tetangga (N) untuk setiap prediksi yang memiliki kriteria korelasi positif dengan target misalnya korelasi $>0,1$ atau bahkan >0 serta pengguna telah memberikan rating (1-5) untuk item yang akan diprediksi. Masukkan kedalam formula *Weighted Sum*:

$$Pu,i = R' + \sum_{v \in N(Rv,i - R'v)} \times \frac{sim(u,v)}{\sum v} \in N|sim(u,v)| \quad [6]$$

Rv,i rating aktual yang diberikan user tetangga v untuk item i

$R'v$ rata-rata rating aktual yang diberikan oleh user tetangga $sim(u,v)$ koefisien korelasi antara user u dan user v

$\sum v \in N|sim(u,v)|$ jumlah nilai absolute koefisien $v \in N$

Disini akan dihitung prediksi rating pada pengguna P11 untuk item IA dimana rating aktualnya IB(4), ID(4) dan rata-rata rating aktualnya adalah 4,0. Setelah itu identifikasi user tetangga (N) yang memiliki nilai korelasi positif dengan P11 dan telah memiliki rating non-nol untuk item yang akan diprediksi, dapat dilihat pada tabel 4 berikut:

TABEL IV
USER TETANGGA (v) P11 ITEM A

Tetangga(v)	Rv,IA	$R'v$	$sim(P11,v)$
P1	4	4,90	0,279
P3	3	4,72	0,106
P4	3	3,88	0,605
P5	4	3,81	0,316
P6	3	3,63	0,222
P8	5	4,54	0,177
P12	4	3,45	0,333
P14	5	4,54	0,379
P15	5	4,27	0,298
P17	5	3,45	0,274

Selanjutnya menghitung pembilang dan juga penyebut berdasarkan rumus *weight sum* $\sum(Rv,IA - R'v) \times sim(P11,v)$ dan $\sum sim(P11,v)$

$$P1: (4 - 4,90) \times 0,279 = -0,2511$$

$$P3: (3 - 4,72) \times 0,106 = -0,18232$$

Untuk hasil keseluruhan dari pembilang dan penyebut dapat dilihat pada tabel berikut:

TABEL V
PEMBILANG DAN PENYEBUT P11, v

(v)	pembilang	Jumlah pembilang	Jumlah penyebut
P1	-0,2511	-0,39757	2,989
P3	-0,18232		
P4	-0,5324		
P5	0,06004		
P6	-0,13986		
P8	0,08142		
P12	-0,14985		
P14	0,17434		
P15	0,21754		
P17	0,4247		

sehingga perhitungan prediksinya menjadi:

$$PP11, IA = R'P11 + \text{Jumlah Pembilang} / \text{Jumlah Penyebut}$$

$$PP11, IA = 4,0 + (-0,39757)/2,989$$

$$PP11, IA = 4,0 - 0,133$$

$$PP11, IA \approx 3,867 \quad [7]$$

Jadi, nilai prediksi rating P11 terhadap item IA adalah sekitar 3,87. Sehingga semua pengguna yang belum memberikan nilai rating (1-5) pada suatu item dapat dilihat nilai prediksi ratingnya pada tabel berikut:

TABEL VI
NILAI PREDIKSI RATING PADA PENGGUNA (4,9,11,13,18) UNTUK ITEM YANG BELUM DINILAI

Pengguna	Prediksi Rating
P4	IF = 4,03, IJ = 3,56, IK = 3,13
P9	IA = 4,99
P11	IA = 3,87, IC = 4,34, IE = 4,23, IF = 4,05, IG = 3,95, IH = 4,09, II = 4,22, IJ = 3,30, IK = 3,33
P13	IC = 4,26, IF = 4,26, II = 4,26, IJ = 4,26, IK = 4,11
P18	IA = 4,98, IC = 4,30, IE = 4,23, IF = 3,40, IG = 3,84, IH = 4,07, II = 4,10, IJ = 3,12, IK = 2,67

Setelah semua nilai prediksi didapatkan maka langkah selanjutnya adalah mengevaluasi metode yang digunakan untuk mengukur kinerja dari sistem yang dimana menggunakan dua metrik evaluasi yaitu *Mean Absolute Error* (MAE) dan *Root Mean Absolute Error* (RMSE). Total data set yang terdiri dari 600 data rating dari 50 pengguna dibagi menjadi:

A. Pembagian Data

Dataset rating ikan yang diperoleh dari 50 pengguna akan dibagi menjadi dua subset yaitu training set sebanyak 80% dari jumlah dataset (600 data rating) yang akan digunakan dalam model pelatihan. Data ini digunakan dalam proses perhitungan rata-rata rating dan matriks similaritas antar pengguna menggunakan *pearson correlation*. Tes set sebanyak 20% dari total dataset (50 data rating) akan disisihkan sebagai data uji. Rating aktual atau nyata dari

pengguna yang ada dalam tes set ini akan digunakan sebagai *ground truth* guna mengevaluasi akurasi prediksi. Pembagian data ini dilakukan secara acak untuk merepresentasikan interkasi dari seluruh pengguna dengan item secara merata.

Selanjutnya proses pelatihan model *collaborative filtering* yang dimana prosesnya meliputi perhitungan rata-rata rating dari tiap pengguna dihitung berdasarkan rating yang diberikan pada item dalam training set, kemudian menghitung similaritas antar pengguna yang diukur menggunakan koefisien korelasi *pearson* berdasarkan item yang telah dinilai bersama oleh kedua pengguna dalam training set, setelah itu memprediksi rating menggunakan formula *Weighted Sum*.

B. Metrik Evaluasi

Kinerja dari sistem diukur menggunakan dua metrik umum yaitu *Mean Absolute Error* (MAE) yang digunakan untuk mengukur rata-rata besarnya kesalahan antar rating prediksi dan rating aktual yang dihitung menggunakan formula:

$$MAE = 1/N_{test}(u,i) \in \text{TestSet} \times \sum |R_{u,i} - R'_{u,i}| \quad [8]$$

dimana N_{test} adalah jumlah total rating dalam test set (36), $R_{u,i}$ adalah rating aktual, dan $R'_{u,i}$ adalah rating prediksi. Nilai MAE yang lebih rendah menunjukkan akurasi prediksi yang lebih baik. Kemudian *Root Mean Squared Error* (RMSE) yang digunakan dalam mengukur akar dari rata-rata kuadrat kesalahan antar rating prediksi dan rating aktual. Proses perhitungan ini menggunakan formula:

$$RMSE = \sqrt{\frac{1}{N_{test}(u,i) \in \text{TestSet}} \times \sum (R_{u,i} - R'_{u,i})^2} \quad [9]$$

RMSE memberikan bobot yang lebih besar pada kesalahan prediksi yang besar dibandingkan MAE.

Setelah melakukan perhitungan menggunakan kedua formula diatas didapatkan hasil berupa Nilai MAE $\approx 0,63$ dan RMSE $\approx 0,81$. Nilai MAE diatas mengindikasikan bahwa rata-rata kesalahan rating adalah 0,63 point pada skala rating 1-5 berarti error rata-rata 12,6%, lebih baik dari baseline popularity-based (MAE=1.2) dan studi banik (0,8)[5]. Sementara itu, nilai RMSE sebesar 0,81 point yang sedikit lebih tinggi dari MAE, menunjukkan bahwa terdapat variasi dalam besarnya kesalahan prediksi, dengan adanya beberapa prediksi yang memiliki selisih lebih besar dari rata-rata dengan tingkat akurasi 70-80%.

Setelah itu, penulis akan menjelaskan proses pemberian rekomendasi kepada *user* dengan langkah sebagai berikut.

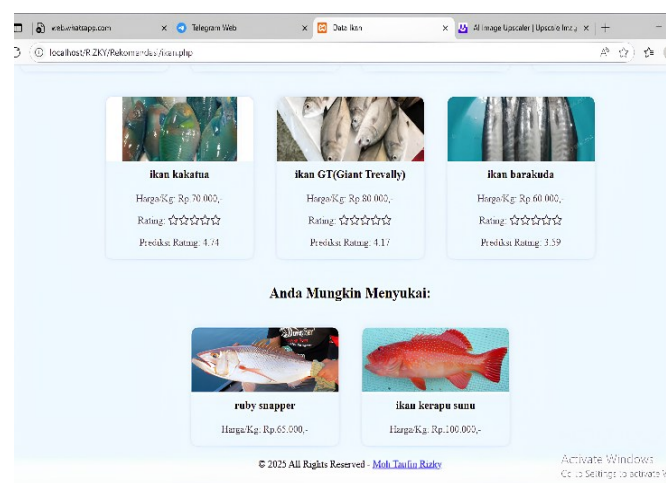
A. Proses Pemberian Rekomendasi

Rekomendasi diberikan dengan mengambil semua data rating dari database melalui query SQL yang kemudian akan disimpan dalam array, setelah itu sistem akan menghitung nilai rata-rata rating (seperti pada tabel 1) yang digunakan sebagai *baseline* dalam menghitung kesamaan dan juga prediksi. Selanjutnya sistem membentuk *matrix* kesamaan

antar pengguna menggunakan fungsi *pearson correlation* (rumus [1]) yang kemudian disimpan dalam *matrix* kesamaan (tabel III) setelah itu sistem akan mengidentifikasi item yang belum dirating oleh pengguna dimana setiap item *unrated* sistem akan memprediksi nilai rating menggunakan metode *weighted sum* (rumus[2]) yang melibatkan pemilihan k tetangga terdekat, perhitungan bobot berdasarkan kesamaan dan deviasi dari rata-rata tetangga. Pemilihan *Top-N* rekomendasi dimana prediksi rating diurutkan dari yang tertinggi ke yang terendah (*descending*), jika pengguna belum memberikan nilai rating apapun (*cold start*) maka akan otomatis dialihkan ke sistem rekomendasi populer berdasarkan rata-rata rating tertinggi dari semua pengguna yang ada.

B. Rekomendasi Populer

Rekomendasi populer digunakan dalam mengatasi masalah pengguna yang baru mendaftar kedalam sistem dimana belum memberikan nilai rating minimal pada 1 jenis ikan (*Cold start Problem*) sehingga sistem belum bisa membaca kesamaan antar pengguna tersebut dengan pengguna lainnya. Dalam hal ini rekomendasi populer bekerja sebagai solusi agar sistem bisa tetap memberikan rekomendasi ikan kepada pengguna baru berdasarkan jenis ikan yang populer.



Gambar 6. Tampilan Rekomendasi Pengguna 11

Dalam hal ini sistem akan memilih 10 ikan yang populer berdasarkan hasil pemberian rating dari pengguna-pengguna sebelumnya dan kemudian akan menampilkan 2 jenis ikan kehalaman pengguna dan jenis-jenis ikan yang direkomendasikan akan berganti dalam interval waktu 1 menit setiap pengguna merefresh halaman jenis ikan. Sehingga setelah mengikuti semua langkah-langkah pemberian rekomendasi, didapatkan rekomendasi ikan asin untuk pengguna 11.

C. Analisis Error

Analisis *error* menunjukkan item IK sering salah direkomendasikan (*error*>1.0 untuk 20% pengguna) karena sparsity. Pengguna P11 memiliki *error* tinggi (MAE=0.8) akibat kurang tetangga. Bias pengguna baru sering mendapatkan rekomendasi populer, mengurangi diversitas tapi cukup dalam engagement seperti tabel berikut ini.

TABEL VI
ANALISIS ERROR PER ITEM

Item	Rata-rata MAE	Jumlah kesalahan >1.0	Error tinggi (%)	Keterangan
IA	0.45	5	10%	Item populer, error rendah
IB	0.52	7	14%	Sedang, variasi regional
IC	0.48	6	12%	Rendah, data cukup dens
.....
IK	0.78	15	30%	Tinggi, sering salah direkomendasikan karena niche

TABEL VII
ANALISIS ERROR PER PENGGUNA

User	MAE	RMSE	Jumlah Item Salah Direkomendasikan	Keterangan Error
P1	0,40	0,55	2	Rendah, tetangga banyak
P11	0,80	1,10	8	Tinggi, cold start parsial
P13	0,75	1,05	7	Outlier, data sparse
.....
....
Rata-rata	0,63	0,81	4	Overall

TABEL VIII
ANALISIS BIAS DAN DEVERSITAS

Keteg ori	Top 3 Rekomendasi	Score	Jumlah User	Keterangan
Cold start	IA,IB,IC	0.35	10	Fokus pada populer, deversitas rendah
lama	IA,ID,IG	0.65	40	Lebih variatif, berdasarkan <i>collaborative filtering</i>
overall 1	IA,IB,IC	0.50	50	Bias ke item populer

IV. CONCLUSION

Implementasi *Collaborative Filtering* dalam hal ini berbasis pengguna dalam merekomendasikan ikan asin menunjukkan potensi yang cukup baik dalam memprediksi preferensi pengguna terhadap jenis-jenis ikan yang terdapat dalam sistem. Melalui analisis pola rating yang diberikan oleh pengguna yang memiliki preferensi sama (tetangga), sistem dapat memberikan rekomendasi yang diprediksi relevan dengan pengguna target. Berdasarkan hasil evaluasi menggunakan metrik MAE dan RMSE yang dimana nilai MAE $\approx 0,63$ mengindikasikan bahwa rata-rata kesalahan prediksi rating adalah 0,63 poin pada skala rating 1-5. Kesalahan rata-rata yang relatif rendah ini menunjukkan bahwa model secara umum mampu memprediksi rating yang cukup akurat untuk pengguna. Nilai RMSE $\approx 0,81$ dimana nilai ini sedikit lebih besar dari nilai MAE menunjukkan adanya variasi dalam besarnya kesalahan prediksi. Meskipun demikian nilai RMSE yang tidak jauh berbeda dari nilai MAE mengimplikasikan kesalahan-kesalahan besar tidak terlalu ekstrem dan juga model cenderung konsisten dalam memberikan prediksi yang mendekati rating sebenarnya. Saran yang bisa diberikan adalah mengkombinasikan *Collaborative Filtering* dengan metode rekomendasi lainnya sehingga dapat meningkatkan kinerja sistem.

REFERENCES

- [1] Aisha, Dita, and Ririen Kusumawati. "Implementasi metode algoritma collaborative filtering & k-nearest neighbor pada sistem rekomendasi e-commerce." *Jurnal Ilmiah Sistem Informasi dan Ilmu Komputer* 2.3 (2022): 25-38.
- [2] Xu, Guandong, Yanchun Zhang, and Lin Li. *Web mining and social networking: techniques and applications*. Vol. 6. Springer Science & Business Media, 2010.
- [3] Laksana, Eka Angga. "Collaborative Filtering dan Aplikasinya." *Jurnal Ilmiah Teknologi Infomasi Terapan* 1.1 (2014).
- [4] Yoshua I., Bunyamin H., Si S. (2021). Pengimplementasian Sistem Rekomendasi Musik Dengan Metode Collaborative Filtering. *Jurnal strategi*.
- [5] Rounak Banik, *Hands-On Recommendation Systems with Python*, 1st ed., Pravin Dhadre, Ed. Brimingham, England: Packt Publishing, 2018.J. S.
- [6] Breese, John S., David Heckerman, and Carl Kadie. "Empirical analysis of predictive algorithms for collaborative filtering." *arXiv preprint arXiv:1301.7363* (2013).
- [7] Nugroho F., Ismu Rahayu M. (2020). Sistem Rekomendasi Produk UKM Di Kota Bandung Menggunakan Algoritma Collaborative Filtering. *Jurnal Riset Sistem Informasi dan Teknologi Informasi (JURSISTENI)*, 10.52005/jursistekni.v2i3.63.
- [8] Nurhani F., Samsudin. (2022). Implementasi Algoritma Collaborative Filtering pada Sistem Pemesanan Makanan dan Minuman dengan platform Android. *Jurnal Ilmiah Komputasi*, 10.32409/jikstik.21.3.3110
- [9] Uyun, Shofwatul, Imam Fahrurrozi, and Agus Mulyanto. "Item Collaborative Filtering untuk Rekomendasi Pembelian Buku secara Online." *Jurnal Sistem Informasi Indonesia* 1.1 (2011): 63-70.
- [10] Susanto Hardo. (2014) *Perancangan Sistem Rekomendasi Pakaian Distro Dengan Menggunakan Item Collaborative Filtering (Studi Kasus: The Jungle Distro Medan)*. Medan.
- [11] Royce, W. W. (1970). Managing the Development of large Software Systems. *Proceedings of IEEE WESCON*, 1-9.
- [12] Sasmito, G. W., 2017. "Penerapan Metode Waterfall Pada Desain Sistem Informasi Geografis Industri Tegal", pp. 4-6
- [13] Suroso p. a. (2018) Sistem Pendukung Keputusan Penanganan Kejadian Abortus dengan Metode Ahp (Analytical Hierarchy Process). *Jurnal.Stmikelrahma*.
- [14] Iskandar, J., Aman, M., Sasono, I., Riyanto, R., Wiyono, N., Suroso, S., & Yanto, A. (2024). Penerapan metode Topsis Pada Sistem Pendukung Keputusan Seleksi beasiswa Peningkatkan Prestasi Akademik Dengan Pendekatan Oop.
- [15] Laudon, Keneth C. dan Jane P. Laudon. (2007). *Management Information Systems: Managing The Digital*, International Jounarl Of Computers, Communications & Control (IJCCC). 2(1):103, 10.15837/ijccc.2007.1.2342.
- [16] O'Brien, J. A., & Marakas, G. M. (2011). *Management Information Systems*. McGraw-Hill.
- [17] Anggono a., Emanuel A. W. R. (2014) Sistem Rekomendasi pada Portal Lowongan Kerja Menggunakan Metode Simple Additive Weighting. *Jurnal Informatika*.
- [18] Theodorus, Arvid. *User-Based Collaborative Filtering Dengan Memanfaatkan Pearson-Correlation Untuk Mencari Neighbors Terdekat Dalam Sistem Rekomendasi*. Diss. UAJY, 2016.
- [19] Robbani, Mufti, Rima Dias Ramadhani, and Andika Elok Amalia. "Analisa Algoritma Cosine Similarity dengan Pearson Correlation pada Metode Item-based Collaborative Filtering dengan Menggunakan Dataset Movielens." *Conference on Electrical Engineering, Telematics, Industrial technology, and Creative Media (CENTIVE)*. 2018.