

Perbandingan Metode *Demographic* dan *Content Based Filtering* Pada Sistem Rekomendasi *Smartphone* Android

Fauzi Cahya Miftahul Sidiq¹, Yuliant Sibaroni²

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹fauzicahya@student.telkomuniversity.ac.id, ²yuliant@telkomuniversity.ac.id

Abstrak

Smartphone sudah menjadi kebutuhan hidup manusia jaman sekarang, terutama di era digital sekarang sudah terlalu banyak alternatif pilihan *smartphone* android dengan banyaknya spesifikasi yang ada sehingga menyebabkan calon pembeli *smartphone* android kebingungan dalam menentukan tipe yang sebaiknya dipilih. Calon pembeli bebas memilih *smartphone* android apa saja, mendapatkan hasil yang maksimal dari apa yang dikeluarkan. Maka dari itu penelitian ini membangun sebuah sistem rekomendasi dengan membandingkan dua metode untuk membantu proses pemilihan *smartphone* android bagi calon pembeli. Metode *demographic filtering* dapat memberikan sebuah rekomendasi dengan landasan kemiripan profil calon pembeli. Sedangkan metode *content-based filtering* dapat memberikan sebuah rekomendasi dengan landasan kemiripan spesifikasi *smartphone* android yang dibutuhkan calon pembeli dengan data-data yang disediakan pada database *smartphone*. Hasil pengujian dalam penelitian ini menghasilkan kesimpulan bahwa *content-based filtering* mendapatkan hasil lebih baik daripada *demographic filtering*.

Kata Kunci : Sistem Rekomendasi, *Demographic Filtering*, *Content Based Filtering*, *Smartphone* Android.

Abstract

Smartphones have become a necessity for human life today, especially in the digital era, now there are too many alternative choices for Android smartphones with a lot of specifications, so it's causing potential buyers to be confused in determining which type of android smartphone to choose. Prospective buyers are free to choose any android smartphone, get the most out of what is issued. Therefore this research builds a recommendation system by comparing two methods to help the process of selecting an Android smartphone for potential buyers. Demographic filtering methods can provide a recommendations based on similar profiles of prospective buyers. While the content-based filtering method can provide a recommendations based on the similarity of the Android smartphone specifications desired by prospective buyers with the data provided in the smartphone database. Based on evaluation on this case, content-based filtering gets better recommendation than demographic filtering.

Keywords : System Recommendation, *Demographic Filtering*, *Content Based Filtering*, *Smartphone* Android.

1. Pendahuluan

1.1. Latar Belakang

Teknologi yang begitu pesat sangat mempengaruhi cara untuk mendapatkan informasi yang kini kian mudah lewat internet. *Smartphone* merupakan salah satu alat untuk mengaksesnya. Dari 276,4 juta populasi penduduk di Indonesia, tercatat ada 353,8 juta ponsel yang terkoneksi [1]. Selain sebagai media untuk mendapatkan informasi, *smartphone* juga bisa berfungsi sebagai alat komunikasi, edukasi, atau bahkan hanya sekedar untuk gaya hidup.

Dengan maraknya pengguna *smartphone* pada saat ini, membuat semakin banyak juga merk-merk *smartphone* yang bermunculan. Akibatnya banyak orang yang kesulitan untuk menemukan *smartphone* yang sesuai dengan kebutuhannya. Spesifikasi *smartphone* yang seringkali sulit untuk dimengerti oleh orang awam, menjadi salah satu faktor dalam kesalahan pembelian *smartphone* yang pada akhirnya hanya memperhitungkan citra brandnya saja, padahal belum tentu sesuai dengan fitur-fitur yang mungkin diperlukan si pembeli.

Penelitian serupa telah digarap oleh Huda dengan membuat sebuah klasifikasi fitur konten berdasarkan karakteristik rasa sebagai dasar penentu rekomendasi pada suatu makanan [2]. Atribut yang digunakan dalam penelitian tersebut antara lain *bitter*, *sweet*, *meat*, *spicy*, *savory*, *sour*, *salty*, *sauce*, dan *vegetable* yang merupakan selera dari pengguna. Metode yang digunakan merupakan *pearson correlation* untuk melakukan perhitungan kesamaan fitur konten antar makanan. Penelitian lain juga yang dilangsungkan oleh Maharani dan Gunawan terkait membuat sebuah sistem rekomendasi dengan membandingkan dua metode [3]. Penelitian tersebut menunjukkan bahwa metode *demographic filtering* masih belum mendapatkan hasil

rekomendasi sebaik metode *content based filtering* dikarenakan atribut *demographic filtering* yang digunakan masih kurang relevan. Penelitian lainnya yang telah dilakukan oleh Titipat Achakulvisut dkk, dimana dalam penelitian tersebut dibuatkan rekomendasi artikel untuk referensi publikasi ilmiah dengan menggunakan sebuah metode yaitu *content based filtering* [4]. Penelitian lainnya telah digarap oleh Richard Firdaus Oeyliawan dkk, yang membuat sebuah aplikasi rekomendasi buku menggunakan *vector space model*, metode *content based filtering* digunakan untuk melakukan pendekatan, hasil penelitian membuahkan sebuah nilai presisi yang tercapai yaitu sebesar 46,5 [5].

Oleh karena itu, perbandingan metode *demographic* dan *content based filtering* dalam kasus rekomendasi *smartphone* android menarik untuk dilakukan. Penelitian ini menggunakan metode *demographic* dan *content based filtering* untuk mendapatkan rekomendasi *smartphone* android. Metode *demographic filtering* merupakan sebuah metode yang digunakan untuk memberikan rekomendasi kepada pengguna baru yang belum memiliki catatan transaksi dengan cara membandingkan nilai atribut atau data demografis pengguna tersebut dengan pengguna lain [3]. Sedangkan *content based filtering* merupakan metode yang merekomendasikan item dengan dasar korelasi antara atribut yang tersedia di database dengan preferensi pengguna sebagai pembandingan [6].

1.2. Topik dan Batasannya

Batasan masalah yang terdapat dalam penelitian tugas akhir ini adalah jenis *smartphone* yang digunakan yaitu *smartphone* android.

1.3. Tujuan

Tujuan yang ingin dicapai dari penelitian ini adalah mengimplementasikan metode *demographic* dan *content based filtering* pada sistem rekomendasi *smartphone* android dan juga mendapatkan hasil perbandingan rekomendasi *smartphone* android yang menggunakan metode *demographic* dan *content based filtering*.

2. Studi Terkait

Penelitian tentang rekomendasi mobil oleh Herastia dan Ferry pada tahun 2014 [3] bertujuan untuk memilih mobil yang sesuai dengan kebutuhan pengguna dikarenakan banyaknya alternatif pilihan mobil yang tersedia. Setiap orang memiliki prioritas yang berbeda dalam memilih sebuah mobil, sehingga bisa saja terjadi perbedaan kecocokan mobil di setiap individu. Penelitian ini membuat sebuah sistem rekomendasi dengan menggabungkan dua metode dengan tujuan membantu calon pembeli untuk memilih mobil. Metode *content based filtering* bertujuan untuk menghasilkan sebuah rekomendasi mobil dengan landasan kemiripan antar spesifikasi mobil yang diinginkan oleh calon pembeli dengan spesifikasi mobil yang tersedia. Sedangkan metode *demographic filtering* bertujuan untuk menghasilkan sebuah rekomendasi mobil dengan landasan kemiripan profil antar pembeli. Hasil pengujian pada penelitian ini menunjukkan bahwa *content based filtering* menghasilkan rekomendasi yang cenderung lebih relevan dibandingkan dengan hasil rekomendasi dari *demographic filtering*.

Pada penelitian yang dilakukan oleh Cut, Herastia dan Nathania tentang *Product Recommendation System* pada tahun 2019 [8] membahas metode yang tepat untuk organisasi yang dapat membantu mempromosikan produk yang paling sesuai berdasarkan karakteristik pelanggan mereka. Organisasi juga perlu menyarankan produk alternatif untuk pelanggan jika produk yang diminta tidak tersedia. Sistem rekomendasi yang diusulkan mengadopsi metode perhitungan *cosine similarity* yaitu skor kesamaan produk dan filter berbasis konten untuk dihitung skor rekomendasi pelanggan dan digunakan sebagai model untuk sistem yang diajukan. Berdasarkan hasil pengujian sistem yang diusulkan, dapat disimpulkan bahwa metode yang dipilih dapat digunakan untuk merekomendasikan produk dan pelanggan produk. Ini menunjukkan bahwa *precision* dan *recall* dari skor kesamaan produk dan rekomendasi pelanggan untuk skor produk adalah 100% dan 93,47%.

Penelitian tentang sistem rekomendasi yang dilakukan oleh Lukas, Chandra, dan Vivi [9] membangun sebuah sistem rekomendasi menggunakan metode *content based filtering* dan algoritma apriori untuk memesan produk. Alasan memilih metode *content based filtering* yaitu karena kelebihanannya yang dapat mengambil kesamaan karakteristik dari item yang disukai pengguna pada sebelumnya sehingga dapat menyarankan item baru ke pengguna lain. Dengan sistem yang dibangun, menjadikan pencarian dan pemesanan produk menjadi lebih mudah sehingga meningkatkan efisiensi transaksi sehingga penjualan jadi meningkat. Sistem rekomendasi yang dibuat mendapatkan hasil yang cukup akurat yaitu sebesar 73,33%.

Pada penelitian yang dilakukan oleh Anderias dan Deni tentang sistem rekomendasi laptop [10] berlandaskan laptop yang saat ini sangat diperlukan, ditambah beragamnya merk dan spesifikasi laptop. Sehingga dibuat sebuah sistem rekomendasi dengan teknik *mixed hybrid* yang merupakan penggabungan dua buah metode yaitu *collaborative filtering* dan *content based filtering* dengan tujuan untuk memberikan

rekomendasi laptop. Hasil yang didapatkan adalah metode *content based filtering* memiliki waktu eksekusi tercepat jika dibandingkan metode lainnya dalam penelitian ini.

Penelitian tentang sistem rekomendasi *content based filtering* yang dilakukan oleh Rhesa, Ardhi dan Winarno [11] tentang wisata kuliner yang sedang menjadi salah satu tren dunia kuliner saat ini. Penggunaan mesin pencari saja dinilai masih kurang cukup, sehingga diperlukan sebuah sistem. Metode *content based filtering* dinilai cocok digunakan dalam ekspansi dalam kasus aplikasi MANGAN, karena metode ini bisa menghasilkan sebuah rekomendasi terkait kemandirian pengguna. Pengujian dilakukan dengan tiga *threshold* berbeda terhadap hasil rekomendasi, diperoleh nilai presisi rata-rata 0,8915 dan nilai akurasi rata-rata 0,5118. Rendahnya nilai akurasi tersebut disebabkan oleh kesalahan atribut yang digunakan.

Dengan demikian, berdasarkan permasalahan yang dihadapi yaitu untuk mendapatkan rekomendasi *smartphone* android yang sesuai, maka penelitian yang dilakukan adalah membuat sebuah sistem rekomendasi dengan pemilihan atribut yang dinilai lebih detail dengan kasus ini yaitu spesifikasi android yang terperinci. Sistem rekomendasi ini akan membandingkan dua buah metode yaitu *demographic filtering* dan *content-based filtering* dengan landasan kedua metode tersebut cocok digunakan dalam kasus ini yang berkaitan dengan *user* dan juga atribut dari *smartphone* android itu sendiri.

2.1. Sistem Rekomendasi

Sistem rekomendasi atau sistem pendukung keputusan merupakan suatu sistem informasi interaktif berbasis komputer. Sistem rekomendasi dapat menghasilkan informasi, pemodelan dan manipulasi data yang dapat digunakan sebagai pengambilan keputusan yang menyediakan sebuah rekomendasi yang berguna untuk pengguna [12].

Selain itu, sistem rekomendasi juga merupakan sebuah model aplikasi yang dihasilkan dari observasi keadaan dan keinginan terhadap pengguna serta mempermudah pengguna dalam mengambil keputusan yang dinilai sesuai dalam menentukan produk [13].

2.2. Demographic Filtering

Demographic Filtering merupakan pendekatan sebuah sistem rekomendasi dengan cara merekomendasikan item dengan parameter data pengguna pada wilayah tertentu. *Demographic filtering* membuat kategori pengguna yang memiliki demografis seperti karakteristik dan perilaku. Seperti teknik kolaboratif, teknik demografis juga membentuk korelasi "*people-to-people*" tetapi menggunakan data yang berbeda. Kolaboratif dan teknik berbasis konten memerlukan riwayat peringkat pengguna yang bukan dari jenis yang diperlukan oleh pendekatan demografis [14].

Sebagai contoh apabila ada user baru akan menggunakan sistem rekomendasi dengan metode *demographic filtering*, maka user tersebut harus menginputkan atribut demografisnya yang nantinya akan dibandingkan dengan data demografis user lain yang tersedia di dataset. Semakin sesuai data demografis dari user baru dengan user dataset, maka semakin direkomendasikan juga hal terkait dari user di dataset, dalam kasus ini yaitu sebuah rekomendasi *smartphone* android.

2.3. Content-Based Filtering

Content-Based Filtering merupakan hasil dari penelitian informasi dengan landasan sistem berbasis konten. Metode sistem rekomendasi ini dimulai dengan memahami kebutuhan pengguna, yaitu terkait preferensi dan kendala apabila tersedia. Informasi tersebut digabungkan dengan *log* dari database *user* untuk membangun sebuah profil pengguna yang nantinya sistem rekomendasi mencocokkan *profil* pengguna dengan parameter informasi tentang suatu produk yang tersedia dalam database [13].

Sebagai contoh apabila ada user baru akan menggunakan sistem rekomendasi dengan metode *content-based filtering*, maka user tersebut harus menginputkan atribut konten yang sebelumnya dimiliki oleh user baru tersebut. Dalam kasus ini, konten yang dimaksud adalah *smartphone* android, dan atribut yang dimaksud adalah spesifikasi dari android itu sendiri. Kemudian akan dibandingkan dengan *smartphone* lain di dataset, semakin mirip spesifikasi yang dimiliki maka akan semakin direkomendasikan merek *smartphone* android tersebut kepada user baru.

2.4. Uji Wilcoxon

Uji wilcoxon merupakan suatu metode analisis yang bertujuan mengukur perbedaan kelakuan terhadap dua sampel berpasangan. Analisis uji wilcoxon dapat menjadi metode pengganti dari uji *paired sample T test*. [16]. Dengan tolak ukur T, dapat digunakan rumus berikut :

$$Z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \quad (1)$$

Keterangan :

T = Hasil penjumlahan rangking bertanda positif atau negatif dengan jumlah minimum

n = Jumlah data dengan selisih 0 dieliminir

Hipotesis :

H_0 : Tidak terdapat perbedaan perlakuan data terhadap kedua hasil rekomendasi yang dihasilkan.

H_a : Terdapat perbedaan perlakuan data terhadap kedua hasil rekomendasi yang dihasilkan.

Dasar pengambilan keputusan pada analisis uji wilcoxon adalah :

- Jika T_{hitung} atau probabilitas (Asymp.Sig) < 0,05 dapat disimpulkan H_0 ditolak dan H_a diterima.
- Jika T_{hitung} atau probabilitas (Asymp.Sig) > 0,05 dapat disimpulkan H_0 diterima dan H_a ditolak.

Sebagai contoh, akan dilakukan uji wilcoxon terhadap 10 pegawai acak untuk mengetahui pengaruh pelatihan terhadap produktivitas pegawai rekam medis sebelum dan sesudah pelatihan.

- Menentukan nilai taraf nyata 0,05 yaitu $T_{0,05(10)} = 8$ sesuai tabel kritis uji wilcoxon berikut :

Tabel 1 Nilai Kritis Uji Wilcoxon

N	0,05	0,02	0,01
10	8	5	3
11	11	7	5
12	14	10	7
13	17	13	10
14	21	16	13
15	25	20	16
16	30	24	20
17	35	28	23
18	40	33	28
19	46	38	32
20	52	43	38

- Menghitung statistik uji wilcoxon :

Tabel 2 Statistik Uji Wilcoxon

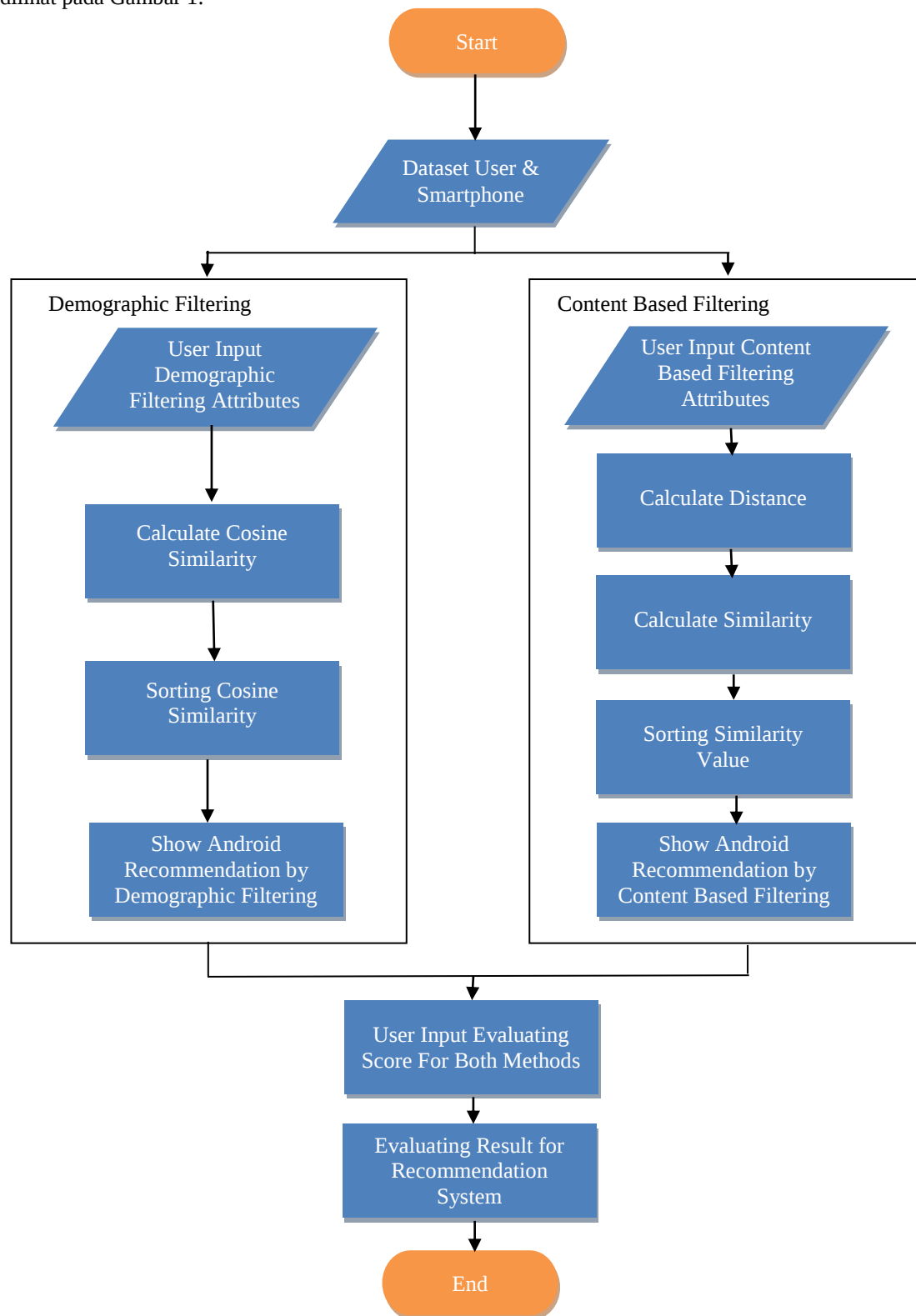
No	Sebelum	Sesudah	Selisih		Rank	Tanda Rank	
	x	y	x-y	x-y		Positif	Negatif
1	100	105	5	5	7,5	7,5	0
2	98	94	-4	4	5,5	0	5,5
3	76	78	2	2	2,5	2,5	0
4	90	98	8	8	9	9	0
5	87	90	3	3	4	4	0
6	89	85	-4	4	5,5	0	5,5
7	77	86	9	9	10	10	0
8	92	87	-5	5	7,5	0	7,5
9	78	80	2	2	2,5	2,5	0
10	82	83	1	1	1	1	0
Jumlah						36,5	18,5

- Diambil T_{hitung} dengan nilai terkecil yaitu 18,5 kemudian dibandingkan dengan T_{tabel} yaitu 8, karena $T_{hitung} > T_{tabel}$ dapat disimpulkan bahwa H_0 diterima yaitu tidak terdapat perbedaan produktivitas kerja.

3. Sistem yang Dibangun

3.1. Skema Umum

Pada sistem yang akan dibangun, peneliti membuat skema umum terkait perbandingan *demographic filtering* dan *content based filtering* pada sistem rekomendasi *smartphone* android. Skema umum dapat dilihat pada Gambar 1.



Gambar 1 Skema Umum

Pada tahap pertama merupakan pengumpulan dataset user dan smartphone android. Kemudian menuju proses demographic filtering, yaitu user menginputkan data demografinya yang kemudian akan dihitung cosine similarity lalu akan di sort nilainya dari yang terbesar ke terkecil sehingga mendapatkan sebuah rekomendasi smartphone android berdasarkan demographic filtering.

Kemudian dilakukan proses content based filtering yaitu user menginputkan data spesifikasi smartphone lamanya yang kemudian akan dihitung jaraknya dan juga kesamaannya. Lalu nilainya akan disorting dari yang terbesar ke terkecil sehingga mendapatkan sebuah rekomendasi smartphone android berdasarkan content based filtering.

Kemudian user mengevaluasi yaitu memberikan penilaian terhadap hasil rekomendasi smartphone android berdasarkan kedua metode yang nantinya hasil penilaian tersebut akan dijadikan dievaluasi dan dianalisis lebih lanjut.

3.2. Dataset

Pada penelitian ini data yang digunakan adalah dataset hasil dari survey terhadap 150 pengguna yang isinya berupa usia, pekerjaan, pendapatan, dan *smartphone* android yang digunakan ditambah dengan 220 *smartphone* android hasil pengembangan dari *smartphone* yang dimiliki oleh user.

Berikut merupakan beberapa contoh dataset user dan *smartphone* android :

Tabel 3 Dataset User

id	id_smartphone	nama_user	usia_user	pekerjaan_user	pendapatan_user
1	1	Asri Lestari R	20 < Usia <= 30	Wiraswasta	Pendapatan <= 2.500.000
2	4	Tuti Helawati	40 < Usia <= 50	Pegawa Swasta	7.500.000 < Pendapatan <= 10.000.000
3	183	Sumanah	40 < Usia <= 50	Pegawa Swasta	2.500.000 < Pendapatan <= 5.000.000
4	5	Eti Suhaeti	Usia > 50	Pegawai Negeri Sipil	5.000.000 < Pendapatan <= 7.500.000
5	2	Rukamah	40 < Usia <= 50	Wiraswasta	5.000.000 < Pendapatan <= 7.500.000
6	6	Wawan Setiawan	Usia > 50	Pegawai Negeri Sipil	Pendapatan > 10.000.000
7	145	Sansan Setiawan	30 < Usia <= 40	Pegawa Swasta	7.500.000 < Pendapatan <= 10.000.000
8	8	Sumi Muniroh	40 < Usia <= 50	Pegawai Negeri Sipil	2.500.000 < Pendapatan <= 5.000.000
9	172	Yeni Marlani	Usia > 50	Wiraswasta	2.500.000 < Pendapatan <= 5.000.000
10	11	Nurhayati	Usia > 50	Pegawai Negeri Sipil	2.500.000 < Pendapatan <= 5.000.000

Tabel 4 Dataset Smartphone

id	merk_android	nama_android	harga_android	ram_android	memori_android	baterai_android	finger_print_android	face_android	kamera_android	foto_android
1	Vivo	Vivo Y33T	3.399.000	8	128	5.000	Tersedia	Tersedia	4	vivo-y33t.jpg
2	Oppo	Oppo F11 (4/64)	3.599.000	4	64	4.020	Tersedia	Tersedia	3	oppo-f11.jpg
3	Oppo	Oppo F11 (4/128)	3.899.000	6	64	4.020	Tersedia	Tersedia	3	oppo-f11.jpg
4	Oppo	Oppo F11 (6/64)	4.299.000	4	128	4.020	Tersedia	Tersedia	3	oppo-f11.jpg
5	Oppo	Oppo F11 (6/128)	4.699.000	6	128	4.020	Tersedia	Tersedia	3	oppo-f11.jpg
6	Samsung	Note 20	10.499.999	8	256	4.300	Tersedia	Tersedia	4	note20.jpg
7	Realme	Realme 8 (4/64)	2.499.000	4	64	5.000	Tersedia	Tersedia	5	realme-8.jpg
8	Realme	Realme 8 (4/128)	2.699.000	4	128	5.000	Tersedia	Tersedia	5	realme-8.jpg
9	Realme	Realme 8 (6/128)	2.999.000	6	128	5.000	Tersedia	Tersedia	5	realme-8.jpg
10	Realme	Realme 8 (8/128)	3.299.000	8	128	5.000	Tersedia	Tersedia	5	realme-8.jpg

Dari data tersebut akan dibuatkan 2 buah database yaitu database user dan database *smartphone*. Database user memiliki atribut usia, pekerjaan, pendapatan, dan id *smartphone*. Database ini akan digunakan dalam metode *demographic filtering*. Lalu database *smartphone* memiliki atribut id *smartphone*, merek android, nama android, harga android, ram android, memori internal, baterai, ketersediaan fingerprint, ketersediaan *face id*, dan jumlah kamera. Database ini akan digunakan dalam metode *content-based filtering*.

Pengklasifikasian bobot dari berbagai atribut ini terbagi menjadi dua, yaitu pengklasifikasian atribut demografis dan pengklasifikasian atribut konten. Pengklasifikasian atribut demografis diambil berdasarkan proses penggabungan antara atribut yang dinilai cukup sesuai dari penelitian tentang rekomendasi mobil oleh Herastia dan Ferry pada tahun 2014 [3] dengan pemilihan bobot yang selanjutnya dipilih dengan menimbang atribut yang dinilai cocok dengan pengguna *smartphone* android. Sedangkan pengklasifikasian atribut konten, dipilih berdasarkan perincian dari spesifikasi *smartphone* android yang digunakan.

Berikut merupakan pengklasifikasian bobot dari berbagai atribut pada *demographic filtering* :

Tabel 5 Klasifikasi Usia User

Usia	Kelas
Usia ≤ 20	1
$20 < \text{Usia} \leq 30$	2
$30 < \text{Usia} \leq 40$	3
$40 < \text{Usia} \leq 50$	4
Usia > 50	5

Indikator yang digunakan dalam rekomendasi *smartphone* android berdasarkan kriteria usia dimana dinilai atau dilihat dari umur calon pembeli. Terdapat 5 kelas yang merepresentasikan bobot jarak.

Tabel 6 Klasifikasi Pekerjaan User

Pekerjaan	Kelas
Pelajar	1
Wiraswasta	2
Pegawai Negeri Sipil	3
Pegawai Swasta	4
Profesional	5

Indikator yang digunakan dalam rekomendasi *smartphone* android berdasarkan kriteria pekerjaan dimana dinilai atau dilihat dari pekerjaan calon pembeli. Terdapat 5 kelas yang merepresentasikan bobot jarak, namun tidak mewakili tingkat baik atau buruknya dari setiap tingkatan kelas.

Tabel 7 Klasifikasi Pendapatan User

Pendapatan	Kelas
Pendapatan $\leq 2.500.000$	1
$2.500.000 < \text{Pendapatan} \leq 5.000.000$	2
$5.000.000 < \text{Pendapatan} \leq 7.500.000$	3
$7.500.000 < \text{Pendapatan} \leq 10.000.000$	4
Pendapatan $> 10.000.000$	5

Indikator yang digunakan dalam rekomendasi *smartphone* android berdasarkan kriteria pendapatan dimana dinilai atau dilihat dari pendapatan calon pembeli. Terdapat 5 kelas yang merepresentasikan bobot jarak.

Berikut merupakan pengklasifikasian bobot dari berbagai atribut pada *content based filtering* :

Tabel 8 Klasifikasi Harga Android

Harga Android	Kelas
Harga $\leq 2.000.000$	1
$2.000.000 < \text{Harga} \leq 4.000.000$	2
$4.000.000 < \text{Harga} \leq 6.000.000$	3
$6.000.000 < \text{Harga} \leq 10.000.000$	4
Harga $> 10.000.000$	5

Indikator yang digunakan dalam rekomendasi *smartphone* android berdasarkan kriteria harga dimana dinilai atau dilihat dari harga *smartphone* yang diinginkan calon pembeli. Terdapat 5 kelas yang merepresentasikan bobot jarak.

Tabel 9 Klasifikasi RAM

RAM	Kelas
RAM $\leq 2 \text{ GB}$	1
RAM 3 GB	2
RAM 4 GB	3
RAM 6 GB	4
RAM $> 6 \text{ GB}$	5

Indikator yang digunakan dalam rekomendasi *smartphone* android berdasarkan kriteria RAM dimana dinilai atau dilihat dari RAM *smartphone* yang diinginkan calon pembeli. Terdapat 5 kelas yang merepresentasikan bobot jarak.

Tabel 10 Klasifikasi Memori Internal

Memori Internal	Kelas
Memori ≤ 16 GB	1
Memori 32 GB	2
Memori 64 GB	3
Memori 128 GB	4
Memori > 128 GB	5

Indikator yang digunakan dalam rekomendasi *smartphone* android berdasarkan kriteria memori internal dimana dinilai atau dilihat dari memori internal *smartphone* yang diinginkan calon pembeli. Terdapat 5 kelas yang merepresentasikan bobot jarak.

Tabel 11 Klasifikasi Baterai

Baterai	Kelas
Baterai ≤ 3000 mAh	1
$3000 \text{ mAh} < \text{Baterai} \leq 4000 \text{ mAh}$	2
$4000 \text{ mAh} < \text{Baterai} \leq 4500 \text{ mAh}$	3
$4500 \text{ mAh} < \text{Baterai} \leq 5500 \text{ mAh}$	4
Baterai > 5500 mAh	5

Indikator yang digunakan dalam rekomendasi *smartphone* android berdasarkan kriteria baterai dimana dinilai atau dilihat dari baterai *smartphone* yang diinginkan calon pembeli. Terdapat 5 kelas yang merepresentasikan bobot jarak.

Tabel 12 Klasifikasi Fingerprint

Fingerprint	Kelas
Tersedia	1
Tidak Tersedia	0

Indikator yang digunakan dalam rekomendasi *smartphone* android berdasarkan kriteria fingerprint dimana dinilai atau dilihat dari *fingerprint* *smartphone* yang diinginkan calon pembeli. Terdapat 5 kelas yang merepresentasikan bobot jarak.

Tabel 13 Klasifikasi Face ID

Face Unlock	Kelas
Tersedia	1
Tidak Tersedia	0

Indikator yang digunakan dalam rekomendasi *smartphone* android berdasarkan kriteria face unlock dimana dinilai atau dilihat dari face unlock *smartphone* yang diinginkan calon pembeli. Terdapat 5 kelas yang merepresentasikan bobot jarak.

Tabel 14 Klasifikasi Kamera

Kamera	Kelas
Kamera ≤ 2	1
Kamera 3	2
Kamera 4	3
Kamera 5	4
Kamera > 5	5

Indikator yang digunakan dalam rekomendasi *smartphone* android berdasarkan kriteria kamera dimana dinilai atau dilihat dari kamera *smartphone* yang diinginkan calon pembeli. Terdapat 5 kelas yang merepresentasikan bobot jarak.

3.3. Implementasi Metode

a) Metode *Demographic Filtering*

Pada tahap ini yang dilakukan adalah menangani pengguna yang baru mendaftar atau belum memiliki informasi tentang *smartphone* android. Setiap user memiliki atribut usia, pekerjaan, pendapatan, dan id *smartphone*. Dengan penilaian dari kesamaan atribut tersebut akan direkomendasikan android sesuai dengan data demografis dari user lain di database.

Dalam penelitian ini digunakan metode cosine similarity untuk menghitung nilai vektor kemiripan dari salah satu user terhadap user lain dengan menggunakan rumus sebagai berikut [15]:

$$\text{Cos}(\alpha) = \frac{\mathbf{A} \cdot \mathbf{B}}{|\mathbf{A}| |\mathbf{B}|} \quad (2)$$

Keterangan :

A = Vektor user A

B = Vektor user B

|A| = Panjang vektor user A

|B| = Panjang vektor user B

Dari rumus diatas, dilakukan penghitungan kemiripan antara vektor user baru dengan semua vektor user yang tersedia di database. Semakin besar kemiripan vektor antar user, semakin tinggi juga nilai *cosine similarity*-nya, dan akan semakin direkomendasikan juga *smartphone* androidnya kepada user baru yang menjalankan sistem rekomendasi.

Pseudocode metode *demographic filtering* dapat dilihat pada Gambar 4.

Algoritma CosineSimilarity

Deklarasi

data_a : array [1...maxN] of integer #maxN adalah jumlah attribut *demographic filtering* yaitu 3
 data_b : array [1...maxM, 1...maxN] of integer #maxM adalah jumlah user di database yaitu 150
 a,i,j,k : integer
 n,x,y,z : integer
 top,div,sum : integer

Begin

```
Function cos(user_demographic)
  data_a ← user_demographic
  data_b ← data_demographic
  a ← 0
  n ← 0
  FOR (i ← 1 ; i ≤ 20 ; i++)
    top ← 0
    div ← 0
    FOR (j ← 1 ; j ≤ 3 ; j++)
      top ← top + (data_a[j] * data_b[i][j])
      div ← div + (data_b[i][j] - data_a[j]) * (data_b[i][j] - data_a[j])
    END FOR
    sum ← top / sqrt(div)
    IF sum > a THEN
      a ← sum
      n ← i
    END IF
  END FOR
  Return data_b[n]
END Function
Arsort (data_b)
END
```

Gambar 4 Pseudocode *Demographic Filtering*

Disediakan dua buah array yaitu *data_a* dengan maxN 3 sesuai attribute *demographic filtering* dan *data_b* dengan maxM 150 sesuai jumlah user di database. Array *data_a* diisi oleh vektor data demografis dari user, sedangkan array *data_b* diisi oleh vektor data demografis user yang tersedia di database. Vektor data demografis user dibandingkan dengan cara dihitung *cosine similarity*-nya dengan semua vektor data demografis user yang tersedia di database. Ketika sudah didapatkan semua nilai *cosine similarity*-nya, akan dilakukan sorting dari terbesar ke terkecil dan dipilih *smartphone* dengan nilai *cosine similarity* terbesar yang akan direkomendasikan kepada *user*.

b) Metode Content Based Filtering

Pada tahap ini user baru akan memasukkan data *smartphone* lamanya, atau bisa juga memasukkan spesifikasi yang dimiliki apabila *smartphone* tidak ada dalam database. Lalu *smartphone* user baru akan dihitung kemiripannya dengan *smartphone* yang ada di database, sehingga akan didapatkan rekomendasi *smartphone* yang memiliki spesifikasi paling mendekati dengan *smartphone* yang diinputkan.

Untuk menghitung *similarity* atau nilai kemiripan antara profil *user* dengan *smartphone* android, maka dilakukan penghitungan jarak antara profil *user* dengan *smartphone* dengan persamaan sebagai berikut [3]:

$$\text{calc_distance}(\text{data_a}, \text{data_b}) = \sum_{i=1}^n (\text{data_a}_i - \text{data_b}_i)^2 \quad (3)$$

Keterangan:

n = Total elemen dalam vektor android
data_a = Nilai parameter dari vektor profil *user*
data_b = Nilai parameter dari vektor android

Dari rumus diatas dihitung jarak antara vektor spesifikasi *smartphone* lama user dengan semua vektor *smartphone* yang tersedia di database. Semakin besar nilai *distance* atau jarak antara kedua vektor, maka akan semakin besar kesempatan *smartphone* di database tersebut akan direkomendasikan kepada user. Namun sebelum itu, setelah diperoleh jarak, maka nilai *similarity* dihitung berdasarkan rumus berikut [3]:

$$\text{sim_dist}(\text{data_a}, \text{data_b}) = \frac{1}{1 + \text{calc_distance}(\text{data_a}, \text{data_b})} \quad (4)$$

Dari rumus diatas dihitung *similarity* atau kesamaan antar kedua vektor *smartphone*. Semakin tinggi nilai *similarity* yang didapatkan, maka akan semakin direkomendasikan *smartphone* dari database tersebut kepada user. Pseudocode metode *content based filtering* dapat dilihat pada Gambar 5.

Algoritma Distance_Similarity

Deklarasi

data_a : array [1...maxN] of integer #maxN adalah jumlah attribut *content based filtering* yaitu 7
data_b : array [1...maxM, 1...maxN] of integer #maxM adalah jumlah *smartphone* di database yaitu 220
a, i, j, n : integer
calc_distance, sim_dist : integer

Begin

Function *calc_distance*(*user_smartphone*)

data_a ← *user_smartphone*

data_b ← *data_smartphone*

a ← 0

n ← 0

FOR (*i* ← 1 ; *i* ≤ 220 ; *i*++)

sim_dist ← 0

FOR (*j* ← 1 ; *j* ≤ 7 ; *j*++)

calc_distance ← (*data_b*[*i*][*j*] - *data_a*[*j*]) * (*data_b*[*i*][*j*] - *data_a*[*j*])

sim_dist ← *sim_dist* + *calc_distance*

END FOR

IF *sim_dist* > *a* THEN

a ← *sim_dist*

n ← *i*

END IF

END FOR

Return *data_b*[*n*]

END Function

Arsort (*data_b*)

END

Gambar 5 Pseudocode Content-based Filtering

Disediakan dua buah array yaitu data_a dengan maxN 7 yaitu jumlah atribut *content based filtering* dan data_b dengan maxM yaitu 220 sesuai jumlah *smartphone* yang tersedia di database. Array data_a diisi oleh vektor *smartphone* lama dari user, sedangkan array data_b diisi oleh vektor *smartphone* yang tersedia di database. Vektor *smartphone* lama user dibandingkan dengan cara dihitung *similarity*-nya dengan semua vektor *smartphone* yang tersedia di database. Ketika sudah didapatkan semua nilai *similarity*-nya, akan dilakukan sorting dari terbesar ke terkecil dan dipilih *smartphone* dengan nilai *similarity* terbesar yang akan direkomendasikan kepada *user*.

c) Hasil Evaluasi Metode Demographic Filtering dan Content-Based Filtering

Pada langkah terakhir ini, *user* baru akan diberikan hasil rekomendasi *smartphone* android berdasarkan metode *demographic filtering* dan *content-based filtering*. Kemudian *user* baru akan memberikan penilaian terhadap hasil rekomendasi *smartphone* android dari kedua metode tersebut. Adapun kriteria penilaian untuk evaluasi *user* seperti pada tabel 15.

Tabel 15 Kriteria Penilaian terhadap metode yang digunakan

Nilai	Penjelasan
1	Tidak Baik
2	Kurang Baik
3	Cukup Baik
4	Baik
5	Sangat Baik

4. Evaluasi

4.1 Hasil Pengujian

Tujuan dari pengujian ini adalah untuk mendapatkan evaluasi penilaian dari *user* terhadap sistem rekomendasi *smartphone* android yang dibangun menggunakan metode *demographic filtering* dan *content based filtering*. Kemudian mengetahui perbandingan performansi dari kedua metode yang dibangun pada sistem rekomendasi *smartphone* android.

Setelah *user* menginputkan data *demographic filtering* dan data *content based filtering*, *user* diberikan dua hasil rekomendasi *smartphone* android yang berbeda berdasarkan kedua metode tersebut. Kemudian *user* memberikan evaluasi terhadap kedua hasil rekomendasi tersebut berdasarkan kriteria penilaian dari tabel 15. Berikut hasil dari pengujian yang diperoleh :

Tabel 16 Penilaian terhadap metode yang digunakan

User	Penilaian Demographic Filtering	Penilaian Content Based Filtering
User-1	3	4
User-2	2	5
User-3	4	5
User-4	4	4
User-5	5	3
User-6	3	3
User-7	2	4
User-8	3	4
User-9	5	5
User-10	4	3
User-11	4	4
User-12	4	5
User-13	5	5
User-14	5	4
User-15	4	3
User-16	5	4
User-17	4	5
User-18	5	5
User-19	4	4
User-20	3	5
Rata-rata	3,9	4,2

Berdasarkan tabel 16, rata-rata hasil penilaian user terhadap sistem yang dibangun menggunakan metode *content based filtering* mendapatkan nilai lebih tinggi daripada menggunakan metode *demographic filtering*. Sistem rekomendasi smartphone android menggunakan metode *content based filtering* memperoleh penilaian rata-rata yaitu sebesar 4,2 sedangkan sistem rekomendasi smartphone android menggunakan metode *demographic filtering* memperoleh penilaian rata-rata yaitu sebesar 3,9.

4.2 Analisis Hasil Pengujian

Dari hasil penilaian yang diberikan oleh user, kemudian dilakukan analisis Wilcoxon untuk menguji perbedaan kelompok data dari hasil penilaian *demographic filtering* dengan *content based filtering*. Berikut hasil dari analisis yang diperoleh :

Tabel 17 Analisis Uji Wilcoxon

User	Demographic	Content Based	Selisih		Rank	Tanda Rank	
	x	y	x-y	x-y		Positif	Negatif
User-1	3	4	-1	1	12	0	12
User-2	2	5	-3	3	20	0	20
User-3	4	5	-1	1	12	0	12
User-4	4	4	0	0	4	0	0
User-5	5	3	2	2	18	18	0
User-6	3	3	0	0	4	0	0
User-7	2	4	-2	2	18	0	18
User-8	3	4	-1	1	12	0	12
User-9	5	5	0	0	4	0	0
User-10	4	3	1	1	12	12	0
User-11	4	4	0	0	4	0	0
User-12	4	5	-1	1	12	0	12
User-13	5	5	0	0	4	0	0
User-14	5	4	1	1	12	12	0
User-15	4	3	1	1	12	12	0
User-16	5	4	1	1	12	12	0
User-17	4	5	-1	1	12	0	12
User-18	5	5	0	0	4	0	0
User-19	4	4	0	0	4	0	0
User-20	3	5	-2	2	18	0	18
Jumlah						66	116

Berdasarkan tabel 17, T_{hitung} yang diperoleh adalah 66. Dengan taraf nyata yaitu 0,05, T_{tabel} yang didapat dari nilai kritis uji Wilcoxon adalah 52. Dalam kasus ini, T_{hitung} lebih besar dari T_{tabel} yang merupakan tidak ada perbedaan perlakuan terhadap kedua hasil penilaian tersebut.

Hasil rekomendasi smartphone android menggunakan metode *content based filtering* mendapatkan penilaian lebih baik dibandingkan hasil rekomendasi smartphone android menggunakan metode *demographic filtering*. Faktor yang menjadikan penilaian dari metode *content based filtering* lebih baik adalah pemilihan parameter yang spesifik dan mampu menggambarkan kebutuhan user terhadap smartphone yang dibutuhkan. Disisi lain pemilihan parameter dari metode *demographic filtering* masih dinilai kurang spesifik sehingga hasil rekomendasi smartphone android yang dihasilkan masih dinilai kurang relevan terhadap kebutuhan dari user.

5. Kesimpulan

Berdasarkan penelitian yang dilakukan, maka dapat diambil kesimpulan bahwa metode *content based filtering* mendapatkan hasil rekomendasi yang lebih baik dibandingkan dengan metode *demographic filtering*. Hal ini dipengaruhi oleh parameter spesifik yang dimiliki metode *content based filtering* untuk merepresentasikan smartphone yang dibutuhkan. Sedangkan parameter yang dimiliki metode *demographic filtering* masih dinilai kurang spesifik untuk merepresentasikan kebutuhan user sehingga perlu dilakukan analisis lebih lanjut untuk mendapatkan data demografis yang sesuai untuk dijadikan parameter.

Saran yang diberikan untuk penelitian kedepannya adalah mencari parameter yang lebih spesifik terutama pada metode *demographic filtering* sehingga dapat meningkatkan performansi pada sistem yang dibangun, juga mendapatkan akurasi yang lebih baik sehingga bisa merepresentasikan kebutuhan user.

Daftar Pustaka

- [1] S. Kemp, "Digital 2023: Indonesia," [www.datareportal.com](https://datareportal.com/reports/digital-2023-indonesia), 2023. <https://datareportal.com/reports/digital-2023-indonesia>.
- [2] A. A. Huda, "Pemanfaatan Karakter Rasa pada Makanan sebagai Dasar Penentuan Rekomendasi Restoran," *JISKA (Jurnal Informatika Sunan Kalijaga)*, p. 2045, 2018.
- [3] H. Maharani and F. A. Gunawan, "Sistem Rekomendasi Mobil Berdasarkan Demographic dan Content-Based Filtering," *J. Telemat.*, vol. 9, no. 2, pp. 64–68, 2014.
- [4] T. Achakulvisut, D. E. Acuna, T. Ruangrong, and K. Kording, "Science Concierge: A fast content-based recommendation system for scientific publications," *PLoS One*, vol. 11, no. 7, pp. 1–11, 2016, doi: 10.1371/journal.pone.0158423.
- [5] R. F. Oeyliawan and D. Gunawan, "Aplikasi Rekomendasi Buku Pada Katalog Perpustakaan Universitas Multimedia Nusantara Menggunakan Vector Space Model," *J. Ultim.*, vol. 9, no. 2, pp. 97–105, 2018, doi: 10.31937/ti.v9i2.639.
- [6] R. Van Meteren and M. Van Someren, "Using Content-Based Filtering for Recommendation," *ECML/MLNET Work. Mach. Learn. New Inf. Age*, pp. 47–56, 2000, [Online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.25.5743&rep=rep1&type=pdf>.
- [7] F. Ricci, L. Rokach, and B. Shapira, *Recommender Systems Handbook*, no. June 2014. 2011.
- [8] C. Fiarni and H. Maharani, "Product Recommendation System Design Using Cosine Similarity and Content-based Filtering Methods," *IJITEE (International J. Inf. Technol. Electr. Eng.)*, vol. 3, no. 2, p. 42, 2019, doi: 10.22146/ijitee.45538.
- [9] L. Tommy, C. Kirana, and V. Lindawati, "Recommender System Dengan Kombinasi Apriori Dan Content-Based Filtering Pada Aplikasi Pemesanan Produk," *J. Teknoinfo*, vol. 13, no. 2, p. 84, 2019, doi: 10.33365/jti.v13i2.299.
- [10] A. E. Wijaya and D. Alfian, "Sistem Rekomendasi Laptop Menggunakan Collaborative Filtering Dan Content-Based Filtering," *J. Comput. Bisnis*, vol. 12, no. 1, pp. 11–27, 2018.
- [11] R. H. Mondy and A. Wijayanto, "Recommendation System With Content-Based Filtering Method for Culinary Tourism in," vol. 8, no. 2, pp. 65–72, 2019.
- [12] M. A. Purwanto, I. M. I. Subroto, and D. Kurniadi, "Sistem Rekomendasi Penerimaan Kartu Indonesia Pintar (Kip) Menggunakan Metode Simple Additive Weighting," *TRANSISTOR Elektro dan Inform.*, vol. 3, no. 2, pp. 111–119, 2018, [Online]. Available: <http://jurnal.unissula.ac.id/index.php/EI/article/view/3986>.
- [13] N. Rosmawarni, "Perancangan sistem rekomendasi untuk pengembangan aplikasi salon terpadu berbasis android," *Rekayasa Inf.*, vol. 6, no. 1, pp. 61–70, 2017.
- [14] I. Ryngksai and L. Chameikho, "Recommender Systems : Types of Filtering," vol. 3, no. 11, pp. 251–254, 2014.
- [15] R. T. Wahyuni, D. Prastiyanto, and E. Suprpto, "Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi," *J. Tek. Elektro*, vol. 9, no. 1, pp. 18–23, 2017, doi: 10.15294/jte.v9i1.10955.
- [16] S. Santoso, *SPSS Versi 10 Mengolah Data Statistik Secara Profesional*, Jakarta: *PT Elex Media Komputindo*, 2011.