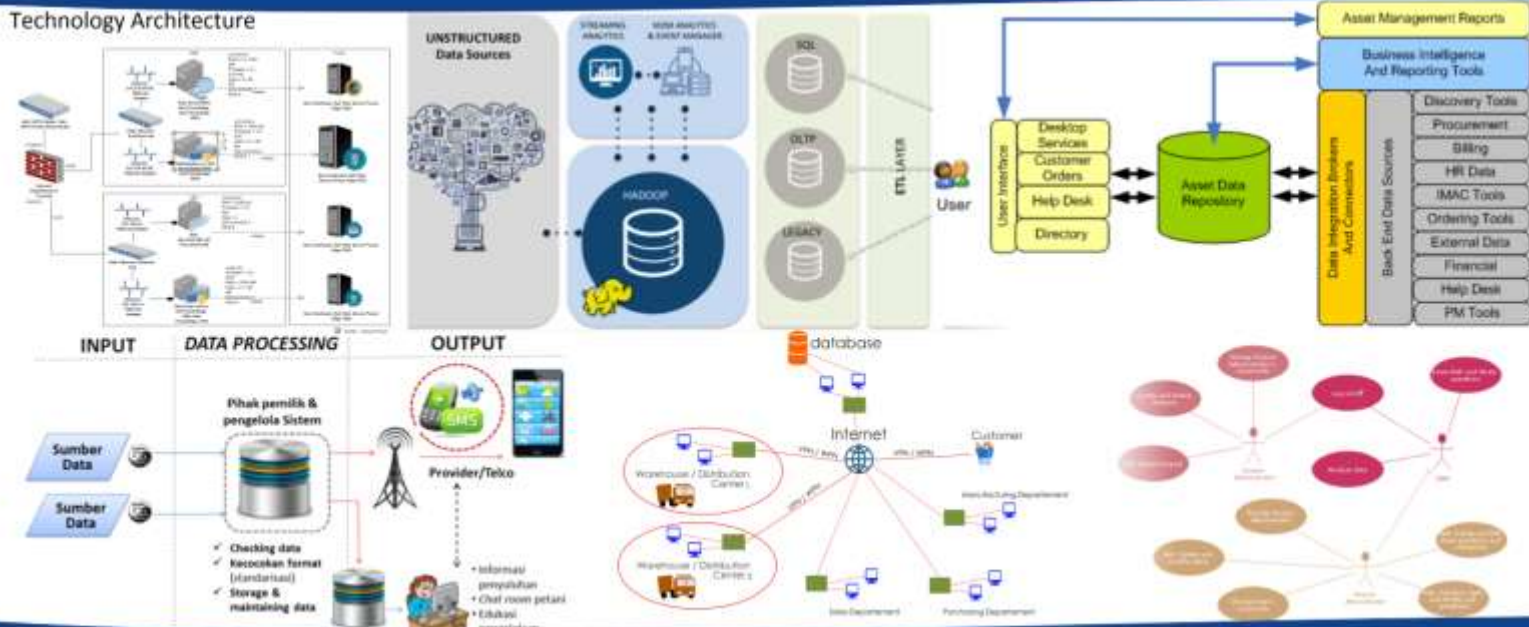




MISI

JURNAL MANAJEMEN INFORMATIKA & SISTEM INFORMASI

Technology Architecture



Diterbitkan Oleh LPPM STMIK Lombok
Jln. Basuki Rahmat No.105 Praya, Lombok Tengah - NTB
Telp dan Fax (0370) 654310 - e-journal.stmiklombok.ac.id/jsi
email. lppm@stmiklombok.ac.id



DEWAN REDAKSI

JURNAL MISI (JURNAL MANAJEMEN INFORMATIKA DAN SISTEM INFORMASI)

Jurnal Manager

Wire Bagye, S.Kom.,M.Kom (STMIK Lombok, SINTA ID : 5992010)

Reviewer :

Resad Setyadi, S.T., S.Si., MMSI., Ph.D (cand)- Institut Teknologi Telkom Purwokerto
SCOPUS ID 57204172534, SINTA ID : 6113570

Yesaya Tommy Paulus, S.Kom., MT., Ph.D. - STMIK Dipanegara Makassar
SCOPUS ID 57202829909, SINTA ID : 6002004

Lalu Mutawalli, S.Kom., M.I.Kom., M.Kom - STMIK Lombok
SCOPUS ID : 57205057118, SINTA ID : 6659709

Saruni Dwiasnati, ST., MM., M.Kom - Universitas Mercu Buana
SCOPUS ID : 57210968603, SINTA ID : 6150854

Ida Bagus Ary Indra Iswara, S.Kom., M.Kom - STMIK STIKOM Indonesia
SCOPUS ID 57203711945, SINTA ID : 183498

Erlin Windia Ambarsari - Universitas Indraprasta PGRI
SCOPUS ID : 56242503900, SINTA ID : 5998887

Wafiah Murniati, ST., MT. - STMIK Lombok
SCOPUS ID : 56242503900, SINTA ID : 5998887

Yuliadi, S.Kom., M.Kom - Universitas Teknologi Sumbawa
SINTA ID : 6730786

Fachrudin Pakaja, S.Kom, M.T - Universitas Gajayana
SINTA ID : 6164357

Ahmad Jufri, S.Kom., M.T - Sekolah Tinggi Teknologi STIKMA Internasional
SINTA ID : 172241

Mohammad Taufan Asri Zaen, ST., MT - STMIK Lombok
SINTA ID : 5992087

Hairul Fahmi, S.Kom., M.Kom - STMIK Lombok
SINTA ID : 5983160

I Ketut Putu Suniantara, S.Si., M.Si - ITB STIKOM Bali
SINTA ID : 6086221

Nawassyarif S. Kom., M.Pd. - Universitas Teknologi Sumbawa
SINTA ID : 6722660

Muhamad Malik Mutoffar, ST., MM., CNSS - Sekolah Tinggi Teknologi Bandung
SINTA ID : 6013819

Editor :

Saikin, Skom., M.Kom. - STMIK Lombok

Vrestanti Novalia Santosa, M.Pd. - IKIP Budi Utomo Malang

Desain Grafis & Web Maintenance

Jihadul Akbar, S.Kom - STMIK Lombok

Secretariat

Maulana Ashari, M.Kom - STMIK Lombok

DAFTAR ISI

- 1** ANALISIS CLUSTERING PROVINSI DI INDONESIA BERDASARKAN TINGKAT KEMISKINAN MENGGUNAKAN ALGORITMA K-MEANS **1 - 8**
Achmad Bahauddin¹, Agustina Fatmawati², Febrianti Permata Sari³
- 2** PEMBOBOTAN MENGGUNAKAN *PAIRWISE COMPARISON* PADA *CASE BASED REASONING* REKOMENDASI HOTEL **9 - 18**
Kukuh Tri Nur Iman¹, Setyawan Wibisono²
- 3** IMPLEMENTASI METODE AHP PADA SISTEM PENDUKUNG KEPUTUSAN PENYELEKSIAN NASABAH PINJAMAN KREDIT **19 - 27**
Irfak Lahumu Darajat¹, Wiwien Hadikurniawati²
- 4** SELEKSI PENERIMAAN BEASISWA BIDIKMISI PADA STMIK INDONESIA PADANG MENGGUNAKAN METODE (AHP) **28 - 35**
Heru Saputra¹, Efendi Mardiono², Ilfa Stephane³, Ratih Purwasih⁴
- 5** PENGELOMPOKKAN JENIS RUMPUT LAUT MENGGUNAKAN FUZZY C-MEANS BERBASIS CITRA **36 - 44**
Franki Yusuf Bisilisin¹, Remerta Noni Naatonis²
- 6** SISTEM REKOMENDASI PRODUCT EMINA COSMETICS DENGAN MENGGUNAKAN METODE CONTENT - BASED FILTERING **45 - 54**
Fatoni Batari Agung Larasati¹, Herny Februariyanti²
- 7** SISTEM INFORMASI BOOKING (STUDI KASUS: REGGAENERASI INK STUDIO) **55 - 62**
Ni Wayan Yesi Mertha Sari¹, Ni Luh Putu Ning Septyarini Putri Astawa², I Nyoman Yudi Anggara Wijaya³
- 8** PENERAPAN METODE SMART DALAM SISTEM PENDUKUNG KEPUTUSAN PEMBERIAN SANKSI PELANGGARAN TATA TERTIB SISWA (Studi Kasus: SMK Negeri 1 Pujut) **63 - 72**
Mohammad Taufan Asri Zaen¹, Baiq Daniatan Janiah², Sofiansyah Fadli³
- 9** RANCANGAN SISTEM INFORMASI PERHITUNGAN PENYUSUTAN *FIXED ASSETS* MENGGUNAKAN *STRAIGHT LINE METHOD* PADA PT FIF GROUP PEMATANGSIANTAR **73 - 77**
Ayu Tiara Defi¹, Dedi Suhendro²
- 10** PERANCANGAN SIMPLE STATELESS AUTENTIKASI DAN OTORISASI LAYANAN REST-API BERBASIS PROTOKOL HTTP **78 - 87**
I Gusti Ngurah Ady Kusuma

SISTEM REKOMENDASI PRODUCT EMINA COSMETICS DENGAN MENGGUNAKAN METODE CONTENT - BASED FILTERING

Fatoni Batari Agung Larasati¹, Herny Februariyanti²

^{1,2}Program Studi Teknik Informatika Universitas STIKUBANK Semarang

Jl. Trilomba Juang No 1 Semarang 50241
Jl. Kendeng V, Bendan Ngisor, Kec. Semarang Sel., Kota Semarang, Jawa Tengah 50241

¹larasatifatoni26@gmail.com, ²herny@edu.unisbank.ac.id

Abstract

*Emina cosmetic is a cosmetic product from Paragon Technology and Innovation Company with the concept of cosmetics for teenagers and young adults. Over time, Emina products will certainly increase in variants. With the increasing number of products, sometimes the customer find it difficult to determine the right product to use, therefore a system is needed that can recommend products according to customer interests and needs. This study designed an application to provide Emina cosmetic recommendations to customers based on the previously searched product. The recommendation process is carried out by comparing the similarities between the products that the customer is looking for with the existing product description. This study uses a content-based filtering method where this method can be used to recommend products based on the availability of content / product descriptions. To calculate the similarity between sentences using the cosine similarity algorithm. Initially, the product description will be weighted using *tfidf*, then the similarity value will be calculated using the cosine similarity algorithm. Products containing keywords will be calculated for their similarity values and ranking based on the highest to lowest similarity values. In this study, the product with the highest similarity was obtained with a value of 0.7195.*

Keywords : *Recommender System, Content-Based Filtering, Cosine Similarity Algorithm, tf-idf, Cosmetic*

Abstrak

Emina cosmetic merupakan produk kosmetik dari PT Paragon Technology and Innovation dengan mengusung konsep kosmetik untuk remaja dan dewasa muda. Seiring berjalannya waktu, produk emina tentunya akan bertambah varian. Dengan semakin bertambahnya jumlah produk, para customer terkadang akan merasa kesulitan dalam menentukan produk yang tepat untuk digunakan, maka dari itu dibutuhkan sistem yang dapat merekomendasikan produk sesuai dengan ketertarikan dan kebutuhan customer. Penelitian ini bertujuan untuk merancang sebuah aplikasi yang dapat memberikan rekomendasi kosmetik emina kepada customer berdasarkan produk yang dicari sebelumnya. Proses rekomedasi dilakukan dengan membandingkan kesamaan antara produk yang dicari customer dengan deskripsi produk yang ada. Penelitian ini menggunakan metode *content-based filtering* dimana metode tersebut dapat digunakan untuk merekomendasikan produk berdasarkan ketersediaan konten/ deskripsi produk. Untuk menghitung kesamaan antar kalimat menggunakan algoritma *cosinesimilarity*. Mulanya deskripsi produk akan dilakukan pembobotan dengan *tfidf*, lalu akan dihitung nilai similaritasnya dengan algoritma *cosinesimilarity*. Produk yang mengandung kata kunci akan dihitung nilai kemiripannya dan dilakukan perangkingan berdasarkan nilai similaritas tertinggi hingga terendah. Dalam penelitian ini, produk dengan similaritas tertinggi didapat dengan nilai sebesar 0,7195.

Kata kunci : *Sistem Rekomendasi, Content-based filtering, Algoritma cosine similarity, tf-idf, kosmetik*

1. PENDAHULUAN

Berkembangnya dunia kosmetik dewasa ini semakin melesat, terlebih dengan hadirnya *beauty vlogger* dan *beautyinfluencer* yang juga ikut mempengaruhi gaya hidup masyarakat tentang penggunaan kosmetik. Menurut

Tranggono (dalam Suhartini, dkk., 2013) penggunaan kosmetik dibagi menjadi 2, yaitu kosmetik perawatan kulit (*skin care*) yaitu kosmetik untuk memelihara, merawat dan mempertahankan kondisi kulit. Sedangkan yang kedua adalah kosmetik riasan (dekoratif atau

make up), yaitu jenis kosmetik yang digunakan untuk mempercantik dan menutup ketidaksempurnaan pada kulit wajah, sehingga menghasilkan penampilan yang lebih menarik.

Dengan beragam jenis *make up* dan *skincare* dari brand emina cosmetics, terkadang membuat para customer emina bingung dalam memilih jenis kosmetik yang diperlukan. Selama ini ketika customer memilih produk dilakukan dengan cara manual yaitu dengan melakukan searching di internet mengenai produk yang diperlukan. Untuk membantu customer mendapatkan lebih banyak pilihan produk yang sesuai, maka diperlukan adanya sebuah system rekomendasi yang dapat memberikan saran ataupun rekomendasi kepada para customer untuk memilih product Emina Cosmetics berdasarkan ketertarikan dan kebutuhan dalam pencarian referensi

Dalam membangun system rekomendasi untuk menentukan produk yang sesuai kebutuhan customer, terdapat beberapa metode yang sering digunakan yaitu *Collaborative Filtering*, *Content-Based Filtering*, dan *Hybrid*. *Collaborative filtering* menggunakan riwayat product yang disukai atau penilaian sebagai dasar untuk menentukan rekomendasi. Pada *content-based filtering* menggunakan kesamaan produk untuk ditawarkan kepada pembeli seperti judul atau deskripsi. Sedangkan metode *Hybrid* menggabungkan metode dua atau lebih metode untuk menghasilkan rekomendasi yang lebih baik. Pada penelitian ini, penulis menggunakan metode *content-based filtering* yang menggunakan ketersediaan konten sebuah item sebagai basis dalam pemberian rekomendasi (Ricci, 2011). *Content-based filtering* dipilih karena dalam proses rekomendasi tidak membutuhkan adanya penilaian pengguna lain (rating) sebagai dasar pemberian rekomendasi. Pada penelitian ini, parameter yang digunakan hanya deskripsi produk yang mirip dengan kata kunci yang dicari oleh customer. Berdasarkan item yang dicari oleh customer, system dapat merekomendasikan item baru yang memiliki kemiripan deskripsi konten dengan item lama. Jenis barang dibagi berdasarkan vektor komponen pembentuknya, misalnya kategori face make up, kategori lip make up, kategori *skincare* dll. Sistem rekomendasi akan memudahkan perusahaan dalam memberikan informasi produk yang sedang dicari maupun yang direkomendasikan oleh sistem kepada customer.

2. TINJAUAN PUSTAKA DAN TEORI

Dalam penelitian yang telah dilakukan oleh Parwita (2019) yang mengusung metode *content-*

based filtering salah satunya adalah penelitian yang menguji akurasi metode *content-based filtering* dalam dokumen UPP dan dokumen publikasi atau penelitian calon dosen pembimbing digunakan sebagai dasar penentuan rekomendasi. Penelitian ini meninjau pengaruh *stopword* dalam tingkat akurasi penentuan rekomendasi dokumen. Penelitian ini menyimpulkan bahwa system rekomendasi dengan proses *stopwordremoval* masih lebih unggul dibandingkan sistem rekomendasi tanpa proses *stopwordremoval*. [1]

Penelitian selanjutnya dilakukan oleh Fiarni C & Maharani (2019) yang merancang sebuah sistem rekomendasi untuk alternative produk kepada customer menggunakan metode *content-based filtering* berdasarkan karakteristik dan history transaksi menggunakan algoritma *cosine similarity*. Algoritma *cosine similarity* digunakan untuk menghitung nilai kesamaan suatu produk. Metode *content-based filtering* dapat menyediakan rekomendasi customer yang sesuai dengan karakteristik perusahaan dan transaksi pembelian yang terjadi.[2]

Penelitian yang dilakukan oleh badriyah & Syarif (2018) menggunakan metode *content-based filtering* algoritma apriori, dimana pola-pola kombinasi item yang dibeli akan dianalisis. Sistem dapat memberikan rekomendasi produk berdasarkan transaksi belanja yang pernah dilakukan oleh pembeli di e-commerce. Sistem dapat menampilkan produk rekomendasi untuk setiap user member yang aktif. Dan pada setiap detail produk yang ditampilkan, dapat direkomendasikan produk lain berdasarkan prosentasi kemunculan dari produk lain dari user yang sedang menampilkan detail produk pada saat itu. Semakin besar nilai support dan confidence yang diberikan maka akan semakin sedikit rekomendasi yang ditampilkan dan semakin kecil nilai support dan confidence yang diberikan maka akan semakin banyak rekomendasi yang ditampilkan.[3]

Penelitian selanjutnya dilakukan oleh Wijaya & Alfian (2018) yang menggunakan dua metode sekaligus yaitu *content-based filtering* dan *collaborative filtering* dengan tujuan memberikan saran ataupun rekomendasi laptop kepada customer berdasarkan ketertarikan dan kebutuhan customer. Algoritma yang digunakan untuk metode *collaborative filtering* yaitu Adjusted-cosine similarity untuk menghitung kemiripan antar customer, dan algoritma weighted sum untuk perhitungan prediksinya, sedangkan untuk metode *content based filtering* algoritma yang digunakan adalah tf-idf untuk pencarian ketersediaan konten yang ada. [4]

Kurniawati S.B (2018) melakukan penelitian mengenai STBI (Sistem Temu Kembali Informasi)

pencarian buku perpustakaan dengan metode *content-based filtering*, dimana pengunjung perpustakaan nantinya akan menginput kata kunci buku yang akan dicari, selanjutnya dengan metode *cosine similarity*, system akan menampilkan buku apa saja yang direkomendasikan berdasarkan kata kunci yang diinput pengunjung.[5]

2.1 Sistem Rekomendasi

Sistem rekomendasi adalah suatu alat atau teknik perangkat lunak yang digunakan untuk melakukan prediksi terhadap suatu objek. Sistem rekomendasi dapat memberikan saran-saran item yang bermanfaat bagi pengguna (Ricci, et al., 2011). Saran tersebut dapat digunakan untuk menentukan pengambilan keputusan oleh user, seperti menentukan item yang akan dibeli, menentukan musik yang akan didengarkan, film apa yang akan ditonton, berita mana yang akan dibaca, dsb.

Sistem rekomendasi memiliki 2 inputan yang berbeda, yaitu implicit input dan explicit input (Hu, et al., 2008). Pada implicit input, didapat dengan cara mengamati kebiasaan pengguna, seperti riwayat pemesanan, riwayat penelusuran, pola pencarian, dll. Sedangkan pada explicit input didapat dengan hasil penilaian yang diberikan oleh user, seperti pemberian rating, pemberian tanda favorite item, ataupun thumbs-up/down pada item tertentu

Klasifikasi sistem rekomendasi dapat dibagi menjadi beberapa tipe: content-based, collaborative-based, hybrid-based. Beberapa peneliti menambahkan metode knowledge based, seperti pada gambar 2.1 yang menunjukkan gambar klasifikasi sistem rekomendasi.



Gambar 2.1 Klasifikasi Sistem Rekomendasi

2.2. Collaborative Filtering

Collaborative Filtering adalah tipe sistem rekomendasi yang didasarkan pada pemberian rating oleh pengguna. Dalam hal ini, dicontohkan rekomendasi film pada sebuah website. Rating user akan merekomendasikan film yang belum pernah ditonton pengguna lain, tetapi pengguna yang menonton dan menyukai film serupa. Untuk menentukan apakah dua pengguna serupa atau

tidak, filter ini mempertimbangkan film yang ditonton keduanya dan bagaimana mereka memberikan rating. Dengan melihat kesamaan item, algoritma jenis ini pada dasarnya akan memprediksi laju film untuk pengguna yang belum menontonnya, berdasarkan pada tarif pengguna yang serupa. Agar dapat bekerja secara akurat, jenis filter ini perlu penilaian pengguna (rating), dan tidak semua pengguna menilai produk secara konstan. Beberapa dari mereka nyaris atau tidak pernah menilai apa pun. Karakteristik lain dari metode ini adalah keragaman dalam rekomendasi, yang dapat dikatakan baik atau buruk, tergantung pada kasusnya. Sebagai contoh, katakanlah pengguna A sangat menyukai film-film dystopian dan komedi gelap. Pengguna B juga menikmati film-film dystopian tetapi tidak pernah menonton komedi gelap. Collaborative Filtering akan merekomendasikan pertunjukan komedi gelap kepada pengguna B, berdasarkan pada selera yang dimiliki kedua pengguna untuk film dystopian. Skenario ini dapat memperoleh 2 hasil: pengguna B sangat menyukai komedi gelap, dalam hal ini berarti sistem rekomendasi berhasil. Atau, pengguna B benar-benar menikmati gaya komedi yang lebih ringan, dan dalam hal ini rekomendasinya belum berhasil.

2.3 Content Based Filtering

Berbeda dengan Collaborative Filtering, Content-Based Filtering tidak melibatkan pengguna lain dalam menentukan rekomendasi, namun hanya pengguna itu sendiri. Berdasarkan apa yang dicari user, algoritma ini hanya akan memilih item dengan konten yang mirip untuk direkomendasikan. Hal ini akan membuat keragaman rekomendasi lebih sedikit, tetapi akan berfungsi baik walaupun pengguna tidak memberikan penilaian (rating) Sebagai contoh, katakanlah pengguna A sangat menyukai film-film dystopian dan komedi gelap. Pengguna B juga menikmati film-film dystopian tetapi tidak pernah menonton komedi gelap. content-based filtering hanya akan terus merekomendasikan film dystopian atau sejenisnya. Tentu saja ada banyak kategori yang dihitung kesamaannya, contoh dalam film dapat dibangun sistem rekomendasi berdasarkan genre saja, atau mungkin sutradara, aktor utama, dsb.

Metode content-based filtering akan mengekstrak informasi yang terdapat pada item kemudian membandingkannya dengan informasi item yang pernah dilihat atau disukai oleh user. Metode tersebut biasa digunakan untuk merekomendasikan, artikel, berita, maupun situs web.

Teknik – teknik yang biasa digunakan dalam content-based filtering seperti TF-IDF, Bayesian Classifiers, Cluster analysis, decision trees dan artificial neural networks. Metode ini memiliki kelebihan yaitu dapat memberikan rekomendasi tanpa diperlukan adanya rating oleh customer, melainkan memberikan rekomendasi berdasarkan informasi item (dalam penelitian ini deskripsi produk), sedangkan kekurangan dari metode conten-based filtering yaitu system tidak dapat memberikan rekomendasi pada user yang belum pernah melakukan aktivitas apapun.

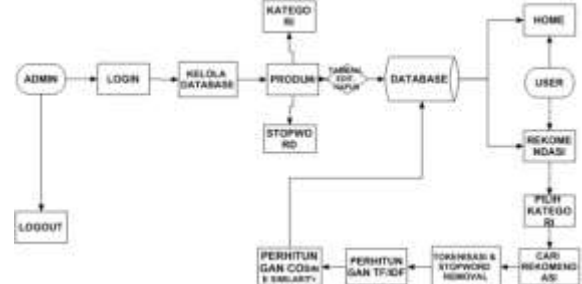
3. METODOLOGI PENELITIAN

Implementasi metode content- based filtering pada pemilihan produk cosmetic emina dapat memudahkan customer untuk mendapatkan saran produk berdasarkan pencarian customer.

Pengelolaan database hanya dapat dilakukan oleh admin, pada awalnya admin akan melakukan login ke dalam aplikasi terlebih dahulu. Pada menu admin, terdapat beberapa halaman yang akan dapat diakses yaitu menu kategori, kosmetik, dan stopword. Pada menu kategori, admin dapat mengelola data kategori seperti menambah kategori, menghapus kategori, dan mengedit kategori yang sudah diinput sebelumnya. Kategori nantinya akan dipilih oleh user sebelum mencari rekomendasi produk. Selanjutnya terdapat menu produk. Pada menu kosmetik berisi seluruh produk yang tampil pada menu home user. Seperti pada menu kategori, admin dapat melakukan penambahan, menghapus, dan mengedit produk. Apabila admin ingin menambahkan produk admin harus mengisi kolom nama produk, kategori, harga, deskripsi produk, dan upload foto. Menu selanjutnya adalah stopword. Admin dapat mengelola data stopword. Stopword adalah kata hubung yang nantinya akan dihilangkan melalui proses *stopword removal*. Untuk keluar dari menu, admin dapat melakukan *logout*.

Pada menu customer terdapat 2 menu yaitu home dan rekomendasi. Menu home menampilkan semua produk yang terdapat pada aplikasi. Pada awalnya, customer akan memilih kategori terlebih dahulu. Kategori diinput oleh admin dengan klasifikasi data tertentu. Selanjutnya customer menginput kata kunci yang diinginkan. Apabila data valid maka sistem akan menyimpan data kata kunci dalam tabel term, kunci, temp, term, dan data rekomendasi produk

akan muncul di interface rekomendasi. Data valid disini adalah apabila produk yang kata kunci nya yang diinput oleh customer tersedia dalam system. Sebagai contoh, customer menginputkan kata kunci “krim pemutih wajah” dalam halaman rekomendasi. Kata kunci tersebut tidak akan menghasilkan rekomendasi karena tidak tersedia produk dengan deskripsi “krim pemutih wajah” dalam system rekomendasi emina cosmetics ini.



Gambar 3.1 Diagram Sistem Rekomendasi

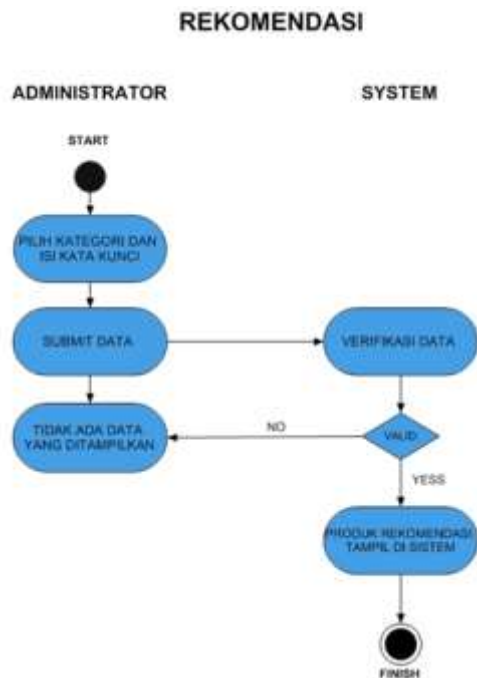
3.1. Pengumpulan Data

Objek dalam penelitian ini adalah Brand Emina Cosmetics yang saat penelitian ini dibuat memiliki 87 varian produk yang berbeda.

Penulis melakukan wawancara langsung dengan assistant trainer di PT Paragon Technology and Innovation, dimana assistant trainer yang memiliki data product knowledge brand kosmetik emina yang merupakan data yang diolah dalam system rekomendasi.

Data produk selanjutnya di klasifikasikan sesuai dengan jenis produk diantaranya adalah :

1. Face Make Up
Berisi produk make up wajah
2. Lip Make Up
Berisi produk make up bibir
3. Eye Make Up
Berisi produk make up mata
4. Cheek Make Up
Berisi produk make up pipi
5. Body Care
Berisi produk perawatan tubuh
6. Lip Care
Berisi produk perawatan bibir
7. Skin care
Berisi produk perawatan wajah
8. Nail
Berisi produk untuk kuku
9. Other
Berisi produk / tools kecantikan



Gambar 3.2 Activity Diagram Rekomendasi

Proses rekomendasi dengan metode content-based filtering pada penelitian ini menggunakan algoritma cosine similarity dilakukan dengan tahapan sebagai berikut :

1. Customer menginput kata kunci pada kolom cari sesuai dengan produk yang akan dicari
2. System akan membaca teks (deskripsi produk) secara baris per baris. Pada proses membaca teks system melakukan proses tokenisasi, yaitu membagi teks yang dapat berupa kalimat, paragraf atau dokumen, menjadi token-token/bagian-bagian tertentu, dan pada waktu yang bersamaan dilakukan juga proses penghapusan karakter tertentu, seperti tanda baca, tagtag html. Sebagai contoh kata “ Emina cc cream dapat menyamarkan pori dan garis halus pada wajah. ” menghasilkan 11 token yaitu “emina”, “cc”, “cream”, “dapat”, “menyamarkan”, “pori”, “dan”, “garis”, “halus”, “pada”, “wajah”
3. Stopword removal, yaitu menghilangkan kata hubung yang kurang penting sebagai contoh : dan, di, ke, dari, yang, dll.
4. Menentukan bobot setiap term dari produk yang mengandung kata kunci. Bobot dalam dokumen dihitung dengan menggunakan tf-idf. TF- IDF dikenal efisien, mudah dan memiliki hasil yang akurat (Robertson, 2006:45). Metode ini menghitung nilai Term Frequency (TF) dan Inverse Document

Frequency (IDF) pada setiap token (kata) di setiap dokumen dalam korpus. Frekuensi kemunculan sebuah kata dalam dokumen tertentu dan inverse frekuensi dokumen yang mengandung kata tersebut digabungkan untuk menghitung bobot. Metode ini akan menghitung bobot setiap token t di dokumen d dengan rumus berikut

$$IDF = (D/DF)$$

$$W = TF * (IDF + 1)$$

Keterangan :

TF : jumlah kemunculan kata atau term dalam dokumen

IDF = inverse document frequency

D = jumlah semua dokumen

DF = jumlah dokumen yang mengandung kata (term)

W : bobot setiap dokumen

Setelah bobot (W) masing-masing dokumen diketahui, maka dilakukan proses pengurutan dimana semakin besar nilai W, semakin besar tingkat similaritas dokumen tersebut terhadap kata kunci, demikian sebaliknya.

5. Hitung kemiripan vektor [kosmetik] query Q dengan setiap dokumen yang ada. Kemiripan antar kosmetik dihitung dengan algoritma cosine similarity menggunakan rumus similarity :

$$\cos \alpha = \frac{A \cdot B}{|A| \cdot |B|} = \frac{\sum_{i=1}^n A_i \cdot B_i}{\sqrt{\sum_{i=1}^n A_i^2} \cdot \sqrt{\sum_{i=1}^n B_i^2}}$$

A = Vektor A, yang akan dibandingkan kemiripannya

B = Vektor B, yang akan dibandingkan kemiripannya

A • B = dot product antara vektor A dan vektor B

|A| = panjang vektor A |B| = panjang vektor B

|A||B| = cross product antara |A| dan |B|

4. HASIL DAN PEMBAHASAN

Proses rekomendasi emina cosmetics dilakukan dengan tahapan text preprocessing, pembobotan TF-IDF, dan perhitungan cosine similarity, Pada penelitian kali ini, akan dihitung nilai produk yang akan direkomendasikan apabila customer memilih kata kunci “loose powder untuk kulit berminyak”:

1. Text Preprocessing

Lakukan text preprocessing pada deksripsi produk dengan tujuan agar data yang dipakai dapat diproses menjadi angka dengan TF-IDF dan *cosine similarity*. Tahapan dalam proses ini adalah tokenisasi dan *stopword* removal. Preprocessing dilakukan pada item yang mengandung kata kunci saja sehingga proses pencarian tidak memakan waktu yang lama, dimana deskripsi produk yang mengandung kata kunci pada penelitian ini ada 10 dokumen . Hasil dari preprocessing dapat dilihat dalam tabel 4.1

Tabel 4.1 Hasil Preprocessing

Doc	Kode	Nama Produk	Setelah Preprocessing
Q	-	Loose powder untuk kulit berjerawat	loose powder kulit jerawat
D1	S0001	Bare with me mineral compact powder 01 fair, 14 gr	complexion powder wajah halus bebas kilap seharian warna kulit merata hasil ringan bebas kilap memiliki varian warna fair light beige amber ebony cocok tampil natural memiliki kandungan soft focus agent kesan halus membantu menyerap minyak praktis dibawa bentuk compact
D2	S0002	Emina bare with me mineral loose powder 01 fair 8 g	complexion powder wajah halus bebas kilap seharian warna kulit merata memiliki tekstur ringan halus tahan memiliki 4 pilihan warna fair light beige amber ebony cocok tampil natural memiliki kandungan soft focus agent kesan halus membantu menyerap minyak
D3	S0003	Emina sebum	complexion powder

		fighter loose powder 8 g	mengandung partikel sebum absorbent menyerap minyak berlebih cocok kulit berjerawat rentan berjerawat bedak translucent berwarna less pigmen cocok warna kulit setting powder ringan kulit bebas kilap shine free
D4	S0004	Emina bright stuff loose powder 55 g	complexion powder mencerahkan tampilan wajah kulit wajah halus micro smooth particle dilengkapi glitter tampilan wajah glowing natural matte finish tekstur ringan nyaman mudah dibaurkan efek mencerahkan natural kulit wajah tingkat coverage diatur sheer to full tahan
D5	S0005	Beauty bliss bb cream light, 20 ml	primer makeup wajah dewy halus kelembaban menutupi kemerahan kulit wajah pelembab foundation ringan dipakai sehari-hari diperkaya vitamin c vitamin e menjadikan kulit halus bercahaya mengandung spf 32 melindungi kulit wajah sinar matahari warna light
D6	S0006	City chic cc cake peach, 12 gr	complexion powder wajah halus bebas kilap seharian warna kulit merata membantu memperbaiki tekstur kulit bertahap kombinasi unik cc

			cream powder hasil halus diperkaya apricot antioksidan membantu rehidrasi kulit mudah diaplikasikan longlasting ringan mudah dibaur tahan bebas kilap
D7	S0007	Emina city chic cc cream natural 20 ml	hadir color changing technology mengubah warna krim putih warna sesuai warna kulit shade pilihan light natural menyamarkan pori pori garis halus kerutan wajah mengandung vitamin c vitamin e berperan antioksidan menjaga sel bahaya radikal bebas menjadikan wajah halus mengandung ekstrak aloe vera allantoin agen pelembab menghidrasi kulit
D8	S0009	Emina bare with me mineral cushion 01 light 15 g	dewy matte finish tampilan matte lembab efek kulit sehat segar kandungan oil absorber kulit berminyak efek 1 tingkat cerah perlindungan sinar uva uvb tekstur ringan tahan buildable coverage
D9	S0010	Emina refill bare with me mineral cushion 01 light 15 g	refill cushion dewy matte finish tampilan matte lembab efek kulit sehat segar kandungan oil absorber kulit berminyak efek 1 tingkat cerah perlindungan

			sinar uva uvb tekstur ringan tahan buildable coverage
D10	S0095	Emina daily matte loose powder 01 light beige 20 g	complexion powder daily matte loose powder ringan bare with me cocok sehari memiliki coverage rendah bare with me selling pointnya tampilan matte bebas kilap ringan sasaran daily matte customer muda

2. Pembobotan TF-IDF

Pembobotan dilakukan pada deskripsi produk yang mengandung kata kunci, setiap dokumen yang mengandung term diberi nilai 1. Pembobotan setiap term dapat dilihat pada tabel 4.2.

Tabel 4.2 Bobot Term

term	TF			
	Loose	Powder	Kulit	Jerawat
Q	1	1	1	1
D1		1	1	
D2		1	1	
D3		1	1	1
D4		1	1	
D5			1	
D6		1	1	
D7			1	
D8			1	
D9			1	
D10	1	1		

Nilai DF merupakan jumlah dari term yang ditemukan pada setiap dokumen. Hitung nilai IDF dengan rumus $IDF = \log(n/DF)$. Hasil

perhitungan idf seperti yang ditampilkan pada tabel 4.3.

Tabel 4.3 Hasil Perhitungan idf

term	DF	D/DF	IDF
Loose	2	5,5	0,7404
Powder	7	1,5714	0,1963
Kulit	10	1,1	0,0414
Jerawat	2	5,5	0,7404

Lakukan perhitungan Wdt dengan rumus $Wdt = tf \cdot idf$, dan didapatkan hasil seperti pada tabel 4.4.

Tabel 4.4 Perhitungan Wdt

term	Wdt=TF.IDF			
	Loose	Powder	Kulit	Jerawat
Q	0,7404	0,1963	0,0414	0,7404
D1	0,0000	0,1963	0,0414	0,0000
D2	0,0000	0,1963	0,0414	0,0000
D3	0,0000	0,1963	0,0414	0,7404
D4	0,0000	0,1963	0,0414	0,0000
D5	0,0000	0,0000	0,0414	0,0000
D6	0,0000	0,1963	0,0414	0,0000
D7	0,0000	0,0000	0,0414	0,0000
D8	0,0000	0,0000	0,0414	0,0000
D9	0,0000	0,0000	0,0414	0,0000
D10	0,7404	0,1963	0,0000	0,0000

3. Perhitungan cosine similarity

Kemiripan vector [kosmetik] query q dihitung dengan setiap dokumen yang ada menggunakan rumus cosine similarity. Hitung hasil perkalian skalar antara Q dan 10 dokumen lain. Hasilnya perkalian dari setiap dokumen dengan Q dijumlahkan dan hasilnya diperlihatkan seperti pada tabel 4.5

Tabel 4.5 Perkalian Skalar

term	WD*Wdi				Total
D1	0,0000	0,0385	0,0017	0,0000	0,0402
D2	0,0000	0,0385	0,0017	0,0000	0,0402
D3	0,0000	0,0385	0,0017	0,5481	0,5883
D4	0,0000	0,0385	0,0017	0,0000	0,0402
D5	0,0000	0,0000	0,0017	0,0000	0,0017
D6	0,0000	0,0385	0,0017	0,0000	0,0402
D7	0,0000	0,0000	0,0017	0,0000	0,0017
D8	0,0000	0,0000	0,0017	0,0000	0,0017
D9	0,0000	0,0000	0,0017	0,0000	0,0017
D10	0,5481	0,0385	0,0000	0,0000	0,5866

Hitung panjang vector setiap dokumen, termasuk Q dengan mengakarkan penjumlahan wdt yang berada pada kolom total. Hasil perhitungan panjang vector dapat dilihat dalam tabel 4.6.

Tabel 4.6 Perhitungan Panjang Vektor

term	Panjang Vektor				Total	Akar
Q	0,5481	0,5481	0,5481	0,5481	1,1364	1,0660
D1	0,0000	0,0385	0,0017	0,0000	0,0402	1,0660
D2	0,0000	0,0385	0,0017	0,0000	0,0402	1,0660
D3	0,0000	0,0385	0,0017	0,5481	0,5883	1,0660
D4	0,0000	0,0385	0,0017	0,0000	0,0402	1,0660
D5	0,0000	0,0000	0,0017	0,0000	0,0017	1,0660
D6	0,0000	0,0385	0,0017	0,0000	0,0402	1,0660
D7	0,0000	0,0000	0,0017	0,0000	0,0017	1,0660
D8	0,0000	0,0000	0,0017	0,0000	0,0017	1,0660
D9	0,0000	0,0000	0,0017	0,0000	0,0017	1,0660
D10	0,5481	0,0385	0,0000	0,0000	0,5866	1,0660

Terapkan rumus cosine similarity dengan menghitung kemiripan antar dokumen D1 sampai D10. Contoh perhitungan hanya dituliskan dalam D1, D3, dan D5 karena pada dokumen yang lain yaitu D2, D4, D6, D7, D8, D9, dan D10 memiliki nilai perhitungannya yang

sama dengan dokumen yang dicontohkan dibawah ini.

$$\begin{aligned} \text{Similarity (Q,D1)} &= 0,0402 / (1,0660 * 0,2005) \\ &= 0,0402 / 0,2137 \\ &= 0,1881 \end{aligned}$$

$$\begin{aligned} \text{Similarity (Q,D3)} &= 0,5883 / (1,0660 * 0,7670) \\ &= 0,5883 / 0,8176 \\ &= 0,7195 \end{aligned}$$

$$\begin{aligned} \text{Similarity (Q,D5)} &= 0,0017 / (1,0660 * 0,0412) \\ &= 0,0017 / 0,0439 \\ &= 0,0387 \end{aligned}$$

Urutan produk yang menjadi rekomendasi dari pencarian dengan kata kunci “loose powder untuk kulit berminyak” dapat dilihat pada tabel 4.7. Produk yang memiliki nilai similaritas yang sama diurutkan berdasarkan abjad. Implementasi pada sistem menghasilkan interface seperti pada gambar 4.1.

Tabel 4.7 Hasil Rekomendasi

Doc	Nama Produk	Nilai Cosine	Ranking
D3	Emina sebum fighter loose powder 8 g	0,7195	1
D10	Emina daily matte loose powder 01 light beige 20 g	0,7195	2
D1	Bare with me mineral compact powder 01 fair, 14 gr	0,1881	3
D6	City chic cc cake peach, 12 gr	0,1881	4
D2	Emina bare with me mineral loose powder 01 fair 8 g	0,1881	5
D4	Emina bright stuff loose powder 55 g	0,1881	6
D5	Beauty bliss bb cream light, 20 ml	0,0387	7
D8	Emina bare with me mineral cushion 01 light 15 g	0,0387	8
D7	Emina city chic cc cream natural 20 ml	0,0387	9
D9	Emina refill bare with me mineral cushion 01 light 15 g	0,0387	10



Gambar 4.1 Hasil Rekomendasi

5. Kesimpulan dan Saran

Metode *content based filtering* menggunakan kesamaan produk untuk ditawarkan kepada pembeli seperti judul atau deskripsi produk, metode ini tidak memerlukan parameter semacam rating untuk menghasilkan suatu rekomendasi. Algoritma *cosine similarity* cocok dipergunakan pada data yang tidak terdapat nilai subjektif seperti similaritas antara teks berdasarkan kemiripan kata dalam teks. Dalam penelitian sistem rekomendasi emina cosmetics dengan metode *content-based filtering* menghasilkan 10 produk rekomendasi dengan hasil akhir perhitungan cosine tertinggi dengan nilai similaritas 0,7195.

Untuk penelitian selanjutnya, produk yang disediakan dalam sistem rekomendasi dapat ditambah menjadi lebih banyak sehingga mendukung hasil rekomendasi produk yang akurat dan semakin beragam serta dapat dikembangkan menjadi platform mobile android.

Daftar Pustaka:

- [1] Parwita, W. G. S. 2019. Pengujian Akurasi Sistem Rekomendasi Berbasis Content-Based Filtering. *Informatika Mulawarman: Jurnal Ilmiah Ilmu Komputer*, 14(1), 27-32.
- [2] Fiarni, C., & Maharani, H. 2019. Product Recommendation System Design Using Cosine Similarity and Content-based Filtering Methods. *IJITEE (International Journal of Information Technology and Electrical Engineering)*, 3(2), 42-48.
- [3] Badriyah, T., Fernando, R., & Syarif, I. 2018. Sistem Rekomendasi Content Based Filtering Menggunakan Algoritma Apriori. *Konferensi Nasional Sistem Informasi (KNSI) 2018*.
- [4] Wijaya, A., & Alfian, D. 2018. Sistem Rekomendasi Laptop Menggunakan Collaborative Filtering Dan Content-Based

- Filtering. *Jurnal Computech & Bisnis*, 12(1), 11-27
- [5] Kurniawati, S. B. 2018. Sistem Temu Kembali Informasi Pencarian Buku Perpustakaan dengan Metode Cosine Similarity (Studi Kasus di Perpustakaan Otoritas Jasa Keuangan Kantor Regional 3 Jawa Tengah dan DIY), Skripsi, Program Studi Teknik Informatika FTI Unisbank, Semarang.
- [6] Harun, R., Pelangi, K. C., & Lasena, Y. 2020. PENERAPAN DATA MINING UNTUK MENENTUKAN POTENSI HUJAN HARIAN DENGAN MENGGUNAKAN ALGORITMA K NEAREST NEIGHBOR (KNN). *Jurnal Manajemen Informatika dan Sistem Informasi*, 3(1), 8-15.
- [7] L. Mutawalli, M. T. A. Zaen, and I. F. Suhriani, "Sistem Identifikasi Persebaran Pecemaran Air Oleh Limbah di Indonesia Menggunakan Average Linkage Dan K-Mean Cluster," *MISI (Jurnal Manaj. Inform. Sist. Informasi)*, vol. 1, no. 2, pp. 36-42, 2018
- [8] M. Nawawi, M. T. A. Zaen, and M. F. Zulkarnaen, "Implementasi Metode Analytic Hierarchy Process (AHP) Untuk Penentuan Penerima Bantuan Kube di Dinas Sosial Lombok Tengah," *MISI (Jurnal Manaj. Inform. Sist. Informasi)*, vol. 2, no. 1, pp. 1-8, 2019.
- [9] Imtihan, K., & Fahmi, H. (2020). ANALISIS DAN PERANCANGAN SISTEM INFORMASI DAERAH RAWAN KECELAKAAN DENGAN MENGGUNAKAN GEOGRAPHIC INFORMATION SYSTEMS (GIS). *Jurnal Manajemen Informatika dan Sistem Informasi*, 3(1), 16-23.
- [10] Sa'adati, Y., Fadli, S., & Imtihan, K. (2018). Analisis Penggunaan Metode AHP dan MOORA untuk Menentukan Guru Berprestasi Sebagai Ajang Promosi Jabatan. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, 3(1), 82-90.
- [11] Imtihan, K., Hadawiyah, R., & Lombok, H. A. S. (2018). Sistem Informasi Penggajian Guru Honorer Menggunakan Konsep Agile Software Development dengan Metodologi Extreme Programming (XP) pada SMK Bangun Bangsa. *IJNS-Indonesian Journal on Networking and Security*, 7(2).